18 CONGRESO COLOMBIANO DE COMPUTACIÓN

MANIZALES, *Del 4 al 6 de Septiembre 2024*

MEMORIAS

Departamento de Informática y Computación
Facultad de Administración
Sede Manizales

Organiza

Sociedad Colombiana de Computación

UNIVERSIDAD NACIONAL DE COLOMBIA

Sociedad Colombiana de Computación

18 Congreso Colombiano de Computación

Memorias
Artículos Cortos
Simposio de Maestria y Doctorado y
Pósteres

Septiembre 4-6

2024

Catalogación en la fuente Universidad Nacional de Colombia.

Incluye referencias bibliográficas al final de cada artículo.

Incluye nombre de los autores despues del título de cada artículo.

ISBN:

Memorias Artículos Cortos, Simposio de Maestría y Doctorado, Pósteres.
18 Congreso Colombiano de Computación - 18CCC

Ilustraciones a color
1. Computación de alto rendimiento, 2. Computación biológica, 3. Computación cuántica 4. Industria 4.0,Transformación digital, 5. Inteligencia artificial, Robótica, 6. Ciberseguridad y seguridad de la información, 7. Ingeniería de software y Automatización, 8. Arquitecturas de Tecnologías de la Información, 9. Datos, 10. información y conocimiento, 11. Interacción Humano-computador, Reconocimiento de patrones y visión por computador, 12. Realidad virtual, aumentada y mezclada, 13. Educación y TIC – e-Learning, 14. Gestión del conocimiento y de la información, 15. Métodos formales en sistemas de computación, 16. Estadística computacional

DD /2024

Editors

Memorias Artículos Cortos, Simposio de Maestría, Doctorado y Pósteres.
18 Congreso Colombiano de Computación
Manizales, Caldas, Colombia.

Ana-Lorena Uribe-Hurtado [ORCID]
Universidad Nacional de Colombia Sede Manizales
Manizales, Caldas, Colombia

Nestor Dario Duque Mendez[ORCID]
Universidad Nacional de Colombia Sede Manizales
Manizales, Caldas, Colombia

Marco Javier Suarez[ORCID]
Universidad Pedagógica y Tecnológica de Colombia
Tunja, Boyacá, Colombia

Dedicamos este libro a todos los científicos, profesores e investigadores, como también a los estudiantes de doctorado, maestría y pregrado que, con sus estudios, generan nuevo conocimiento y buscan soluciones a problemas relevantes apoyados en las ciencias de la informática y la computación. Su esfuerzo contribuye al progreso no solo de nuestro país, sino también del mundo entero.

We dedicate this book to all scientists, professors and researchers, as well as to PhD, Master's and undergraduate students who, through their studies, generate new knowledge and seek solutions to relevant problems supported by computer science and computing. Their efforts contribute to the progress not only of our country, but also of the entire world.

Dedicamos este livro a todos os cientistas, professores e pesquisadores, bem como aos alunos de doutorado, mestrado e graduação que, com seus estudos, geram novos conhecimentos e buscam soluções para problemas relevantes apoiados na informática e nas ciências da computação. Os seus esforços contribuem para o progresso não só do nosso país, mas também do mundo inteiro.

Preface

In the digital age, data science, mining, and artificial intelligence are not only scientific disciplines but also transformative forces that impact our society. Their influence extends beyond laboratories and mathematical equations, manifesting in innovative solutions and decisions based on concrete results. The Master's and Doctoral Symposium (SMD2024), within the framework of the 18th Colombian Congress of Computing, brings together postgraduate students from Colombia and around the world. At this event, they present progress on their thesis proposals and partial results in scientific articles. Each project emphasizes or proposes a contribution to solving specific problems in various contexts. But this event goes beyond that: it is a call for collaboration, an opportunity to build a community of researchers passionate about Informatics and Computer Science. Together, we chart a path toward a smarter, more inclusive, and promising future for our nation and the world.

Manizales, Colombia
September 2024

Ana lorena Uribe
Marco Javier Suárez Barón

Prefacio

En la era digital, las ciencias de datos, la minería y la inteligencia artificial no solo son disciplinas científicas, sino también fuerzas transformadoras que afectan nuestra sociedad. Su impacto va más allá de los laboratorios y las ecuaciones matemáticas, manifestándose en soluciones innovadoras y decisiones basadas en resultados concretos. El Simposio de Maestría y Doctorado (SMD2024), en el marco del 18 Congreso Colombiano de Computación, reúne a estudiantes de postgrado de Colombia y el mundo. En este evento. Ellos presentan avances de sus propuestas de tesis y resultados parciales en artículos científicos. Cada proyecto enfatiza o propone una contribución a la solución de problemáticas concretas en diversos contextos. Pero este evento va más allá: es un llamado a la colaboración, una oportunidad para construir una comunidad de investigadores apasionados por la Informática y la Computación. Juntos, trazamos un camino hacia un futuro más inteligente, más inclusivo y más prometedor para nuestra nación y el mundo.

Manizales, Colombia
September 2024

Ana lorena Uribe
Marco Javier Suarez

Acknowledgements

En la realización de este libro, queremos expresar nuestra gratitud a la Sociedad Colombiana de Computación y al grupo de profesores colaboradores de las universidades del pais, por su apoyo constante y su compromiso con la investigación y la educación de Colombia en el campo de la informática y la computación a través del Congreso Colombiano de Computación hoy en su 18^a *version. Su dedicación ha sido fundamental para el éxito del SMD2024.*

In the creation of this book, we want to express our gratitude to the Colombian Society of Computing and the group of collaborating professors from universities across the country for their constant support and commitment to research and education in Colombia's field of informatics and computer science through the 18th Colombian Congress of Computing. Their dedication has been crucial to the success of SMD2024

Na criação deste livro, queremos expressar nossa gratidão à Sociedade Colombiana de Computação e ao grupo de professores colaboradores de universidades de todo o país por seu constante apoio e compromisso com a pesquisa e a educação na área de informática e ciência da computação da Colômbia através do 18º Congresso Colombiano de Computação. A sua dedicação foi crucial para o sucesso do SMD2024

Contents

Part I
Artículos Cortos

En esta sección, se presentan investigaciones recientes que abordan diferentes temas de las áreas de informática y computación. Los trabajos incluidos exploran desde innovaciones técnicas hasta aplicaciones prácticas en contextos reales, ofreciendo una visión integral de los desafíos y oportunidades en el reconocimiento de patrones, la inteligencia artificial, y su impacto en el mundo real. Estas contribuciones no solo amplían el conocimiento académico en estas áreas, sino que también sientan las bases para futuras investigaciones y aplicaciones en la vida cotidiana.

In this section, recent research addressing various topics in the fields of computer science and computing is presented. The included works explore everything from technical innovations to practical applications in real-world contexts, offering a comprehensive view of the challenges and opportunities in pattern recognition, artificial intelligence, and their real-world impact. These contributions not only expand academic knowledge in these areas but also lay the groundwork for future research and applications in everyday life.

Nesta seção, são apresentadas pesquisas recentes que abordam diferentes tópicos nas áreas de informática e computação. Os trabalhos incluídos exploram desde inovações técnicas até aplicações práticas em contextos reais, oferecendo uma visão abrangente dos desafios e oportunidades no reconhecimento de padrões, na inteligência artificial e em seu impacto no mundo real. Essas contribuições não apenas ampliam o conhecimento acadêmico nessas áreas, mas também estabelecem as bases para futuras pesquisas e aplicações na vida cotidiana.

1
Optimization of personalized learning paths with Deep Q-Learning

Manuel F. Caro[1][0000-0001-6594-2187]
Juan C. Giraldo[2][0000-0002-6327-0230]
Claudia C. Lengua-Cantero[3][0000-0001-8100-3016]

Abstract This study explores the optimization of personalized learning paths using the Deep Q-Learning (DQN) algorithm in an educational context. An Advisor module integrated into an intelligent tutoring system leverages artificial intelligence techniques to adapt learning paces and provide individualized recommendations. The DQN algorithm models student progress and predicts action values to guide the learning process. Intelligent tutoring systems are examined for their ability to adapt to individual student needs, emphasizing an experiential learning approach to promote active participation. The methodology involves designing and implementing a DQN model tailored for educational con- texts, using synthetic student profiles to simulate real-world scenarios. The evaluation of the algorithm reveals effective action predictions and high adaptability to various student profiles. Consistent improvement in rewards over time indicates increasingly optimal decision-making outcomes. Analysis of learning pathways demonstrates significant progress and effective stabilization in the learning process. This study contributes to the growing field of AI in education by presenting a DQN-based model capable of learning effective strategies for selecting educational activities based on a student's current knowledge state, providing empirical evidence of the model's ability to improve its performance over time, and demonstrating the model's capacity to generate distinct, appropriate learning paths for different student profiles, from beginners to advanced learners.

Key words: Personalized education, Learning path, Deep Q-Learning, Intelligent tutor system, AI in education.

[1,2] Universidad de Córdoba, Montería, Colombia
e-mail: \{manuelcaro,jgiraldo\}@correo.unicordoba.edu.co
[3] CECAR, Sincelejo, Colombia
e-mail: claudia.lengua@cecar.edu.co

1.1 Introduction

The continued advancement in the development of artificial intelligence (AI) and machine learning technologies has opened up new possibilities in personalized education. Intelligent tutoring systems (ITS) have emerged as AI tools to adapt learning experiences to individual student needs [1, 2]. Among the various AI techniques, reinforcement learning, particularly Deep Q-Learning (DQN), has shown promising results in creating adaptive systems across diverse domains [3, 4] This study explores the application of DQN to optimize personalized learning paths within an educational context. A learning path is the route chosen by a student through a variety of learning activities, allowing him or her to progressively develop knowledge [5]. By integrating a DQN-based Advisor module into an intelligent tutoring system called SGAMEs21 [6, 7], we aim to integrate AI functionalities to adapt learning paces and provide individualized recommendations tailored to each student's unique learning profile.

The potential of DQN in education consists of its ability to model complex decision-making processes and learn optimal strategies through interaction with an environmentv(see Section 2). In our context, the DQN algorithm models student progress and predicts action values to guide the learning process, potentially leading to more efficient and effective educational outcomes. Our research is grounded in the principles of experiential learning [8], which emphasizes active participation and personalized experiences in the learning process [9]. The methodology employed in this study involves designing and implementing a DQN model specifically tailored for educational contexts. To simulate real-world scenarios, we utilize synthetic student profiles, allowing us to test and refine our algorithm under controlled conditions. Our experiments demonstrate that the Deep Q-Learning approach is effective in creating personalized learning paths that adapt to individual student needs. The algorithm implementation shows consistent improvement in its ability to select appropriate learning activities, resulting in more efficient knowledge acquisition for students. The key contributions of this work include:

1. A DQN-based model capable of learning effective strategies for selecting educational activities based on a student's current knowledge state.
2. Empirical evidence of the model's ability to improve its performance over time, as shown by the learning curve and knowledge gain results.
3. Demonstration of the model's capacity to generate distinct, appropriate learning paths for different student profiles, from beginners to advanced learners.

This paper is structured as follows: Section 1.2 presents a formal model for personalized learning paths using Deep Q-Learning. Section 1.3 describes our validation and experimentation process, including the experimental setup and results analysis. Finally, Section 1.4 concludes the study, discussing the implications of our findings and outlining directions for future research in this field.

1.2 Formal model for personalized learning paths using Deep Q-Learning

This section presents a formal model for personalizing learning paths using deep Q-learning. The model integrates concepts from reinforcement learning, neural networks, and educational theory to create an adaptive system that optimizes learning trajectories for individual students.

1.2.1 State and action spaces

We define the state space S as the set of all possible combinations of a student's current knowledge and skills. Each state $s \in S$ represents a unique configuration of the student's educational status. The action space A comprises all possible learning activities or content recommendations available in the system. Each action $a \in A$ represents a specific educational intervention or task that can be assigned to the student.

1.2.2 Q-Function and Neural Network approximation

The core of our model is the Q-function, $Q : S \times A \longrightarrow \mathbb{R}$ which estimates the expected cumulative reward of taking action a in state s. Given the high dimensionality of the state and action spaces in educational contexts, we approximate the Q-function using a deep neural network:

$$f_\theta : S \longrightarrow \mathbb{R}^{|A|} \tag{1.1}$$

where θ represents the parameters of the network, and $f_\theta(s)_a \approx Q(s, a)$.

1.2.3 Experience replay

To stabilize learning and improve sample efficiency, we employ experience replay. Experiences are stored in a replay memory D:

$$D = \{(st, at, rt, s_{t+1}) \,|t = 1, \ldots, T\} \tag{1.2}$$

where s_t is the state at time t, a_t is the action taken, r_t is the reward received, and st+1 is the resulting state.

1.2.4 Exploration-Exploitation strategy

We balance exploration and exploitation using an ϵ-greedy policy:

$$\pi_{\in}(a|s) = \begin{cases} 1- \in + \frac{\in}{|A|}, \text{ if } a = arg\max_{a'} f_\theta(s)_{a'} \\ \frac{\in}{|A|} \quad otherwise \end{cases} \tag{1.3}$$

This policy ensures that the system explores new learning activities while still exploiting known effective strategies.

1.2.5 Learning update

The neural network is trained by minimizing the loss function:

$$L(\theta) = \mathbb{E}_{(s,a,r,s')} \backsim D \left[(r + \gamma \max_{a'} f_\theta - (s')_{a'} - f_\theta(s)_a)^2 \right] \tag{1.4}$$

where θ^- are the parameters of a target network, updated periodically to stabilize training, and γ is the discount factor.

1.2.6 Recommendation system

The recommendation system R is defined as:

$$R : S \times G_s \times AP_s \times H_s \longrightarrow P_s \tag{1.5}$$

where G_s is the set of student goals, AP_s is the set of student preferences,H_s is the student's historical performance, and P_s is the personalized learning path.

1.2.7 Reinforcement learning agent

The reinforcement learning agent Q is defined similarly:

$$Q : S \times G_s \times P_s \times H_s \longrightarrow P_s \tag{1.6}$$

This agent refines the recommendations based on the student's progress and the learned Q-function.

1.2.8 Reward function

The reward function combines cognitive and environmental components:

$$R(s,a) = CC_a + \lambda EI_a \tag{1.7}$$

where CC_a is the cognitive component, EI_a is the environmental impact, and λ is a weighting factor.

1.2.9 Optimal Q-Function and policy

The optimal Q-function is defined as:

$$Q^*(s,a) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t R(s_t,a_t)|s_0 = s, a_0 = a, \pi^*\right] \tag{1.8}$$

From this, we derive the optimal policy:

$$\pi^*(s) = arg\max_a Q_{(s,a)} \tag{1.9}$$

1.2.10 Deep Q-Learning algorithm design and implementation

The Deep Q-Learning algorithm iteratively refines the Q-function approximation, leading to increasingly optimal personalized learning paths for each student. In this way, the Algorithm 1 for personalized learning paths integrates several key components to create an adaptive educational system. At its core, the algorithm utilizes a neural network to approximate the Q-function, which estimates the value of taking specific actions (learning activities) in various states (student knowledge configurations).

The algorithm's most critical aspects include: (1) the use of experience replay, storing and randomly sampling past experiences $(s_t, a_t, r_t, s_{(t+1)})$ to stabilize learning and improve sample efficiency; (2) the implementation of an ϵ-greedy policy to balance exploration of new activities with exploitation of known effective strategies; (3) the employment of a target network, updated periodically, to stabilize the training process; and (4) the iterative update of the Q-network parameters θ using gradient descent on the loss function. This approach allows the system to continuously refine its understanding of optimal learning paths, adapting to individual student needs and progressively optimizing their educational trajectories. The algorithm's strength lies in its ability to handle the high-dimensional state and action spaces typical in educational contexts, while simultaneously learning from accumulated experiences and adjusting to each student's unique learning experience.

Algorithm 1 Deep Q-Learning for Personalized Learning Paths

Initialize Q-network f_θ with random weights θ
Initialize target network $f_{\theta-}$ with $\theta- = \theta$
Initialize replay memory D
for each apisode **do**
  Initialize state s_0
  for each time step t **do**
    Select action at using ϵ-greedy policy
    Execute a_t, observe reward r_t and next state s_{t+1}
    Store experience (s_t, a_t, r_t, s_{t+1}) in D
    Sample random minibatch of experiences from D
    for each experience (s_j, a_j, r_j, s_{j+1}) in minibatch **do**
      Compute target: $yj = r_j + \gamma \max_{a'} f_{\theta^-}(s_{j+1})_{a'}$
      Compute loss: $L_j = (y_j - f_{\theta^-}(s_j)a_j)^2$
    end for
    Update θ by gradient descent on $\sum_j L_j$
    if $t \mod C == 0$ **then**
      Update $\theta- = \theta$
    end if
  end for
  Update student profile and learning path
end for
return Optimized Q-network and policy

1.3 Validation and experimentation

To validate our proposed Deep Q-Learning model for personalized learning paths, we conducted a series of experiments using a simulated educational environment. We call DQN Agent the module that we incorporate into the tutor system in order to personalize the learning paths. As shown in Table 1.1, our experimental setup includes detailed software and hardware specifications, along with key algorithm parameters and simulated environment characteristics.

Table 1.1: Experimental Setup, Algorithm Parameters, and Simulated Environment Details

Experimental Setup	
Software Specifications	
Python	3.8.10
PyTorch	1.9.0
NumPy	1.21.0
Matplotlib	3.4.2
Pandas	1.3.0
Hardware Specifications	
CPU	Intel Core i7-10700K @ 3.80GHz
RAM	32GB DDR4
GPU	NVIDIA GeForce RTX 3080 (10GB VRAM)
Algorithm Parameters	
Learning rate (α)	0.001
Discount factor (γ)	0.99
Exploration factor ($\in$)	1.0 to 0.01 over 100,000 steps
Replay memory size	100,000 experiences
Batch size	64
Target network update frequency	Every 1000 steps
Number of episodes	10,000
Maximum steps per episode	1000
Simulated Environment	
Number of knowledge components	50
Number of learning activities	200
Number of simulated students	1000

Each student's initial knowledge state was randomly initialized, and learning activities were designed to affect multiple knowledge components with varying degrees of impact.

1.3.1 Results and analysis

Figure 1.1 shows the learning curve of our DQN algorithm over the course of training. We observe a steady increase in the average reward per episode, indicating that the algorithm is successfully learning to create more effective personalized learning paths over time.

Figure 1.2 illustrates the average knowledge gain of students over the course of their personalized learning paths. We can see that as the algorithm improves, students are able to acquire knowledge more efficiently.

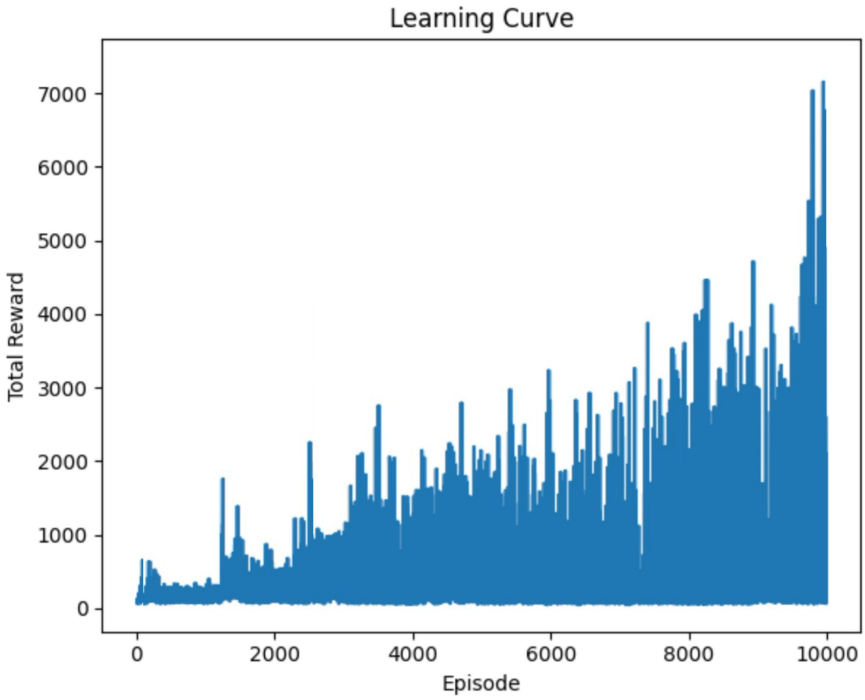


Fig. 1.1: Learning curve of the Deep Q-Learning algorithm

1.3.2 Personalized learning paths

The heatmaps of recommended learning paths for different student profiles (Figure 1.3) provide compelling evidence of the DQN agent's ability to personalize educational experiences. The DQN agent demonstrates its ability to personalize learning experiences to individual needs through differentiated paths for various student profiles. We observe distinct patterns for each profile. For beginners, the agent recommends activities with broader impact across multiple knowledge components, as evidenced by the more uniform color distribution in the heatmap, indicating a strategy to build a foundational understanding across all areas. In contrast, the intermediate profile shows a mix of broad and focused activities, suggesting a balanced approach to reinforce existing knowledge while introducing more advanced concepts, as seen in the varied intensity of colors across the heatmap. Moving on to advanced students, the recommended path focuses on activities with high impact on specific knowledge components, shown by intense colors in particular rows of the heatmap, addressing the few areas where the student's knowledge is not yet complete and optimizing the learning process for nearly-proficient learners. Additionally, the mixed profile demonstrates the agent's adaptability to uneven initial knowledge states, with the heatmap showing a varied pattern of intense colors in some areas

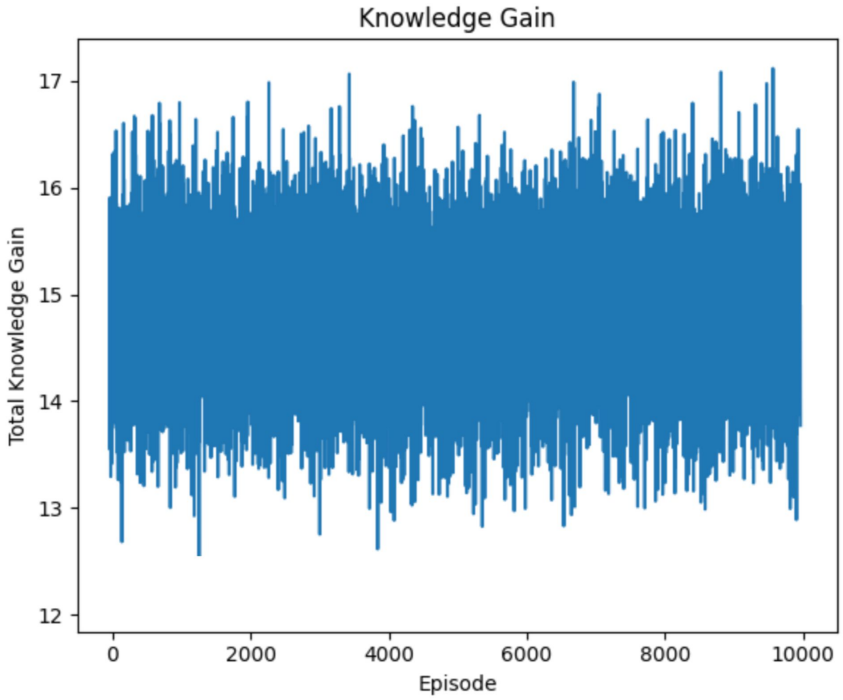


Fig. 1.2: Average knowledge gain of students over time capture

(addressing knowledge gaps) and more diffuse colors in others (advancing areas of existing strength). Ultimately, these differentiated paths highlight the DQN agent's capacity to create efficient and effective learning processes tailored to individual needs.

1.3.3 Limitations

While our findings are promising, it is important to acknowledge the limitations of this study. The simulated learning environment, while complex, does not fully capture the complexities of real-world educational settings. The synthetic student profiles, while diverse, are representations with limited levels of complexity.

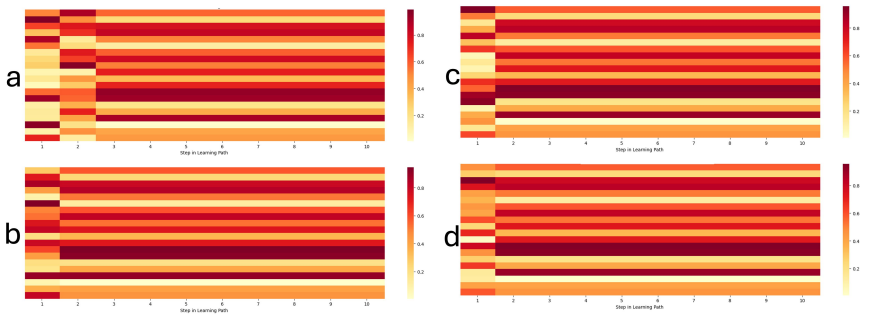


Fig. 1.3: Recommended learning path for different student profiles: a) Recommended path for beginner student 1. b) Recommended path for intermediate student 2. c) Recommended path for advanced student 3. d) Recommended path for mixed student 4.

1.4 Conclusion

This study presents a novel application of Deep Q-Learning to personalize learning paths. Our findings demonstrate the potential of reinforcement learning techniques to create personalized learning experiences. This research makes contributions to the field of adaptive learning systems. First, we present a Deep QLearning- based model that effectively selects educational activities based on a student's knowledge state. Second, our model demonstrates continuous improvement across its learning curve and knowledge acquisition. Lastly, we show the model's ability to generate diverse and appropriate learning paths for various student profiles. While this study provides valuable insights, future research is needed to address its limitations. Real-world validation with real students, improved student modeling, long-term optimization, and handling dynamic environments are important areas for further exploration. Despite these challenges, our results provide a foundation for applying Deep Q-Learning to educational technology. Acknowledgments. This work has been supported by funding from Project No. 2021000100219 – SGR of Universidad de Córdoba, and the Ministry of Science, Technology, and Innovation of Colombia. We extend our sincere gratitude to these institutions for their invaluable financial support, which has been instrumental in advancing our research endeavors. This funding has enabled us to explore the frontiers of AI in education and contribute to the ongoing discourse on personalized learning paths.

Acknowledgements This work has been supported by funding from Project No. 2021000100219 – SGR of Universidad de Córdoba, and the Ministry of Science, Technology, and Innovation of Colombia. We extend our sincere gratitude to these institutions for their invaluable financial support, which has been instrumental in advancing our research endeavors. This funding has enabled us to explore the frontiers of AI in education and contribute to the ongoing discourse on personalized learning paths.

References

1. VanLehn, K.: The relative effectiveness of human tutoring, intelligent tutoring systems,and other tutoring systems. Educational Psychologist, 46(4), 197-221. (2011).
2. Mousavinasab, E., Zarifsanaiey, N., R. Niakan Kalhori, S., Rakhshan, M., Keikha, L., & Ghazi Saeedi, M.: Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods. Interactive Learning Environments, 29(1), 142-163.(2021).
3. Mnih, V., Kavukcuoglu, K., Silver, D., et al.: Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533.(2015).
4. Fan, J., Wang, Z., Xie, Y., & Yang, Z. (2020, July). A theoretical analysis of deep Q-learning. In Learning for dynamics and control (pp. 486-489). PMLR.
5. Nabizadeh, A. H., Leal, J. P., Rafsanjani, H. N., & Shah, R. R.: Learning path personalization and recommendation methods: A survey of the state-of-the-art. Expert Systems with Applications, 159, 113596. (2020).
6. EDUTLAN Homepage, https://edutlan.co, last accessed 2024/08/07
7. SGAMEs21 Tutorial Homepage, https://manual-sgames21.vercel.app/Manual, last accessed 2024/08/07
8. Kolb, D.: Experiential learning: Experience as the source of learning and development. Prentice-Hall.(1984).
9. Morris, T. H. (2020). Experiential learning–a systematic review and revision of Kolb's model. Interactive learning environments, 28(8), 1064-1077.

2
Enhancing *K*-Centres Algorithm Efficiency in Python across Multi-Core and Many-Core Platforms

Santiago Monroy Heredia[0009-0001-7371-54401]
Ana-Lorena Uribe-Hurtado[0000-0002-1424-23722]

Abstract The article presents a parallel implementation of the *K*-Centres algorithm in Python and evaluates the performance of these implementations in terms of execution time. The performance experiments were compared with the sequential version of the algorithm. The parallelization was conducted using CPU (Multiprocessing), CuPy, Numba, and PyCUDA. Twenty-five executions of the datasets were carried out to compare each strategy, averaging the execution times. The quantitative results revealed an improvement in the algorithm's performance with parallel strategies on the GPU, with PyCUDA being the most efficient, closely followed by Numba and CuPy. The parallel implementation on the CPU was the least effective in terms of execution time. In addition to the quantitative results, qualitative aspects were considered, such as the complexity of implementation and the ease of use of each strategy. Although the GPU implementations offered better performance, the choice of strategy should weigh both speed and the complexity and resources available. This study shows different parallelization strategies in Python that improve the performance of the *K*-Centres clustering algorithm.

Key words: *K*-Centres, Parallel computer, multi-core (CPU), many-core (GPU) architectures.

2.1 Introduction

High-performance computing refers to the use of specialized computational architectures and information systems to address computational solutions that require significant processing power. An essential component of high-performance computing is the parallelization of algorithms, an approach that allows computational tasks

Universidad Nacional de Colombia - Sede
Manizales, Kilómetro 7 vía al Magdalena, Manizales, Caldas, CO
e-mail: \{smonroyh,alhurtadou\}@unal.edu.co

to be divided into multiple sub-processes or threads to be processed simultaneously, thereby achieving better response times for current speed requirements. This technique effectively leverages the capabilities of multiple central processing unit (CPU) or graphics processing units (GPUs), significantly accelerating the processing speed of an algorithm to perform scientific simulations and process large-scale data. The parallelization of algorithms is a fundamental pillar in the pursuit of efficiency and scalability in high-performance computing [6].

Clustering algorithms do not require data to be pre-labeled and are able to identify hidden patterns and inherent structures in datasets, which makes them an essential tool for identifying common patterns in many fields such as Data Mining [10].

This paper focuses on the parallelization of the unsupervised learning K-Centres clustering algorithm as implemented by Duin [1], exploring its processing capabilities on multiple architectures, including many-core GPU and multi-core CPU. We implement the sequential algorithm on the CPU, aiming to analyze and compare the advantages and disadvantages of different parallelization approaches both with the sequential reference and among themselves. Specifically, five implementations of the algorithm are considered: the sequential CPU algorithm, the parallelized CPU algorithm with threads, the parallelized PyCUDA algorithm, the parallelized Numba algorithm, and the parallelized CuPy algorithm. These variants allow for a comprehensive evaluation of how parallelization affects the performance of the K-Centres algorithm on different computational architectures, from sequential execution on the CPU to the acceleration provided by GPU. This approach offers a thorough perspective on the potential speedup and scalability improvements achievable in practical applications.

An emerging research question is: ¿What are the appropriate parallel strategies and multicore architectures to speed up the K-Centres algorithm in Python?. Despite the availability of various implementations of the K-Centres algorithm, there is a lack of parallelization in the Python programming language. To answer this question, this paper shows the implementation of four different parallelization strategies for the K-Centres algorithm, evaluating their performance in terms of execution time and considering qualitative aspects such as the complexity of each implementation and the ease of use on many-core and multi-core architectures.

The parallelization shown on CPU with multithreading, on GPU with PyCUDA, Numba and CuPy, offer efficient and accessible alternatives for the analysis of this clustering algorithm in Python, particularly in scenarios where performance and scalability are important aspects.

It is important to notice that this version of the K-Centres algorithm development in [1] was not parallelized for GPU or CPU in python, for this reason, it is not found in any Python API and so far it has only been parallelized in C language, as defined by [7] and publishes in [9].

The remainder of this paper is organized as follows: Section 2.2 presents the four parallel implementations along with their corresponding code. Section 2.3 discusses the speedup and elapsed time of the proposed solutions. Finally, concluding remarks are provided in Sections 2.2 and 2.3.

2.2 Methods

This section presents explanations of the algorithm K-Centres (see Section 2.2.1), as well as the multithreading (CPU) (see Section 2.2.2) and many-core (GPU) (see Section 2.2.4) parallel implementations. Due to space constraints in short paper, the parallelization details of the Python programs are documented within the code itself, you can see the codes in https://github.com/alhurtadou/K-Centres-Python-ParallelCPU-GPU. The programs were implemented and executed on a Google Colab notebook, the CPU and GPU were those provided by these platforms.

2.2.1 K-Centres Algorithm

The CPU-based sequential implementation of K-Centres was described in detail in [8]. In the following lines we describe it again for clarity in this paper. The algorithm assumes that a square matrix $D(N \times N)$, containing all the pairwise dissimilarities between the N objects of the data set, has been pre-computed. We use the euclidean dissimilarity $d = \sqrt{\sum (x_i - y_j)^2}$ to calculate the distance between each object respect to each other, where x and y are objects in the data set.

1. A set of K objects, out of N ones, is selected at random. They constitute the initial centres and are stored in a vector **p**. The indexes of entries in **p** are $j = \{0, \ldots K-1\}$.
2. For all the N objects in the data set, the nearest centres are found and the indexes of the corresponding associations are stored in a vector of labels **l**. The indexes of l are $i = 0 \ldots N-1$.
3. Afterwards, in order to update the centres which are now stored in a vector **c**, for each group (objects having the same labels l_i) the new centre c_j is found as the object whose maximum distance to all the objects in the current group is minimum.
4. The algorithm converges in case that the updated centres **c** are the same previous ones **p**; otherwise, it returns to step 2 storing in **p** the centres found in step 3.

The advantages that we analyzed of the algorithm with respect to others such as the well-known K-Means are: it selects centres that are real objects in the data set and by calculating the maximum of the minima as described in step 3, it ensures that objects far from the edges of each group are well clustered.

2.2.2 Parallelization of the K-Centres algorithm in Python

For the parallel computation of the K-Centres algorithm in Python, four programs were implemented: one program on the CPU using multithreading, and three pro-

grams on the GPU using the PyCUDA, CuPy, and Numba libraries. This section describe briefly each of them.

2.2.3 K-Centres multithreading in CPU

In this context, parallel implementation on the CPU was explored using Python and the multiprocessing library. Four independent functions were developed to execute the steps of the K-Centers algorithm. The first function, "EDM", computes the distance matrix in parallel using Euclidean distance. The second function, "findMaxima", identifies the maximum (or border) values of each cluster. The third function, "findMinimumOfMaxima", selects the minimum among these maximum values. Finally, the "KCentres" function integrates these steps to implement the algorithm.

To parallelize the K-Centers algorithm, the dataset is divided into partitions (chunks). The size of these chunks is determined by dividing the number of objects by the number of CPU cores. Using Python's multiprocessing library, a pool of processes is created based on the number of cores available. This setup defines the number of tasks each core will execute. The pool function is then used to map the data according to the calculated number of chunks. It is crucial to define these partitions carefully, ensuring they align with the number of cores to prevent overloading any single core.

2.2.4 GPU strategies

In this section, we explore the primary strategies for implementing the K-centres algorithm on GPUs. We have chosen three different Python libraries for this purpose: PyCUDA see 2.2.4.1, CuPy see 2.2.4.2, and Numba 2.2.4.3. While all three libraries allow GPU code execution, they differ in their approaches to data mapping, compilation, and execution.

2.2.4.1 PyCUDA in Python

The implementation of the K-Centres algorithm, already available in CUDA by [9], is being explored, taking advantage of the computational power of GPU. Additionally, the adaptation of this implementation in PyCUDA will be investigated, a library that allows integrating CUDA code into Python applications. This exploration aims to evaluate the differences in performance and flexibility between both implementations, as well as their ability to address the clustering problem in large datasets. By implementing it in PyCUDA, the goal is to determine the balance between efficiency and ease of development offered by this tool, contributing to the understanding of how GPU and Python libraries can be effectively used in clustering tasks. For the cal-

culation of the dissimilarity matrix, it is based on the already implemented approach known as Euclidean Dissimilarity Matrix on GPU (EDMGPU) [7].

To calculate the dissimilarity matrix in GPU kernel with PyCUDA an important strategy is to handle the transposition of the dataset matrix when manipulating the array inside the kernel. Consequently, this kernel aims to return the dissimilarity matrix of the transposition of a data matrix. When working with GPU and CUDA, data organization and access are critical aspects to achieve optimal performance.

The unexpected data transposition when working on the GPU often arises from differences in how data is stored and accessed in GPU memory compared to CPU memory. In the CPU, data is often stored in a matrix format where elements are contiguous in memory. When working on the CPU, it is more natural to access elements of a matrix in a "row by row" fashion. This means that when accessing an element in memory, it sequentially traverses rows in memory. In contrast, in the GPU, the memory organization and data access patterns can be different. Data is stored in a way that favors collective and parallel access by multiple execution threads in a block or warp. This can lead to a memory layout where elements are stored "column by column" instead of "row by row" [4].

2.2.4.2 Cupy

specifically, for distance matrix computation with CuPy, a technique known as broadcasting is used when performing a vectorized operation, such as addition, between two matrices with the same shape, it is clear what should happen. Through broadcasting, dimensions are allowed to differ and produce results that are intuitive. A trivial example is the addition of a scalar value to an array, but broadcasting also generalizes to more complex examples. In broadcasting, one or both matrices are virtually replicated (i.e., without copying data into memory), such that the shapes of the operands match [3]. CuPy also provides a significant advantage in preparing variables for use in GPU kernels by eliminating the constant need to transfer data between the host and device. By allowing direct creation of variables on the GPU, CuPy greatly simplifies manipulation and performing operations, such as comparisons, without requiring transfers between the host and device. This feature of CuPy significantly simplifies variable handling compared to PyCUDA by eliminating additional transfer steps.

2.2.4.3 Numba

Numba uses a Just-In-Time (JIT) compilation approach that generates machine code during runtime. This means that kernel functions are written in pure Python and decorated with the @jit decorator to tell Numba to compile these functions. Numba attempts to automatically infer the data type and memory layout for the GPU. Unlike most JIT compilers for interpreted languages, Numba does not perform backtrace or replace the interpreter. Instead, it relies on the user to actively transform Python functions that need to be compiled at runtime. In Python, this is achieved by applying

a decorator to the function. The decorator replaces the original Python function with a special object that compiles the function just in time when it is first called. As mentioned above, Numba inspects the types of the arguments and performs local type inference on the function, as a result, the compiler can have accurate information about the type of each value in the function without backtracking the execution. Numba facilitates GPU programming by supporting GP-GPU backends that expose parallel execution of the hardware model, similar to CUDA-C and OpenCL. In this approach, a Python function can be decorated as a GPU kernel. This strategy has the advantage of simplifying implementation, since GP-GPU backends can share much of the code generation logic with the CPU backend. Numba offers both automatic and manual memory management between GPU device memory and host CPU memory. When an ndarray is used as an argument, memory is automatically transferred to the device and copied back to the host once the kernel has completed. While this transfer is transparent to the user, it can lead to unnecessary transfers since Numba cannot determine whether an ndarray has ever been modified. For maximum control, users can explicitly perform memory transfer using the copy_to_host and copy_to_device functions, see [5].

2.3 Dataset and results

2.3.1 Dataset

We pre-processed a total of 3012 WAV audio files obtained from the web page[1] to create the Yataros1.csv dataset. Utilizing the Guyot and Eldridge [2] API from a public repository[2], we computed 40 acoustic indices for each audio file and stored them in the Yataros1.csv dataset, resulting in 3012 objects, each containing 40 acoustic indices. Leveraging this dataset along with four parallel implementation of the K-Centres algorithm wih $K = 5$, we aimed to identify the five most representative indices of the forest.

2.3.2 Results

Table 24.1 presents the elapsed times (ET) and the speed ups of 25 executions of the sequential and parallel implementations, the elapsed times were averaged to achieve meaningfully representations of performance. The speed up was calculated as $SUp = \frac{ET_{sequential}}{ET_{parallel}}$.

The larger elapsed time with CPU-threads implementation compared to the sequential implementation, see Table 24.1, may be due to the overhead associated with

[1] http://colecciones.humboldt.org.co/rec/sonidos/publicaciones/MAP/JDT-Yataros/

[2] https://github.com/patriceguyot/Acoustic_Indices/

Table 2.1: Elapsed times of K-Centres in CPU and GPU implementations

	Sequential	**CPU-threads**	**CuPy**	**Numba**	**PyCUDA**
ET	0.00299	0.080072	0.000875	0.001109	0.000178
SUp	1	0.037341	3.417142	2.696122	16.797752

inter-process communication and coordination. With the use of Multi-processing in Python, the focus is on the ability to execute several tasks at the same time using multiple processes, this can be beneficial for tasks that are intensive in terms of CPU usage, however, there is a downside; the creation and coordination of these additional processes introduces some "overhead" or additional costs. The overhead can be more noticeable when the tasks being parallelized are simpler or run faster with a sequential implementation that benefits the CPU rather than scheduling and managing multiple processes.

As you can see in Figure 14.1 the speed up of all parallel implementations was from $2x$ up to $16x$ better than the sequential version (red point). The variation in the elapsed times between the different approaches implemented is an evidence of the diversity of processing strategies.

Sequential execution, despite its simplicity, outperforms the implementation of multi-threading on the CPU in this scenario due to the computational overhead incurred by coordinating the multi-thread scheduler handling small tasks.

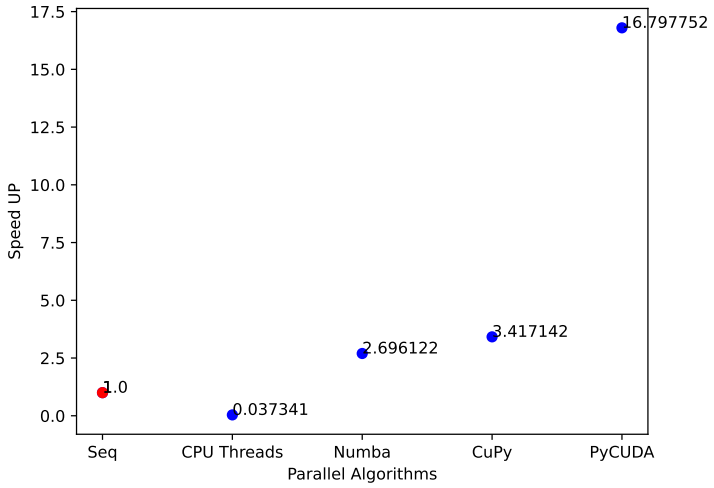


Fig. 2.1: Parallel implementation speed up

When considering parallel processing, CuPy stands out for its ability to operate directly on the GPU, resulting in significantly enhanced performance thanks to its extensive parallelization capabilities. Numba, leveraging Just-In-Time technology to compile functions into machine code through decorators [3], shows further efficiency gains. Meanwhile, PyCUDA shines as the most efficient option, seamlessly integrating CUDA code into Python and harnessing CUDA is intrinsic optimizations for compute-intensive tasks. To ensure that the results of the implementations are comparable, the same centers are chosen in the parallel executions.

2.4 Discussion

Key metrics, such as elapsed times and processing speed ups, are explored with the objective of highlighting the improvements achieved through the introduction of parallelism. In addition, specific cases are examined to provide a deeper understanding of the scenarios in which parallel implementations demonstrate their usefulness. This analysis will allow a comprehensive evaluation of the contributions made and their impact on the efficient resolution of specific problems.

The results illustrate a key insight into the trade-offs between sequential and parallel computing approaches, particularly in the context of small, fast tasks. While sequential execution might seem less sophisticated, it can outperform multi-threading on CPUs due to the overhead involved in coordinating multiple threads. The increased elapsed time in the multi-threaded implementation suggests that inter-process communication and the need for synchronization add significant computational costs.

These findings highlight that for certain types of tasks—particularly those that are simple or execute quickly—sequential execution may actually be more efficient. This is an important consideration for researchers working on parallel computing, as it suggests that the complexity and overhead associated with parallelism might outweigh the benefits for smaller tasks.

However, the study also demonstrates the power of parallel computing when applied correctly. The parallel implementations leveraging CuPy, Numba, and PyCUDA achieved substantial speedups, indicating that for more complex or larger-scale tasks, parallelism can significantly enhance performance. This suggests that research in parallel computing should focus on identifying task characteristics that can benefit from parallelization, while also accounting for the potential overhead.

2.5 Conclusion

Despite the simplicity of the sequential execution, it demonstrated superior performance over the multi-threading approach on the CPU, mainly due to the computational overhead associated with coordinating the multi-thread CPU scheduler handling in small tasks. The longer elapsed time observed with the CPU-threads im-

plementation compared to the sequential implementation suggests potential overhead from inter-process communication and coordination.

While multiprocessing in Python enables concurrent execution of multiple tasks using multiple processes, it introduces additional costs and complexities, which are particularly noticeable when parallelizing simpler or faster tasks. Moreover, the synchronization of multiple processes to avoid race conditions adds to the overhead, as does the process creation and termination. Despite these challenges, the parallel implementations, taking advantage of CuPy, Numba, and PyCUDA, showcased significant speedups ranging from 2x to 16x compared to the sequential version, highlighting the diverse processing strategies and the potential for enhanced efficiency with parallel computing approaches. Implementations using Cupy and Numba minimize the need for developers to understand the underlying computer architecture. By efficiently handling data transfer between the CPU and GPU, these libraries facilitate faster parallel development on GPU architectures. As a result, users can focus more on computational implementation without delving into intricate architectural details.

Acknowledgements The authors gratefully acknowledge the "Universidad Nacional de Colombia - Sede Manizales" and research group "Ambientes inteligentes adaptativos (GAIA)", for their support in the development of this paper.

References

1. Duin, R., Juszczak, P., Paclik, P., Pekalska, E., De Ridder, D., Tax, D., Verzakov, S.: Prtools 4.1. A Matlab toolbox for pattern recognition (2007)
2. Eldridge, A., Guyot, P., Moscoso, P., Johnston, A., Eyre-Walker, Y., Peck, M.: Sounding out ecoacoustic metrics: Avian species richness is predicted by acoustic indices in temperate but not tropical habitats. Ecological Indicators **95**, 939–952 (2018). https://doi.org/https://doi.org/10.1016/j.ecolind.2018.06.012
3. Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M., Haldane, A., del Río, J.F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E.: Array programming with numpy. Nature pp. 357, 362 (2020)
4. Hijma, P., Heldens, S., Sclocco, A., van Werkhoven, B., Bal, H.E.: Optimization Techniques for GPU Programming. ACM Comput. Surv. **55**(11) (mar 2023). https://doi.org/10.1145/3570638
5. Lam, S.K., Pitrou, A., Seibert, S.: Numba: A llvm-based python jit compiler. ACM Digital Library (2015)
6. Trobec, R., Slivnik, B., Bulić, P., Robič, B.: Why Do We Need Parallel Programming, pp. 3–7. Springer International Publishing, Cham (2018). https://doi.org/10.10079783319-988337_1
7. Uribe-Hurtado, A.L.: Aceleración de algoritmos de clasificación basados en disimilitudes, utilizando arquitecturas computacionales con múltiples núcleos. Ph.D. thesis, Universidad Nacional de Colombia. (2022)
8. Uribe-Hurtado, A.L., Orozco-Alzate, M., Lopes, N., Ribeiro, B.: KCentres algorithm in GPU for clustering on manycore architectures: A preliminary approach. In: RECPAD 2018. 24th Portuguese Conference on Pattern Recognition - Proceedings. pp. 81–82 (2018)
9. Uribe-Hurtado, A.L., Orozco-Alzate, M., Lopes, N., Ribeiro, B.: Gpu-based fast clustering via k-centres and k-nn mode seeking for geospatial industry applications. Computers in Industry

122, 103260 (2020). https://doi.org/https://doi.org/10.1016/j.compind.2020.103260, https://www.sciencedirect.com/science/article/pii/S0166361520304942

10. Zou, H.: Clustering algorithm and its application in data mining - wireless personal communications. Wireless Personal Communications (2020). https://doi.org/10.1007/s11277-019-06709-z

3
Plataforma Tipo Banco De Recursos Educativos Digitales para Facilitar el Acceso a Materiales Educativos en Zonas Rurales sin Conectividad a Internet

Michelle Ceballos[1] [0009-007-67480909]
Bryjeth Ceballos[2] [0000-0001-7735-3959]
Néstor Duque-Méndez[3] [0000-0002-4608-281X]
Patricia Paderewski[4] [0000-0001-6626-9633]

Resumen Esta investigación busca contribuir a la disminución de la brecha digi-tal educativa en zonas rurales, la cual se atribuye en parte a la escasa disponibili-dad de contenidos educativos digitales adaptados a las características del con-texto y las necesidades específicas de los estudiantes, así como la limitada infra-estructura eléctrica y el escaso acceso a internet. La solución presentada en este trabajo se denomina RuralTIC, una plataforma tipo banco de recursos educativos digitales abiertos diseñada para facilitar el acceso tanto de docentes como estu-diantes a materiales educativos de calidad en zonas rurales de difícil acceso a internet. RuralTIC permite el almacenamiento y distribución de recursos educa-tivos de manera offline, fomentando el aprendizaje y la enseñanza de forma con-tinua y accesible en estas comunidades.

Palabras clave: Tecnologías de la información y la comunicación (TIC), zonas ru-rales, recursos educativos digitales abiertos, plataforma educativa digital

3.1 Introducción

La era digital ha transformado la educación, convirtiendo los recursos educativos digi-tales en herramientas esenciales. Sin embargo, la pandemia de COVID-19 marcó los desafíos, especialmente en comunidades rurales, donde el acceso a tecnología es limi-tado, revelando fallos en los procesos gubernamentales y educativos [1]. Las escuelas rurales sufren de infraestructura deficiente, con instalaciones en mal estado, recursos tecnológicos obsoletos y acceso limitado a servicios básicos como

[1,3,4,5]Universidad Nacional de Colombia, Manizales, Colombia
e-mail: dceballosc@unal.edu.co
[2]Universidad de Granada, Granada, España

electricidad e inter-net [2][3]. La falta de conectividad restringe el acceso a contenidos educativos, exacerbando la exclusión educativa y dificultando la transición hacia el aprendizaje en línea [4]. Ade-más, la baja calidad de los recursos educativos, agravada por la insuficiente capacita-ción docente y el acceso limitado a tecnologías, afecta negativamente el aprendizaje y la capacidad de implementar metodologías innovadoras [5]. Para abordar estos desafíos, este estudio propone RuralTIC, un software diseñado para funcionar sin necesidad de conexión a internet, permitiendo tanto a los docentes como a los estudiantes visualizar e interactuar con los recursos digitales abiertos disponibles en la plataforma. La plataforma abarca asignaturas fundamentales como len-guaje, matemáticas y ciencias naturales, con el objetivo de proporcionar material edu-cativo e interactivo a estudiantes de primero a quinto grado de primaria.

3.2 Marco téorico

Para sustentar esta investigación, se utilizó la metodología de Watson y Webster [6]. Esta revisión reveló hallazgos clave sobre el acceso a recursos educativos digitales en zonas rurales y el impacto de las tecnologías en la educación en estos contextos. En particular, se destacan características comunes entre los diferentes repositorios e inicia-tivas analizadas, como el enfoque en el uso TIC para mejorar la educación rural y los retos comunes en su implementación. Además de esto la revisión de literatura permitió conocer que la infraestructura deficiente y las bajas competencias digitales de los do-centes son otros factores que impactan en la brecha digital que se da entre las zonas rurales y las zonas urbanas. A continuación, se destacan las principales características de repositorios actuales como Portal Colombia Aprende [7], MECRED [8], MERLOT [9], Kolibri [10], Khan Academy [11] y Aula Libre, frente a lo que RuralTIC ofrece (Ver Tabla 3.1).

3.3 Metodología

Para el desarrollo de RuralTIC, se adoptó la metodología MOA-UX [12], con el obje-tivo de mejorar la satisfacción del usuario y la eficiencia del desarrollo. La meta prin-cipal es crear una plataforma que facilite el acceso a recursos educativos digitales y mejore la calidad educativa en zonas rurales (ver Fig. 3.1).

Table 3.1: Comparativa de Características entre repositorios actuales y la plataforma.

Característica	**Repositorios Actuales**	**RuralTIC**
Acceso a recursos Educativos	Acceso en línea, con limitadas opciones of-fline (Kolibri, Khan Academy Offline, Aula Libre).	Funcionamiento offline completo para comunidades rurales sin acceso a Internet.
Compatibilidad tecnológica	Limitaciones con la infraestructura rural (Colombia Aprende, MECRED, Kolibri).	Compatibilidad garantizada con la infraestructura tecnológica disponible en las escuelas rurales colombianas.
Interfaz y usabilidad	Varían en complejidad; algunas plataformas son difíciles de usar para personas con baja alfabetización digital, por ejemplo: MERLOT	Diseño intuitivo y amigable, optimizado para accesibilidad y usabilidad en contextos rurales.
Contenido educativo	Amplia variedad de recursos organizados por nivel educativo y asignatura, pero algunos requieren conexión a Internet como es el caso de MECRED y MERLOT	Recursos cuidadosamente seleccionados y organizados, alineados con las mallas de aprendizaje del gobierno colombiano, disponibles para descarga y uso offline.
Creación de planes de estudio personalizados	La mayoría de repositorios no ofrecen la capacidad de personalizar planes de estudio.	Creación y personalización de planes de estudio según las necesidades de los docentes y estudiantes.
Soporte multiplataforma	Dependencia de una plataforma en línea; algunas ofrecen versiones offline como Kolibri, Khan Academy Offline, Aula Libre.	Disponible a través de múltiples medios: web, USB, correo electrónico, Drive, permitiendo acceso flexible a los recursos.
Reutilización de recursos	Promueven la reutilización de recursos, pero dependen de la conectividad y la capacidad técnica de los usuarios.	Enfoque en la reutilización de recursos que son fáciles de acceder y utilizar sin necesidad de conectividad constante.
Evaluación y retroalimentación	Algunas plataformas permiten la evaluación de los recursos y retroalimentación por parte de usuarios de forma limitada.	Evaluación de la plataforma mediante cuestionarios y observación directa, con retroalimentación iterativa para mejorar la usabilidad y efectividad de los recursos.
Soporte y capacitación	Requiere capacitación continua para uso efectivo, especialmente en plataformas más complejas que están disponibles en otros idiomas diferentes al español.	Disponibilidad de tutoriales que orientan a los usuarios en el uso efectivo de la plataforma, asegurando una curva de aprendizaje más suave para los docentes rurales.
Enfoque en comunidades rurales	Enfocado principalmente en contextos urbanos o con infraestructuras robustas; algunas excepciones con opciones offline que pueden ser utilizados en zonas rurales son Kolibri, Khan Academy Offline y Aula Libre	Diseñada específicamente para las necesidades de las comunidades rurales, con re-cursos optimizados para contextos con infraestructura limitada y sin conexión a Internet.

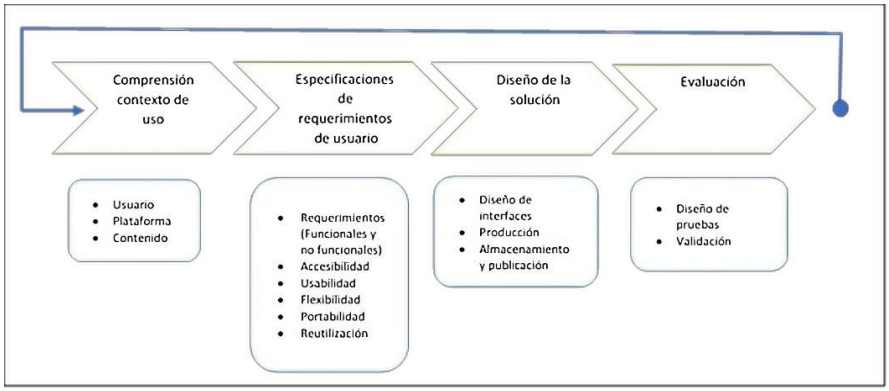


Fig. 3.1: Metodología MOA-UX adaptada a RuralTIC.

3.3.1 Comprensión del Contexto de Uso

3.3.2 Especificaciones de Requerimiento de Usuario

Requerimientos funcionales La plataforma brinda a los profesores la oportunidad de personalizar planes de estudio, categorizando los recursos por áreas de aprendizaje y grados. Además, los recursos estarán alineados con los estándares educativos colom-bianos, asegurando su relevancia. Para facilitar el acceso, los recursos podrán descar-garse en dispositivos USB, enviarse por correo electrónico o almacenarse en Google Drive, proporcionando una distribución flexible que se adapta a las limitaciones de co-nectividad en zonas rurales.

Requerimientos no funcionales La plataforma es interactiva, intuitiva y rápida, con un tiempo de respuesta máximo de 3 segundos por página para asegurar una experiencia fluida. Es compatible con diversos dispositivos y sistemas operativos, accesible 24/7, y diseñada para incluir a usuarios con discapacidades. La plataforma también es modular y escalable, permite una fácil personalización y expansión conforme crece la base de usuarios y recursos. Los recursos educativos son portables y reutilizables en distintos contextos educativos, aumentando su valor y eficacia.

3.3.3 Diseño de la solución

Diseño de interfaces. En el diseño de interfaces de la plataforma educativa, se abordaron tres aspectos fundamentales: el diseño computacional, el diseño de la interfaz y el diseño de contenidos. La plataforma fue desarrollada siguiendo una arquitec-

tura en ca-pas [13], cada una con funciones específicas. La capa de presentación se diseñó para ser intuitiva y accesible, facilitando la búsqueda, selección, visualización y descarga de recursos. La capa de negocio incluye funcionalidades como la creación de planes de estudio y una búsqueda organizada de recursos, mientras que la capa de datos permite el almacenamiento local de recursos para uso offline en zonas rurales con acceso limi-tado a Internet. La implementación se realizó utilizando tecnologías web como HTML, CSS y JavaScript, y se empleó CANVA para logotipos e imágenes, modificando dise-ños de usuarios como Sketechify, Sparklestroke, Gambar-Pixcel-3D y Sketchifyedu.

Producción. Durante esta etapa, se construyó la plataforma integrando los recursos recolectados de manera estructurada, organizada y comprensible, facilitando la interac-ción del usuario. Este proceso fue iterativo, con mejoras continuas basadas en la retro-alimentación de expertos. La interfaz de usuario se diseñó a partir de mockups detalla-dos que representan visualmente los elementos de la plataforma. En la página principal, el usuario puede acceder a los grados disponibles y obtener información sobre RuralTIC, así como explorar recursos por área de aprendizaje o grado específico y acceder a los recursos educativos digitales correspondientes. Cada recurso seleccionado ofrece detalles y opciones para agregar a favoritos, descargar, pre-visualizar, ver créditos y organizar en carpetas personalizadas, facilitando la gestión y el acceso eficiente a los materiales educativos (ver Fig. 3.2). Además, la interfaz propor-ciona una sección dedicada para gestionar los recursos favoritos y organizar el estudio de manera personalizada y ordenada (ver Fig. 3.3).

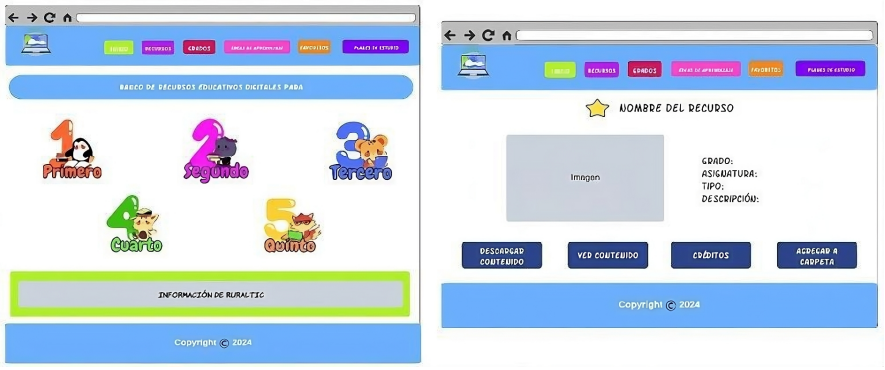


Fig. 3.2: Interfaz principal con grados e información de RuralTIC y gestión del recurso seleccio-nado

Almacenamiento y publicación. Los recursos educativos se almacenan de forma local, categorizados por áreas de aprendizaje y niveles educativos para garantizar su accesi-bilidad y gestión eficiente. Se desarrolló una interfaz intuitiva que permite a los usuarios encontrar rápidamente los recursos mediante una categorización clara. Además, se apli-caron criterios específicos para seleccionar y gestionar los recur-

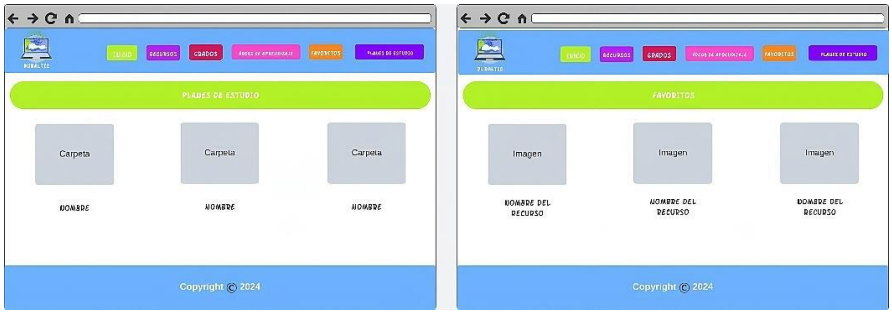


Fig. 3.3: Gestión de Recursos Favoritos y Organización de Carpetas en RuralTIC

sos, asegurando su re-levancia y adecuación para el contexto educativo rural. Los recursos fueron revisados y aprobados antes de ser publicados en la plataforma para asegurar su calidad. Se implementó un sistema de soporte y un canal de comentarios para resolver pro-blemas, responder preguntas y recibir sugerencias. Finalmente, se distribuyó la plata-forma ampliamente, asegurando acceso a través de USB, correo electrónico y Google Drive, incluso en zonas rurales con baja conectividad.

3.3.4 Evaluación

Diseño de pruebas y validación. Se seleccionó un grupo de 10 expertos en TIC y educación mediada por TIC, provenientes de la Universidad Nacional de Colombia, la Universidad de Granada (España) y la Pontificia Universidad Católica del Ecuador. El criterio de selección de los expertos fue profesionales en el área de tecnología y/o edu-cación que dictaran o hayan dictado clases en zonas rurales.

Dimensiones de evaluación. La evaluación se enfocó en cuatro dimensiones clave, siguiendo la guía de los autores Yorely Ceballos, Iván Jaramillo, Néstor Duque, y Car-los Arturo [14]. En primer lugar, la dimensión de contenido evaluó la relevancia, ade-cuación y organización de los recursos educativos, asegurando que fueran pertinentes y accesibles para el contexto rural. La dimensión de presentación se centró en la clari-dad, el atractivo visual y la estructura de la información, verificando que el diseño fuera intuitivo y fácil de navegar. La dimensión de interacción examinó la facilidad de navegación y acceso a los recursos, incluso en condiciones de conectividad limitada, garan-tizando una experiencia de usuario eficiente. Finalmente, la dimensión educativa eva-luó la efectividad de la plataforma en mejorar el aprendizaje y apoyar a los docentes, determinando su utilidad en la enseñanza en entornos rurales.

Instrumento de evaluación. Como se muestra en la Tabla 3.2 del formulario de evaluación utilizado, se incluyeron 12 preguntas, distribuidas según las dimensiones mencionadas. Cada pregunta se calificó en una escala *Likert* de cinco puntos, desde

"totalmente en desacuerdo" (1) hasta "totalmente de acuerdo" (5). Al final del formulario, se dejó un espacio para que los expertos pudieran añadir comentarios o sugerencias.

Table 3.2: Evaluación Integral de la Plataforma RuralTIC

Evaluación RuralTIC
Pregunta
Dimensión de contenido
¿Considera que el contenido proporcionado en la plataforma es relevante para el contexto educativo rural?
¿Cree que los materiales educativos disponibles son adecuados para las caracterísicas y necesidades específicas de los estudiantes en zonas rurales?
¿Los recursos educativos están bien organizados dentro de la plataforma?
Dimensión presentación
¿La información presentada en la plataforma es clara?
¿El diseño visual de la plataforma es atractivo?
¿La organización y estructura dela plataforma faciita el acceso y uso de los recursos?
Dimensión interacción
¿La plataforma apoya el proceso de enseñanza de los docentes?
¿Recomendaria esta platatorma para ser ulizada por las comunidades rurales?
Sugerencias
¿Tiene alguna sugerencia específica para mejorar la plataforma?

3.4 Análisis de resultados

El análisis de los resultados del cuestionario de evaluación de RuralTIC mostró un pre-dominio de opiniones positivas (ver Fig. 3.4). En las cuatro dimensiones evaluadas, las respuestas superaron consistentemente un valor de 4 en una escala de 1 a 5, reflejando un alto nivel de satisfacción y aceptación hacia la plataforma. Los expertos destacaron que RuralTIC ofrece recursos educativos bien organizados y relevantes para el contexto rural, con una interfaz gráfica atractiva y fácil de navegar, lo que facilita el acceso a los recursos y la adquisición de conocimientos. En la dimensión de interacción, se concluyó que la plataforma es intuitiva y eficiente, permitiendo un acceso rápido a los recursos sin necesidad de una conexión estable a internet. En cuanto a la dimensión educativa, la mayoría de los expertos consideraron que RuralTIC es una herramienta útil para me-jorar los procesos de enseñanza y aprendizaje, recomendándola por su capacidad para hacer el aprendizaje más agradable e interactivo. Se sugirió incrementar los recursos interactivos y desarrollar manuales y tutoriales para docentes.

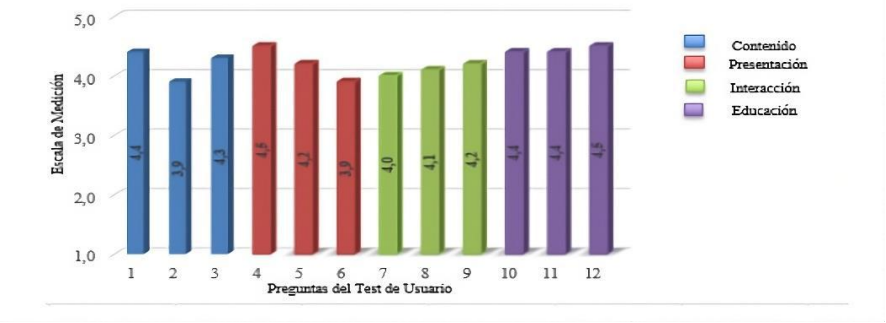


Fig. 3.4: Resultados de la Evaluación de la Plataforma RuralTIC

3.5 Conclusiones y trabajo fututo

RuralTIC es un avance crucial en el acceso a recursos educativos digitales en zonas rurales con infraestructura tecnológica limitada y baja conectividad a internet. A diferencia de otras plataformas, sobresale por su capacidad de operar eficientemente sin necesidad de una conexión constante, permitiendo a docentes y estudiantes acceder a materiales educativos offline. Esto es especialmente relevante en contextos rurales donde la tecnología es obsoleta y la conectividad inestable o inexistente. Además de facilitar el acceso a recursos de calidad, RuralTIC se adapta a las limitaciones tecnoló-gicas de estas comunidades, promoviendo un aprendizaje continuo y efectivo. La plataforma no solo aborda problemas de acceso, sino que también se enfoca en la formación continua de docentes y en el fortalecimiento de competencias digitales, ase-gurando un impacto sostenible. Al equipar a los docentes con las herramientas y habi-lidades necesarias, RuralTIC contribuye a cerrar la brecha digital y a construir una co-munidad educativa más resiliente y autónoma. Este enfoque integral hace de RuralTIC una solución superior a las alternativas existentes, respondiendo efectivamente a los desafíos educativos en áreas rurales. En el futuro, se trabajará en la construcción de un entorno que permitirá a los docen-tes cargar, editar, publicar y compartir sus propios recursos educativos, garantizando la veracidad de la información y la integración del recurso. Además, se implementará un proceso continuo de actualización de contenidos para asegurar que los docentes y estu-diantes siempre tengan acceso a materiales relevantes y de alta calidad. En resumen, RuralTIC no solo ofrece una solución técnica adecuada, sino que también se distingue por su capacidad de adaptación y enfoque en el desarrollo sostenible de la educación en comunidades rurales.

Referencias

1. N. Landa, S. Zhou, and N. Marongwe, “Education in emergencies: Lessons from COVID-19 in South Africa,” Int. Rev. Educ., vol. 67, no. 1–2, pp. 167–183, 2021.
2. S. D. Mwapwele, M. Marais, S. Dlamini, and J. Van Biljon, “Teachers’ ICT Adoption in South African Rural Schools: A Study of Technology Readiness and Implications for the South Africa Connect Broadband Policy,” African J. Inf. Commun., no. 24, pp. 1–21, 2019.
3. M. Mahdum, M. Safriyanti, and H. Hadriana, “Exploring teacher perceptions and motivations to ICT use in learning activities in Indonesia,” Technol. Educ. Res., vol. 18, pp. 293–317, 2019.
4. B. Dube, “Rural online learning in the context of COVID-19 in South Africa: Evoking an inclusive education approach,” Multidiscip. J. Educ. Research, vol. 10, no. 2, pp. 135–157, 2020.
5. K. Okoye et al., Impact of digital technologies upon teaching and learning in higher education in Latin America : an outlook on the reach , barriers , and bottlenecks. Springer US, 2023.
6. R. T. Watson and J. Webster, “Analysing the past to prepare for the future : Writing a literature review a roadmap for release 2 . 0,” J. Decis. Syst., vol. 29, no. 3, pp. 129–147, 2020.
7. M. de E. N. MEN, “Colombia Aprende,” 2024. [Online]. Available: https://www.colombiaaprende.edu.co/.
8. Ministerio de educación de Brasil, “MEC RED,” Plataforma MEC de Recursos Educacionais Digitais, 2023. [Online]. Available: https://plataformaintegrada.mec.gov.br/sobre.
9. MERLOT, “MERLOT,” Multimedia Educational Resource for Learning and Online Teaching, 2024. [Online]. Available: https://www.merlot.org/merlot/index.htm.
10. Learning Equality, “Kolibri,” Kolibri, 2022. [Online]. Available: https://learningequality.org/kolibri/.
11. Organización Khan Academy, “Khan Academy,” 2020. [Online]. Available: https://es.khanacademy.org/.
12. G. Lafuente and J. Filippi, “Un enfoque metodológico contemplando la experiencia de usuario en el desarrollo de objetos de aprendizaje,” Rev. Académica Investig. Docencia y Extensión las Ciencias Vet., vol. 3, no. April, 2022.
14. Y. Ceballos, I. Jaramillo, N. Duque, and C. Niaza, “Diseño y creación de recursos digitales etnoeducativos con contenido lúdico: Pueblo indigena Embera Chami,” Rev. Vínculos, vol. 13, no. 1, pp. 35–44, 2016.

4
Representación de dominios de software científico: un aprendizaje continuo a partir de estructuras matemáticas de los esquemas preconceptuales

Paola Noreña-Cardona[1][0000-0002-3509-7354]
Elizabeth Suescún[2][0000-0001-7872-7638]
Johnathan Calle[3][0000-0002-7297-1362]
Carlos M. Zapata[4][0000-0002- 0628-4097]
Carlos Orozco[5][0000-0003-2615-3294]
Carlos Alvarez[6]

Resumen Los dominios de software científico (DSC) permiten describir un conjunto común de elementos, incluyendo vocabulario y funcionalidades específicas, en sistemas de software aplicados a campos científicos. Estos dominios suelen incorporar modelos matemáticos en su especificación. Aunque algunos modelos de cómputo se utilizan para representar estructuras matemáticas básicas en un DSC, carecen de capacidades para representar ecuaciones matemáticas complejas, así como el vocabulario y las funcionalidades completas de los DSC. Los esquemas preconceptuales (EP) son modelos de representación del conocimiento que se utilizan para caracterizar elementos y relaciones de un dominio. La notación de los EP integra características tanto comportamentales como estructurales, incluyendo notación matemática, a diferencia de otros modelos. Por lo tanto, se propone el uso de las estructuras matemáticas de los EP para representar los DSC, tener trazabilidad desde el espacio del problema al espacio de la solución y con ello facilitar el aprendizaje continuo de los científicos en los procesos de desarrollo. La notación de los EP permite a científicos, desarrolladores y analistas modelar y codificar aplicaciones científicas desde la perspectiva de la ingeniería de software, reduciendo así la brecha entre esta disciplina y la ciencia.

Palabras clave: Aprendizaje continuo, esquemas preconceptuales, dominios de software científico, software científico

[1,2,5,6]Escuela de Ciencias Aplicadas e Ingenierías, Universidad EAFIT, Medellín, Colombia
e-mail: \{panoreaac,esuescu1,corozcog,calvar52\}@eafit.edu.co
[3]Facultad de Ingenierías, Instituto Tecnológico Metropolitano, Medellín, Colombia
e-mail: johnathancalle@itm.edu.co
[4] Facultad de Minas, Universidad Nacional de Colombia, Medellín, Colombia
e-mail: cmzapata@unal.edu.co

4.1 Introducción

Los dominios de software científico (DSC) comprenden aplicaciones de software con un alto componente de conocimiento científico, incluyendo vocabulario, funcionalidad y elementos comunes de contextos científicos. Estas aplicaciones se basan en modelos matemáticos, métodos numéricos y fenómenos físicos que buscan facilitar la investigación en diferentes campos, por ejemplo: las matemáticas, ingeniería, medicina, química, entre otros [1]. En particular, las aplicaciones desarrolladas en este contexto se pueden clasificar en tres categorías: (i) software científico industrial, enfocado en atender las necesidades de procesos complejos desarrollado por un grupo de investigación empresarial; (ii) software científico de investigación, que busca apoyar la investigación de un área científica específica y se limita a un grupo pequeño de investigación académico; y (iii) software científico a pequeña escala, desarrollado en el contexto de un proyecto con un alcance mayor. En general, las personas que participan en el desarrollo de software científico no cuentan con una capacitación formal en las actividades que se deben llevar a cabo para implementar soluciones que sigan los lineamientos que propone la ingeniería de software. Sin embargo, las personas adquieren estas habilidades a través de un proceso de aprendizaje continuo, que permite a las partes involucradas abstraer los elementos necesarios y suficientes para llevar a cabo su solución [2]. Gracias a ello, los científicos pueden adquirir habilidades en software y profesionales en software pueden adquirir habilidades científicas. Con el paso de los años, el campo de la ingeniería de software ha realizado esfuerzos en el desarrollo de estrategias que faciliten el entendimiento y aplicación de software en diferentes dominios mediante la definición de técnicas de modelado, por ejemplo: diagramas de lenguaje unificado de modelado (UML) [3] o modelos para representar procesos de negocio (BPMN, EPC) [4], [5]. En particular, estos modelos incluyen operadores lógicos y relacionales simples para caracterizar los dominios de software. Sin embargo, estas técnicas buscan representar soluciones agnósticas, limitándose a representar vistas de comportamiento o estructurales, dejando de lado aspectos específicos del negocio. Como resultado, se observa que las estrategias de modelado no incluyen estructuras relacionales y/o lógicas complejas que son propias de los DSC, por ejemplo: expresiones lógicas matemáticas complejas o vocabularios propios del contexto científico. De acuerdo con lo anterior, surge la necesidad de proponer o estudiar la posibilidad de implementar un modelo de representación del conocimiento que integre una vista comportamental y estructural de los DSC, el cual se apoye en un lenguaje común que permita aplicar el conjunto de mejores prácticas propuestas desde el punto de vista de la ingeniería de software y del desarrollo de software científico. Para llevar a cabo esta tarea, se considera pertinente considerar el uso de un concepto conocido como Esquemas preconceptuales (EP). Los EP son modelos de representación del conocimiento que permiten caracterizar un dominio de software mediante la comprensión del conocimiento entre las partes interesadas y el analista para traducir el conocimiento en requerimientos explícitos que se puedan caracterizar en productos de software. Los EP integran estructuras gráficas, lingüísticas y matemáticas de la vista comportamental y estructural de un dominio de software en un mismo modelo desde su notación. Para lograrlo, los EP

se basan en el uso de lenguaje natural y el discurso de las partes interesadas, lo que facilita el modelado y validación de dominios complejos. Estos EP se han usado en dominios académicos e industriales lo que puede ser un buen referente para trasladar su uso a los dominios científicos. La notación de los EP se divide en cuatro grupos [6]: (i) nodos: concepto, condicional, referencia, operador y clase-concepto; (ii) relaciones: relación estructural, relación dinámica, relación de logro y relación eventual [7]; (iii) enlaces: implicación, concepto-nota, operadores y conjunción/disyunción; y (iv) aglutinadores: marco, nota, especificación, restricción y evento. Considerando todos los aspectos mencionados anteriormente, este artículo propone el uso de los EP para apoyar las etapas iniciales en la implementación de un DCS (análisis y diseño). Al aplicar los EP, se espera proveer un lenguaje común entre analistas de software y científicos que fomente el aprendizaje continuo de desarrolladores científicos en torno al modelado, desarrollo y validación de sus soluciones desde la perspectiva de la ingeniería de software [1]. El documento está estructurado de la siguiente manera: en la sección 4.2 se presenta la revisión de literatura y el problema; en la sección 4.3 se proponen las estructuras matemáticas de los EP para el aprendizaje continuo de DSC. En la sección 4.4 se valida la comprensión de las estructuras matemáticas de los EP. Finalmente, se discuten las conclusiones y el trabajo futuro.

4.2 Antecedentes

En general, los científicos subestiman la importancia de la ingeniería de software en el desarrollo de sus propias aplicaciones [8], enfocando sus esfuerzos principalmente en la etapa de codificación, dejando de lado las etapas de diseño de la solución [9]. Como resultado, las soluciones propuestas en contextos científicos son difíciles de entender, mantener y extender a lo largo del tiempo, dificultando la reutilización y escalabilidad del software para fines científicos [10]. Con el propósito de fomentar el uso de buenas prácticas en el contexto de los DCS, se han desarrollado diferentes estrategias, por ejemplo: análisis de negocio utilizando modelado UML, BPMN y EPC para representar dominios de software [7]. UML es un lenguaje semiformal con diagramas estructurales y de comportamiento para analizar, diseñar e implementar sistemas basados en software (máquinas de estados, secuencias y actividades, entre otros). BPMN [4] es una técnica de modelado utilizada para describir procesos y lógica de negocio apoyado en estructuras lógicas (and, or, entre otros) y estructuras relacionales (¡, ¿, =, !=, entre otros) [11]. Finalmente, EPC es un modelo utilizado para representar procesos de control de flujo [5]. De acuerdo con lo anterior, se observa que las técnicas de modelado pueden ser una herramienta útil para representar diferentes dominios. No obstante, al ser diseñados para solucionar problemas en el contexto de la ingeniería de software, estos modelos carecen de las estructuras matemáticas y el vocabulario utilizado para representar problemas en los DSC [12]. Además, el formalismo de estos modelos requiere experiencia en ingeniería de software para entenderlos y usarlos, dificultando su uso para comunicar y validar un DSC [12], [13]. A pesar de estas limitaciones, las técnicas de modelado son utilizadas en el DSC

como una técnica complementaria a los modelos científicos, por ejemplo: diagramas de bloques, modelos de Markov, redes neuronales, entre otros, que pueden incluir ecuaciones matemáticas y estructuras complejas [7]. Sin embargo, estas soluciones (en principio híbridas) obligan al científico o analista a crear múltiples diagramas para representar las características de su solución [7], [14]. En este sentido, se considera pertinente definir un modelo general que provea una estructura homogénea que facilite el diseño de software científico y permita reducir la brecha que impide realizar un análisis exhaustivo en el DSC desde el punto de vista del análisis funcional [7], [9], [15]. Para lograrlo, este artículo propone utilizar los EP [6] como un mecanismo de apoyo para representar un DSC. En particular, el uso de los EP como estrategia de integración responde a varios motivos: (i) son modelos que incluyen estructuras matemáticas, (ii) facilitan la comunicación entre los analistas de negocio y los stakeholders, e (iii) incluye un vocabulario común entendimiento cercano al lenguaje natural que facilita la comprensión y comunicación de un dominio específico [6].

4.3 Representación de DSC: un aprendizaje continuo a partir de estructuras matemáticas de los EP

En la Fig. 4.1 se muestra el flujo de actividades propuesto para facilitar la integración entre científicos y analistas de negocio. En las siguientes secciones se presenta el detalle de cada una de las etapas descritas en la Fig. 4.1.

Fig. 4.1: Representación del proceso de diseño.

4.3.1 Identificación de estructuras matemáticas para representar un DSC

En esta etapa se identifican todos los elementos representables utilizando un EP: enlaces, nodos (concepto, condicional, variable, parámetro, arreglo, entre otros), operadores, y aglutinadores (valor-nota y condiciones iniciales) [7], [14], [16].

4.3.2 Definición de reglas para el aprendizaje continuo de DSC mediante la representación de estructuras matemáticas de los EP

Regla 1: Cuando se define un vocabulario usando variables, funciones y parámetros, se deben representar en su forma conceptual; es decir, su nombre completo debe ser entendible, por ejemplo: la variable "y" representa el "valor de la población" (ver la Fig. 4.2). Los corchetes son utilizados para representar arreglos y ecuaciones matemáticas, por ejemplo: la Ecuación 4.3.2 transformada a la Ecuación 4.3.2 (ver la Fig. 4.2) y la Ecuación 4.3.2 transformada en la Ecuación 4.3.2 (ver la Fig. 4.3) se representan gráficamente usando un árbol binario. **Regla 2:** Cuando una

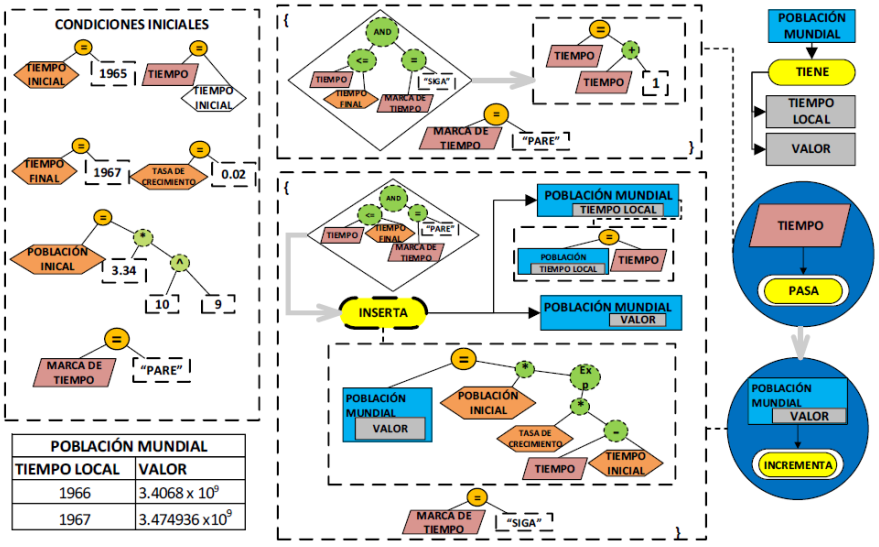


Fig. 4.2: Representación del crecimiento de la población mundial con un EP (Fuente propia).

funcionalidad se representa usando procesos que ejecutan conceptos animados que incluyen ecuaciones, estas ecuaciones se deben representar dentro de una especificación conectada a una relación dinámica, por ejemplo: "Biólogo Clasifica Átomo"

(ver la Fig. 4.3). **Regla 3:** Cuando una funcionalidad se representa usando procesos automáticos que incluyan ecuaciones, estas ecuaciones se deben representar dentro de una especificación conectada con un evento, por ejemplo: "Inserta Valor de Población" en el evento disparador (ver la Fig. 4.2).

$$y(t) = y_0 e^{r(t-t_0)} \tag{4.1}$$

$$Poblacion.valor = Poblacion\_inicial \times e^{tasa\_de\_crecimiento(tiempo - tiempoInicial)} \tag{4.2}$$

$$Circulo.area = Pi \times Circulo.Radio^2 \tag{4.3}$$

$$Rectangulo.Diagonal = \sqrt{Rectangulo.Base^2 + Rectangulo.Altura^2} \tag{4.4}$$

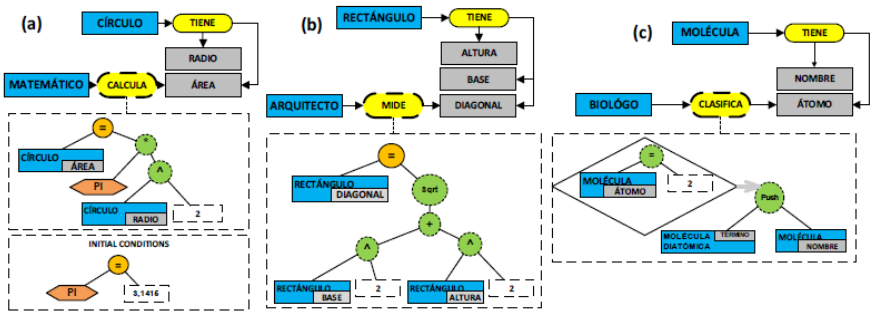


Fig. 4.3: Representación de procesos realizados por científicos con un EP (Fuente propia).

4.3.3 Representación del vocabulario y funcionalidad de DSC mediante EP

Aprendizaje continuo en procesos automáticos desde DSC. Un modelo matemático de crecimiento de la población mundial se representa en la Fig. 4.2 con estructuras

matemáticas del EP desde el dominio de estadística. El modelo se representa mediante la Ecuación diferencial 4.3.2 para obtener el valor incremental de la población mundial en varios años como un proceso automático en este dominio. Para la representación, se traduce la Ecuación 4.3.2 a su forma conceptual utilizando la **Regla1** en la Ecuación 4.3.2 y después, se representa gráficamente con la notación de los EP en la Fig. 4.2. La especificación de las condiciones iniciales permite establecer los valores iniciales de los parámetros (tiempo inicial = 1965 y tiempo final = 1967, tasa de crecimiento = 0.02 y población inicial = 3.34 x 109) y de las variables (tiempo = tiempo inicial y marca de tiempo = "siga", la marca de tiempo se usa para determinar el inicio y el fin de un evento). El flujo del sistema inicia con el evento "tiempo pasa", el cual permite controlar el tiempo del sistema. La especificación del evento "tiempo pasa" se usa para asignar el valor del tiempo en la variable independiente "tiempo" incrementando en '1' después de cada ejecución mediante la operación de suma (si tiempo ¡= tiempo final y marca de tiempo = "siga"). El valor se determina en años, es decir,1965, 1966, etc. Cuando el tiempo cambia, la "marca de tiempo" es "pare" para guardar la información. El evento "valor de la población mundial incrementa" se dispara por la condición: tiempo ¡= tiempo final y marca de tiempo = "pare". Las relaciones dinámicas "insertar tiempo local" y "valor de la población mundial" dependen del "tiempo" para insertar un nuevo valor; por ello, cuando un valor se inserta la "marca de tiempo" cambia a "siga". La relación estructural población "mundial tiene tiempo local" y "valor" indica que "población mundial" es una clase mientras que "tiempo local" y "valor" son conceptos hoja en el vocabulario del dominio. Las relaciones estructurales también se pueden representar con el símbolo de concepto-clase, por ejemplo: el concepto-clase "valor de la población mundial". Las condiciones iniciales se requieren en el funcionamiento del proceso. En el concepto "población mundial" se pueden observar los datos que se insertan automáticamente.

Aprendizaje continuo en procesos realizados por científicos en DSC: Como se ha mencionado anteriormente, el aprendizaje continuo es clave para la sostenibilidad en el desarrollo de software, permitiendo a los equipos actualizar y mejorar constantemente sus conocimientos y habilidades; utilizar prácticas de ingeniería de software a través del uso de los EP facilita la comprensión, modificación y mejora de un DSC. En este contexto y para ilustrar la propuesta se representa el proceso "matemático calcula área del círculo" en la Fig. 4.3 (a) mediante la especificación de la Ecuación 4.3.2 en la relación dinámica "calcula". En este EP, se usa también una condición inicial para describir el parámetro PI = '3,1415' que se usa en la ecuación, para calcular el área del círculo. Se incluye la relación estructural "círculo" (clase) tiene "radio" y "área" (conceptos hoja) como vocabulario del dominio. Además, se representa el proceso "arquitecto mide diagonal del rectángulo" en la Fig. 4.3 (b), que incluye la especificación del proceso con la Ecuación 4.3.2 utilizando el operador "Sqrt" para representar la raíz cuadrada de la ecuación. Se representa la relación estructural "rectángulo" tiene "base", "altura" y "diagonal" para indicar su relación con los conceptos hoja en el vocabulario del dominio. Finalmente, se representa el proceso biólogo clasifica átomos en la Fig. 4.3 (c). El concepto "molécula" se clasifica reconociendo su concepto hoja "átomo" en el condicional si "molécula átomo

= 2". Si la condición se cumple, entonces el concepto hoja "nombre" se almacena utilizando el operador "Push" en la lista "molécula diatómica".

4.4 Validación

Para validar la comprensión de las estructuras matemáticas de los EP en DSC, se empleó una encuesta como instrumento de recolección de datos. El grupo focal estuvo compuesto por 21 participantes: 7 profesionales con posgrado, 5 profesionales, 8 universitarios en áreas científicas y de software, y 1 estudiante de bachillerato sin conocimiento previos en EP. La validación tuvo tres momentos: (i) Identificación de la Ecuación 4.3.2 junto con sus condiciones iniciales, sin usar notación EP: el 57, 7% no pudo identificar la relación de la ecuación, mientras que el 42.9% respondió afirmativamente, relacionando la ecuación con una función exponencial, aunque sin precisar el dominio del uso de la función. (ii) Interpretación del dominio usando notación matemática EP: el 90.4% proporcionaron respuestas relacionadas con el crecimiento de la población mundial a lo largo del tiempo, mientras que, el 9.6% no lograron interpretar el dominio utilizado con la ecuación. Por último (iii) Evaluación de comprensión de la notación matemática EP (véase la Fig. 4.4): los participantes describieron el dominio y evaluaron su comprensión, la descripción precisa del dominio realizada por la mayoría de los participantes, junto con el alto porcentaje de respuestas que indicaron una mayor facilidad de comprensión del EP, sugieren que hubo una comprensión adecuada del dominio. Esto implica que los interesados en aplicaciones científicas podrían también comprender la representación al utilizar un lenguaje común con los científicos y analistas de negocio. Las limita-

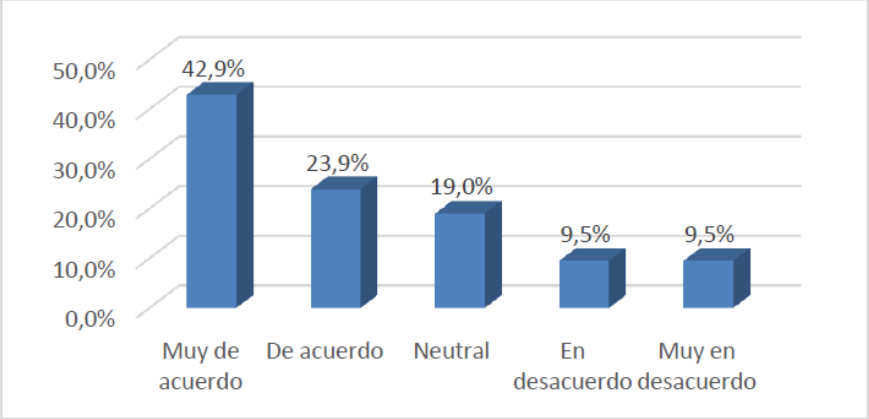


Fig. 4.4: Evaluación de comprensión de la notación matemática EP (Fuente propia).

ciones de esta validación incluyen el tamaño del grupo focal (21 participantes) y su

variada formación académica, lo que puede afectar la representatividad. El uso de encuestas puede introducir sesgos de respuesta y malinterpretaciones. La evaluación se realizó en un entorno controlado, no necesariamente reflejando otros contextos. La subjetividad en la interpretación de datos, la variabilidad en el tiempo dedicado a las encuestas y la falta de familiaridad con la notación de EP también pueden influir en la validez de los hallazgos.

4.5 Conclusiones y Trabajo Futuro

Los DSC son campos de ingeniería, matemáticas y ciencias donde se desarrollan aplicaciones científicas. Los EP son modelos de representación del conocimiento que vinculan estructuras gráficas, lingüísticas y matemáticas mediante elementos dinámicos y estructurales para una representación completa de cualquier dominio. En este artículo, se propuso el uso de estructuras matemáticas y gráficas de los EP. Estas estructuras permiten un aprendizaje continuo de prácticas en ingeniería de software para software científico mediante la representación completa del vocabulario y la funcionalidad de DSC como: ecuaciones matemáticas, procesos autónomos y procesos que realizan los científicos, las cuales se basan en la notación matemática y elementos presentes en estos dominios. En adición, se incluyen reglas para el uso de estas estructuras en la representación del vocabulario y la funcionalidad de DSC. Finalmente, se presentan los EP como un modelo de representación del conocimiento que permite validar DSC previo a su desarrollo. La mayoría de los participantes en la validación comprende el dominio y sus elementos a partir del EP. De este modo, los científicos, desarrolladores y analistas pueden tomar ventaja de los EP para integrarlo como una herramienta que les permite un aprendizaje continuo de DSC mediante su lenguaje común para modelar el dominio y validarlo en etapas tempranas del desarrollo de aplicaciones científicas. El aprendizaje continuo mediante los EP permite reducir la brecha entre la ingeniería de software y la ciencia permitiendo el modelado, validación y codificación de dominios complejos. Se sugiere como trabajo futuro el modelado y automatización de procesos en DSC mediante EP para la producción de aplicaciones científicas. También, se podría considerar la realización de un ejemplo extendido para comparar con otras propuestas, lo cual enriquecería el estudio.

Agradecimientos Este artículo es un producto de la Tesis Doctoral Una extensión a los esquemas preconceptuales para el refinamiento en la representación de eventos y la notación matemática y del proyecto Un modelo de sostenibilidad humana para productos de software con infraestructura tecnológica limitada.

Referencias

1. Kelly, D.: Scientific software development viewed as knowledge acquisition: Towards understanding the development of risk-averse scientific software. Journal of Systems and Software 109, 50–61 (2015)
2. Noreña, P. A., Suescún, E., González, L.: Caracterización de prácticas sostenibles en el desarrollo de software y la sostenibilidad humana usando Essence. In: CISTI (2024)
3. Object Management Group: OMG Unified Modeling Language TM (OMG UML) Superstructure 2.5. http://www.omg.org/spec/UML/2.5/ (2015)
4. Haisjackl, C., Soffer, P., Lim, S. Y., Weber, B.: How do humans inspect BPMN models: an exploratory study. Software & Systems Modeling 17(2), 655–673 (2018)
5. Shuo, L., Haibo, S., Xisheng, L.: Based event model application research on configurable integrated platforms. In: Proceedings of 2011 International Conference on Modelling, Identification and Control, pp. 1–5 (2011)
6. Zapata, C. M.: The UNC-Method revisited: elements of the new approach. Lambert, Saarbrucken (2012)
7. Noreña, P. A.: An extension to Pre-conceptual schemas for refining Event Representation and Mathematical Notation. Tesis Doctoral, Universidad Nacional de Colombia, Medellín, Colombia (2020)
8. Wilson, G., Aruliah, D. A., Brown, C. T., Hong, N. P. C., Davis, M., Guy, R. T., Wilson, P.: Best practices for scientific computing. PLoS Biol 12(1), e1001745 (2014)
9. Arvanitou, E. M., Ampatazoglou, A., Chatzigeorgiou, A.: Software engineering practices for scientific software development: A systematic mapping study. The Journal of Systems & Software 172, 1–13 (2021). https://doi.org/10.1016/j.jss.2020.110848
10. Michailidis, P. D., Margaritis, K. G.: Scientific computations on multi-core systems using different programming frameworks. Applied Numerical Mathematics 104, 62–80 (2016)
11. Kanewala, U., Bieman, J. M.: Testing scientific software: A systematic literature review. Information and Software Technology 56(10), 1219–1232 (2014)
12. Morris, C., Segal, J.: Lessons learned from a scientific software development project. IEEE Software 29(4), 9–12 (2012)
13. Ahmed, Z.: Essential Design Modeling for Scientific Software Solutions Development. Peerj 3, e1423v1 (2015)
14. Calle, J. M.: Identificación de patrones de diseño para software científico a partir de esquemas preconceptuales. Tesis de Maestría, Universidad Nacional de Colombia, Medellín, Colombia (2016)
15. Hannay, E., MacLeod, C., Singer, J., Langtangen, H. P., Pfahl, D., Wilson, G.: How Do Scientists Develop and Use Scientific Software? In: SECSE'09, p. 8. IEEE Computer Society, Vancouver (2009)
16. Noreña, P. A., Zapata, C. M.: Simulating Events in Requirements Engineering by Using Preconceptual-Schema-based Components from Scientific Software Domain Representation. Advances in Systems Science and Applications 21(4), 1–15 (2022)

5

DockerWizard: Optimización del Servidor de Investigaciones de la UNIAJC Colombia mediante la Automatización de la Gestión y Despliegue de Contenedores Docker

Alexis Alberto Ramirez Orozco[1][0009-0004-1448-0890]
Santiago Ruiz Posso[2][0009-0006-9527-5438]
Carlos Lino Rengifo Renteria[3][0000-0003-3550-9707]

Resumen The engineering faculty's research server at UNIAJC University facilitates research projects and student theses that necessitate web services and databases. These projects must be packaged in Docker containers, with the server employing a Docker reverse proxy to enable concurrent project execution. Nonetheless, the manual management of containers through the server terminal results in configuration errors, extended service disruptions, monitoring challenges, and complex user administration, thereby diminishing efficiency and causing delays. To address these challenges, an integrated solution was developed to automate and optimize the management of Docker containers. This operational solution enhances efficiency and minimizes errors through features including an intuitive graphical user interface, automated container creation and configuration, and real-time monitoring capabilities. **Palabras clave:** Docker, Automation, Security.

5.1 Introducción

Antes del surgimiento de Docker, la virtualización era la principal tecnología utilizada para empaquetar y ejecutar aplicaciones de manera aislada [4]. Las máquinas virtuales (VM) permitían a los desarrolladores y administradores de sistemas aislar aplicaciones y entornos, evitando conflictos entre diferentes dependencias y configuraciones [11]. Sin embargo, las VMs eran pesadas y requerían la instalación de un sistema operativo completo, lo que implicaba un uso ineficiente de recursos [14]. Docker revolucionó el panorama de la virtualización al introducir los contenedores, unidades más ligeras y eficientes que comparten el núcleo del sistema operativo host, mejorando significativamente el desempeño y la portabilidad de las aplicaciones, asegurando coherencia entre diferentes entornos de desarrollo y producción [5],[13]. Las plataformas de gestión de Docker simplifican la administración y orquestación de contenedores a través de interfaces gráficas de usuario amigables [10]. Estas herramientas permiten a los usuarios desplegar, supervisar y escalar contenedores sin tener que interactuar directamente con la línea de comandos, mejorando así la accesibilidad y eficiencia operativa [1],[2]. Estas plataformas también proporcionan funcionalidades avanzadas, como supervisión en tiempo real, gestión de volúmenes y redes, así como la integración con sistemas de autenticación para el control de acceso. Dichas herramientas son fundamentales para simplificar la complejidad de los entornos de producción a gran escala, garantizando una

[1,2,3]Institución Universitaria Antonio José Camacho Santiago de Cali, Av 6a Nte.#28N-102, COL
e-mail: \{aramirez,clino\}@admon.uniajc.edu.co
e-mail: santiagoruiz@estudiante.uniajc.edu.co

administración más eficaz de los recursos [15]. La infraestructura del servidor de investigación de la Facultad de Ingeniería de la Institución Universitaria Antonio José Camacho (UNIAJC) se creó con el propósito de dar soporte a proyectos de investigación y trabajos de grado que precisan de servicios de red y bases de datos [9]. Para acceder a dichos servicios, se requiere que los proyectos se desplieguen utilizando contenedores Docker. La actual infraestructura del servidor incluye un proxy inverso Docker que permite la ejecución simultanea de proyectos. No obstante, la creación manual de contenedores aumenta la probabilidad de errores de configuración y afecta el desempeño del proyecto, a la vez que demanda tiempos de inactividad prolongados, generando retrasos en el avance de los proyectos y afectando la disponibilidad del servicio. Además, la gestión manual de usuarios dificulta la colaboración de los participantes. En respuesta a estas problemáticas, se ha implementado una solución integral para automatizar y mejorar la administración de contenedores Docker en el servidor. Esta solución, actualmente en funcionamiento, mejora la eficiencia y disminuye los errores de configuración.

5.2 Metodología

La metodología empleada en este proyecto se enfocó en la identificación y resolución de las deficiencias detectadas en la gestión manual de contenedores Docker en el servidor de investigaciones de la Facultad de Ingeniería de la UNIAJC. En un principio, se realizó un análisis de la situación actual para detectar problemas críticos como errores de configuración, interrupciones prolongadas del servicio y dificultades en el monitoreo, que tenían un impacto negativo en la eficiencia operativa. A partir de este diagnóstico, se diseñó e implementó un sistema de automatización llamado DockerWizard (DW) con el objetivo de optimizar la creación y gestión de contenedores, empleando tecnologías como Angular para la interfaz gráfica y Node.js para los scripts de automatización, a continuación, se detalla la metodología desarrollada.

5.2.1 Análisis de la situación Actual

Inicialmente, se realizó un análisis de requisitos y del sistema de la configuración y administración que se tenía de los contenedores Docker. Este proceso reveló varios problemas fundamentales, entre ellos la creación manual de contenedores, configuraciones propensas a errores, interrupciones prolongadas del servicio, dificultades en el monitoreo y errores humanos. Este diagnóstico fue fundamental para comprender las necesidades y áreas que requerían mejora.

5.2.2 Diseño del sistema de Automatización

Después de identificar las necesidades y deficiencias en la etapa de análisis, se diseñó un sistema de automatización con el objetivo de mejorar la creación y gestión de contenedores Docker. Teniendo como premisa la reducción de errores y mejora la eficiencia operativa. El diseño del sistema siguió las mejores prácticas recomendadas por [12], como la modularidad, la reutilización de componentes, la escalabilidad y la facilidad de mantenimiento. Además, se priorizó el diseño de una interfaz de usuario amigable, la automatización de la configuración de contenedores y un sistema de monitoreo en tiempo real para garantizar una gestión eficiente y confiable de los recursos.

5.2.3 Desarrollo e implementación del sistema de Automatización

Para el desarrollo de la aplicación de automatización de contenedores, se seleccionaron herramientas y tecnologías apropiadas para la gestión de contenedores Docker. Angular fue utilizado para la interfaz visual, mientras que Node.js se empleó para los scripts de shell, facilitando así la creación y configuración de los contenedores. Además, se implementaron algoritmos específicos para el despliegue de los contenedores, con el fin de garantizar configuraciones consistentes y exentas de errores. En este contexto, se establecieron procedimientos automatizados para el despliegue de nuevos contenedores Docker (ver Fig. 5.1). Esto incluyó la definición de imágenes específicas para cada tipo de servicio requerido y la monitorización de las imágenes personalizadas. Además, el proceso de creación de contenedores se automatizó a partir de las imágenes definidas, utilizando scripts que reducen la intervención manual. Los contenedores creados se configuraron automáticamente y se sometieron a diversas pruebas para garantizar su correcto funcionamiento y compatibilidad con los servicios requeridos.

5.2.4 Implementación de medidas de Seguridad

La seguridad es un aspecto crítico durante todo el proceso [6],[7]. Siguiendo la guía del estándar NIST SP 800-190 [8], se implementaron varias medidas y ajustes de seguridad con el propósito de asegurar la integridad del sistema. Se garantizó que las imágenes Docker utilizadas fueran seguras y estuvieran actualizadas, evitando el uso de imágenes no verificadas o con vulnerabilidades. Además, se configuraron políticas de control de acceso estrictas, asegurando que solo usuarios autorizados pudieran gestionar los contenedores Docker. Se implementaron medidas apropiadas de aislamiento de contenedores para restringir el acceso solamente a usuarios con la debida autorización. También se estableció un procedimiento para aplicar periódicamente actualizaciones y parches de seguridad en los contenedores y el sistema operativo del servidor. Se implementaron herramientas de monitoreo para detectar actividades sospechosas y responder rápidamente a posibles incidentes.

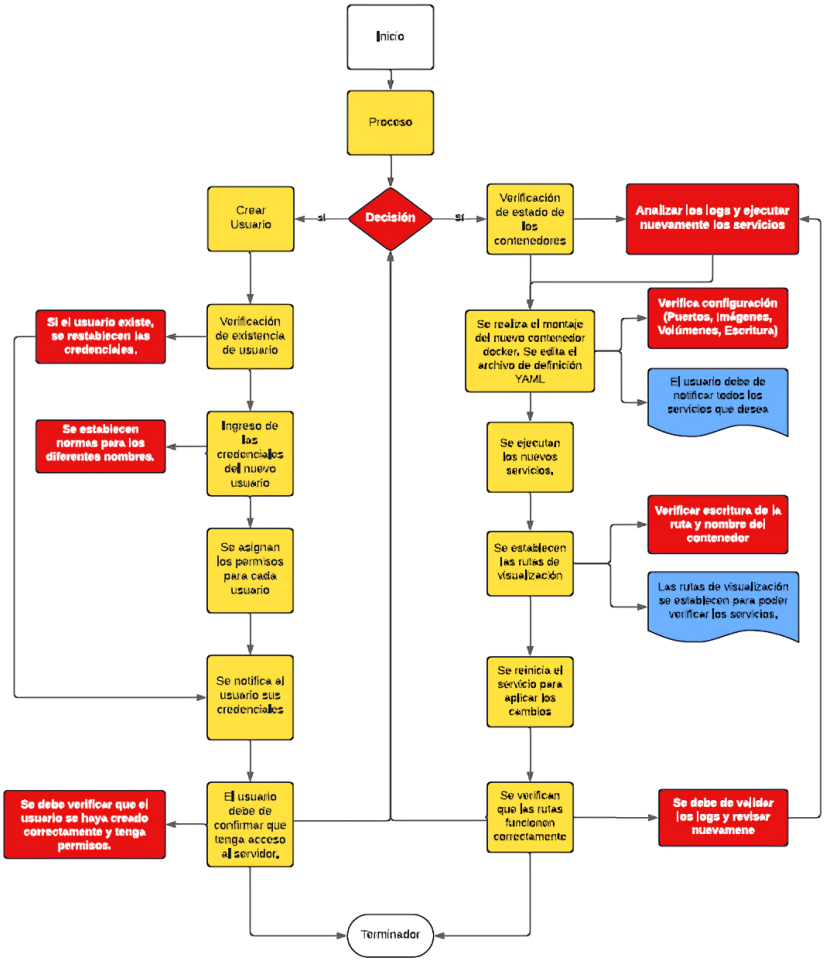


Fig. 5.1: Diagrama de flujo de despliegue de contenedor. (Fuente propia)

5.2.5 Evaluación y mejora continua

Finalmente, se estableció un ciclo de evaluación y mejora continua basado en la retroalimentación obtenida durante el funcionamiento del sistema automatizado. Este ciclo permitió ajustar y optimizar los algoritmos de despliegue, los procedimientos de monitorización y las medidas de seguridad, asegurando una gestión más eficiente y fiable de los contenedores Docker a lo largo del tiempo. En última instancia, se dispuso de un ciclo de evaluación y mejora continua fundamentado en la retroalimentación obtenida durante el funcionamiento del sistema automatizado. Este ciclo permitió ajustar y optimizar los algoritmos de despliegue, los procedimientos de monitoreo y las medidas de seguridad, garantizando una gestión más eficiente y confiable de los contenedores Docker a lo largo del tiempo.

5.3 Resultados

El proyecto se centró en mejorar la experiencia de uso de los investigadores y estudiantes y la gestión del servidor de investigaciones de la Facultad de Ingeniería de la UNIAJC. A continuación se presentan los hallazgos y logros del proyecto.

5.3.1 Creación de usuarios y Asignación de permisos

Para garantizar el acceso y gestión de los contenedores Docker por parte de usuarios autorizados, se implementó un sistema de creación y verificación de usuarios. Este sistema verifica si el usuario ya está registrado en el sistema; en caso contrario, se procede a crear el usuario y asignarle las credenciales correspondientes. Además, se establecieron permisos específicos para cada usuario, asegurando que cuenten con los privilegios adecuados para sus funciones, asignando si se requiere las credenciales. La creación de usuarios es un proceso crítico en la gestión de proyectos de

Table 5.1: Comparación de tiempos y eficiencia entre plataformas

Plataforma	Tiempo	Eficiencia
DockerWizard (Vm)	10,65 s	70%
Shell (Vo)	35,5 s	
DockerWizard (Vm)	11 s	63%
Shell (Vo)	30,07 s	
DockerWizard (Vm)	12 s	53%
Shell (Vo)	25,56 s	

investigación. Para medir la eficiencia de este proceso se procedió a comparar el tiempo que tarda un administrador en crear un usuario en usando el interprete de comandos o shell con respecto a el tiempo que se tarda haciendo la misma tarea usando DW. Para realizar el cálculo de la eficiencia se usa la fórmula de análisis de rendimiento [3].

$$\left(\frac{V_0 - V_m}{V_0}\right) \times 100 \tag{5.1}$$

Donde V_0 es el tiempo original o que tarda un administrador realizando un despliegue con el shell. V_m es el tiempo modificado, es decir, el tiempo que se obtiene al realizar un despliegue desde la interfaz web de DW. Como se puede detallar en la Tabla 5.1 se llega a la conclusión que hay una mejoría en el despliegue de nuevos usuarios para el servidor, además la asignación de credenciales y permisos de escritura y lectura de su espacio personal de trabajo y permiso de acceso al servidor. Esta eficiencia resulta en un promedio del 62% de eficiencia, este porcentaje es favorable ya que permite evidenciar que el entorno gráfico para el despliegue de usuario cumple con su objetivo y ayuda con el despliegue de nuevos usuarios. Esto será fundamental en el momento que se requiera desplegar múltiples usuarios en un mismo requerimiento.

5.3.2 Despliegue Automatizado de contenedores Docker

Se implementó un sistema automatizado para el despliegue y configuración de contenedores Docker. Este sistema utilizó herramientas como Angular para las interfaces gráficas y Node.js para los scripts de shell. El proceso incluyó la definición de imágenes Docker específicas para cada servicio necesario, así como la automatización del proceso de ejecución de los contenedores a partir de estas imágenes, minimizando la intervención manual. Además, los contenedores generados fueron configurados de manera automática y sometidos a pruebas para asegurar su correcto funcionamiento e integración con los servicios requeridos.

5.3.3 Montaje de contenedores y configuración de servicios

Utilizando archivos de definición YAML, se automatizó el montaje de nuevos contenedores y la configuración de rutas necesarias para la visualización y operación de los servicios. Como se pude observar en la Tabla 5.2, se llevaron a cabo dos tipos de pruebas para medir el rendimiento en el despliegue de los servicios Docker. Estas pruebas se realizaron en dos entornos diferentes; el primer entorno se utilizó DW y en el segundo se utilizó el shell del servidor, Se desplegaron dos tipos de servicios, un servicio web únicamente y un servicio web con base de datos MySQL y PhPMyadmin incluido. A partir de estas pruebas, se concluyó que existe una diferencia significativa en el tiempo de despliegue y creación de servicios, obteniendo una eficiencia promedio del 81%. Este resultado permite afirmar que la optimización de los contendores Docker a través del entorno de DW mejora el rendimiento en cuanto a tiempos y reduce la probabilidad de errores humanos. Para la Tabla 5.2 se realizó una prueba de despliegue de servicios, con el objetivo de medir el nivel de rendimiento de la interfaz gráfica de DW. Resultando en un promedio de rendimiento de 80%, se aplicó la misma formula utilizada para la prueba mostrada en la Tabla 5.1.

Table 5.2: Comparación de tiempos, objetivos y eficiencia entre plataformas

Plataforma	**Tiempo (s)**	**Objetivo**	**Observaciones**	**Eficiencia**
DW (V_m)	30, 76	Creación de Docker de solo servicio Web	NA	82.83%
Shell (Vo)	179, 15		NA	
DW (V_m)	37, 76	Creación de Docker de solo servicio Web	Se tardó en subir el servicio	63.28%
Shell (V_o)	119,05		Se tardó en salir el servicio y se realizó el creador web. Se corrigió error creando red y PHP	
DW (Vm)	25, 94	Creación de Docker de solo servicio Web	NA	88.65%
Shell (Vo)	217, 15		NA	
DW (Vm)	74, 48	Creación de Docker con 3 servicios web y phpMyAdmin	NA	88.46%
Shell (Vo)	655, 54		Se usó con el editor VSCODE	

5.3.4 Monitoreo y medidas de seguridad

La monitorización continua y la implementación de medidas de seguridad fueron fundamentales para asegurar la operatividad y protección del servidor. Se desplegaron herramientas para la detección temprana de fallos y el registro de eventos. La seguridad de los contenedores se priorizó mediante la supervisión de las imágenes Docker, así como la aplicación periódica de actualizaciones y la configuración de políticas de acceso estrictas. Además, se garantizó un aislamiento adecuado de los contenedores, de modo que solo los usuarios autorizados pudieran acceder y gestionarlos. Así mismo, se estableció un procedimiento para aplicar regularmente actualizaciones y parches de seguridad en los contenedores y en el sistema operativo del servidor.

5.4 Conclusión

La implementación de una solución integral para la automatización y optimización de la administración de contenedores Docker en el servidor de investigación de la facultad de ingeniería UNIAJC, ha mejorado significativamente la eficiencia operativa, reduciendo errores de configuración y tiempos de inactividad. Las características principales, como contenedores, y un sistema de monitoreo en tiempo real, han permitido un manejo seguro de los recursos. Además, la adopción de medidas de seguridad rigurosas, el control de acceso y el aislamiento de contenedores garantiza la protección del servicio y datos, la evaluación continua y el ajuste de los algoritmos de despliegue y procedimientos de seguridad aseguran una gestión sostenible y fiable a lo largo plazo. Esta solución supera ineficiencias anteriores, facilitando el progreso de los proyectos de investigación y de grado, y asegurando la disponibilidad y calidad de los servicios esenciales para la comunidad académica de la facultad.

Acknowledgements Los autores desean expresar su más sincero agradecimiento a la Institución Universitaria Antonio José Camacho y al decanato de investigaciones por su valioso apoyo en los procesos de investigación. Además de la colaboración y el compromiso con los semilleros de investigación SAIDO, SIMEC y SELECT, de la facultad de ingenierías cuya participación ha sido fundamental para el desarrollo de esta investigación.

Referencias

1. Abhishek, M.K., Rao, D.R., Subrahmanyam, K.: Framework to deploy containers using kubernetes and ci/cd pipeline. International Journal of Advanced Computer Science and Applications 13(4) (2022).https://doi.org/10.14569/IJACSA.2022.0130460
2. Boza, E.F., Abad, C.L., Narayanan, S.P., Balasubramanian, B., Jang, M.: A case for performance-aware deployment of containers. In: Proceedings of the 5th International Workshop on Container Technologies and Container Clouds. p. 25–30. WOC '19, Association for Computing Machinery, New York, NY, USA (2019). https://doi.org/10.1145/3366615.3368355
3. Casalicchio, E., Perciballi, V.: Measuring docker performance: What a mess!!! In: Proceedings of the 8th ACM/SPEC on International Conference on Performance Engineering Companion. p. 11–16. ICPE '17 Companion, Association for Computing Machinery, New York, NY, USA (2017). https://doi.org/10.1145/3053600.3053605

4. Dhanse, S., Chandane, M.M., Chavan, A.: Dockomation system. In: 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT). pp. 1–5 (2023). https://doi.org/10.1109/ICCCNT56998.2023.10307180
5. Khan, K.: Mastering Docker: A Comprehensive Guide to Containerization (07-2023)
6. Mantilla, K.I., Flórez, A.: Development of a monitoring module for physical and virtual servers: Advanced computing center case of the universidad pontificia bolivariana, bucaramanga, colombia. Journal of Physics: Conference Series 1513(1), 012006 (mar 2020). https://doi.org/10.1088/1742-6596/1513/1/012006, https://dx.doi.org/10.1088/1742-6596/1513/1/012006
7. Mullinix, S.P., Konomi, E., Townsend, R.D., Parizi, R.M.: On security measures for containerized applications imaged with docker (2020), https://arxiv.org/abs/2008.04814
8. Murugiah Souppaya, John Morello, Karen Scarfone: Application container security guide. NIST Special Publication 800-190, National Institute of Standards and Technology (2017), https://csrc.nist.gov/publications/detail/sp/800-190/final, accessed:2024-08-11
9. Pastrana, M.A., Rojas, A.M., González, H.A.: Smart campus modelo y resultados (2020), https://repositorio.uniajc.edu.co/handle/uniajc/846
10. Piedade, B., Dias, J.a.P., Correia, F.F.: An empirical study on visual programming docker compose configurations. In: Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings. MODELS '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3417990.3420194, https://doi.org/10.1145/3417990.3420194
11. Potdar, A.M., Narayan, D., Kengond, S., Mulla, M.M.: Performance evaluation of docker container and virtual machine. Procedia Computer Science 171, 1419–1428 (2020)
12. Reis, D., Piedade, B., Correia, F.F., Dias, J.P., Aguiar, A.: Developing docker and docker-compose specifications: A developers' survey. Ieee Access 10, 2318–2329 (2021)
13. Schenker, G.: Docker: Up and Running: Build and deploy containerized web apps with Docker and Kubernetes (English Edition). Bpb Publications (2023), https://books.google.com.co/books?id=j7m6EAAAQBAJ
14. Sollfrank, M., Loch, F., Denteneer, S., Vogel-Heuser, B.: Evaluating docker for lightweight virtualization of distributed and time-sensitive applications in industrial automation. IEEE Transactions on Industrial Informatics 17(5), 3566–3576 (2021). https://doi.org/10.1109/TII.2020.3022843
15. Yepuri, V.K., Polamarasetty, V.K., Donthi, S., Gondi, A.K.R.: Containerization of a polyglot microservice application using docker and kubernetes (2023), https://arxiv.org/abs/2305.00600

6

Guía para la generación de un conjunto de datos en estudiantes sobre recursos educativos abiertos mediante reconocimiento facial de expresiones en un JETSON NANO

Miller D. Quilindo V.Heredia[1] [0009-0009-5780-7589]
Néstor M. Díaz M.[2] [0000-0002-6074-9764]
Carolina G. Serrano[3] [0000-0002-1216-4876]

Resumen Este artículo propone una guía para la recolección de emociones en un aula de clases, donde se especifica un modelo lógico de componentes y las tecnologías adecuadas para su implementación en un dispositivo de bajo costo. Utiliza técnicas de aprendizaje automático como el reconocimiento facial de expresiones (FER) basado en visión por computadora. El objetivo es estudiar imparcialmente los estados emocionales y cognitivos de los estudiantes para evaluar su experiencia con Recursos Educativos Abiertos (REA). Se genero un conjunto de datos a partir de 105 estudiantes de un colegio de la zona rural del suroccidente de Colombia, quienes interactuaron individualmente con los REA durante 15 minutos. El conjunto de datos se recopilo de manera anónima e incluye una variedad de medidas tales como la presencia de las siete emociones básicas de Ekman, la tasa de parpadeo, posición de la cabeza y el estado de compromiso.

Palabras clave: Redes Neuronales Convolucionales, Análisis de emociones, Series de tiempo, Jetson Nano.

6.1 Introduction

El término REA (Recurso Educativo Abierto) surgió en el año 2002 durante el foro de Software Educativo Abierto de la UNESCO. Este refiere a materiales destinados a la enseñanza, el aprendizaje e investigación, que pueden ser modificados y reutilizados sin restricción alguna por parte de terceros [1]. En las últimas dos décadas REA ha presentado un auge impulsado por la creación de diversos repositorios institucionales para sus publicaciones, sin embargo, el aumento de estos materiales ha generado cierto escepticismo entre profesionales de la educación debido a la dificultad para hallar aquellos que satisfagan criterios esenciales como: contenidos precisos y actualizados, objetivos de aprendizaje claros y un diseño instruccional apropiado [2]. Para evaluar los REA se han empleado técnicas como la clasificación de 5 estrellas, revisión por pares y aprendizaje de máquina (ML) basado en el lenguaje. Aunque estas estrategias funcionan correctamente, aún presentan problemáticas relacionadas con altos costos (tanto en tiempo como en dinero) y, la exclusión del impacto emocional y cognitivo que el recurso puede generar en el individuo, tal como se expone [3], fundamentado por la Teoría de Respuesta al Ítem (IRT). Investigaciones recientes han examinado la posibilidad de utilizar modelos de aprendizaje profundo (DL) como las CNN (Redes Neuronales

Universidad del Cauca, Popayan, Cauca, CO
e-mail: \{mdquilindo,nediaz,cgonzals\}@unicauca.edu.co

Convoluciones) para medir el estado emocional de los estudiantes al afrontar sesiones de clases o lecturas [4]. Varios estudios han tenido éxito en este proceso, logrando inferir emociones con una precisión de hasta el 72.02% [5] en escenarios controlados empleando máquinas con suficientes recursos computacionales. Sin embargo, la mayoría de los trabajos no consideran el uso de métodos automatizados para analizar los resultados. Antes de aplicar técnicas sofisticadas para el análisis de emociones, es necesarios contar con repositorios de datos fiables y completos. Por lo tanto, en el presente artículo se propone una guía para la generación de un conjunto de datos emocionales implementado en una serie de tecnologías livianas para su ejecución en un dispositivo de bajo costo Jetson Nano, así mismo como su visualización preliminar empleando técnicas de minería de datos para validar los resultados mediante la identificación de patrones y tendencias similares a las reportadas por la literatura.

6.2 Trabajos relacionados

En [6, 7] implementan técnicas FER (Reconocimiento Facial de Expresiones) junto con modelos DL [8] como las CNN para la inferencia de emociones donde alcanzan precisiones de hasta 65.7% y 62.02%, respectivamente. Los anteriores trabajos solo reciben como entrada a sus modelos una única imagen, por lo tanto, en [9] deciden experimentar en la implementación de modelos RNN (Red Neuronal Recurrente) como las LSTM (Long Short-Term Memory) el cual considera una secuencia de imágenes a fin de hallar variaciones entre las emociones, en este trabajo obtienen una precisión del 75%. Según los objetivos de cada estudio para la medición de las emociones, en [10, 11] proponen un par de métricas llamadas valencia y estimulación, que tratan de indicar el nivel de intensidad de las emociones. Por el contrario, [12] emplean el modelo de emociones propuesto por Ekman que compone de: enojo, disgusto, susto, felicidad, tristeza y sorpresa; en [13] agregan neutralidad. En cuanto al análisis de emociones, en [14] determinan un puntaje de aprendizaje basado en una operación algebraica validada por un grupo de docentes. A diferencia, en [15] genera un modelo de regresión lineal que determina la influencia en el aprendizaje en base a su motivación, los recursos de enseñanza, el trabajo en equipo, y la interacción entre profesor y estudiante, entre otros. Para [12, 16] solo se limitan a describir comportamientos basados en las observaciones de gráficas en series de tiempo.

6.3 Materiales y métodos

Este estudio se enmarca en la resolución No. 8.4.2-90.2/835 de 2023 del Consejo de la Facultad de Ingeniería Electrónica y Telecomunicaciones de la Universidad del Cauca. La guía se desarrolló enfocándose en el dispositivo de bajo costo Jetson Nano, una mini-computadora con GPU integrada y 4GB de RAM compartida, cuya capacidad de cómputo es limitada para modelos de IA de última generación. A nivel de hardware, se integraron una cámara de 8 megapíxeles, una pantalla táctil de 7 pulgadas, un mini teclado, un enrutador de internet y una lampara. A nivel de software, se diseñaron módulos siguiendo la arquitectura mostrada en la Fig. 14.1, permitiendo el despliegue de modelos de DL para capturar, visualizar y almacenar emociones en tiempo real, además de crear un tablero interactivo que informa sobre la experiencia emocional de cada estudiante. Para evitar fuga de datos, se creó una red local mediante el enrutador y la descarga de datos al equipo de análisis se realizó mediante el protocolo de comunicación SSH. Por último, en la Fig. 8.2 se detalla el proceso y flujo de información desde la captura de datos del estudiante y REA, hasta la lectura y aplicación de modelos DL.

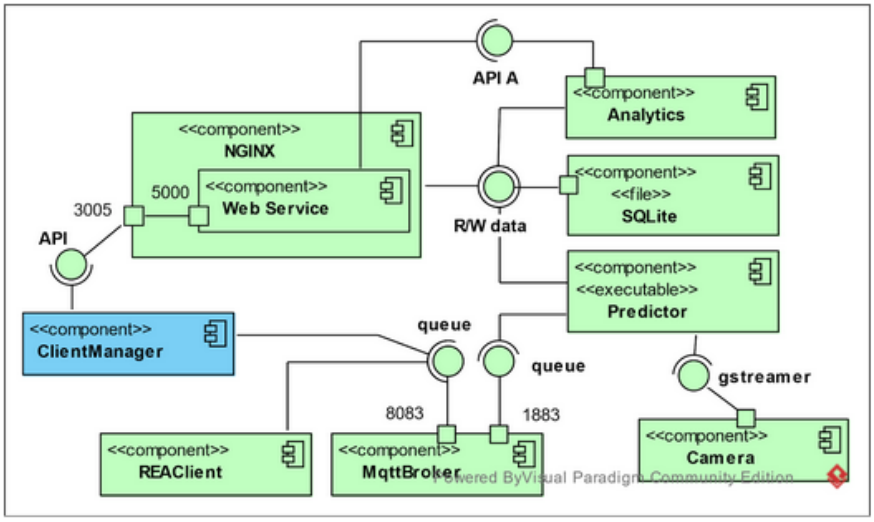


Fig. 6.1: Arquitectura

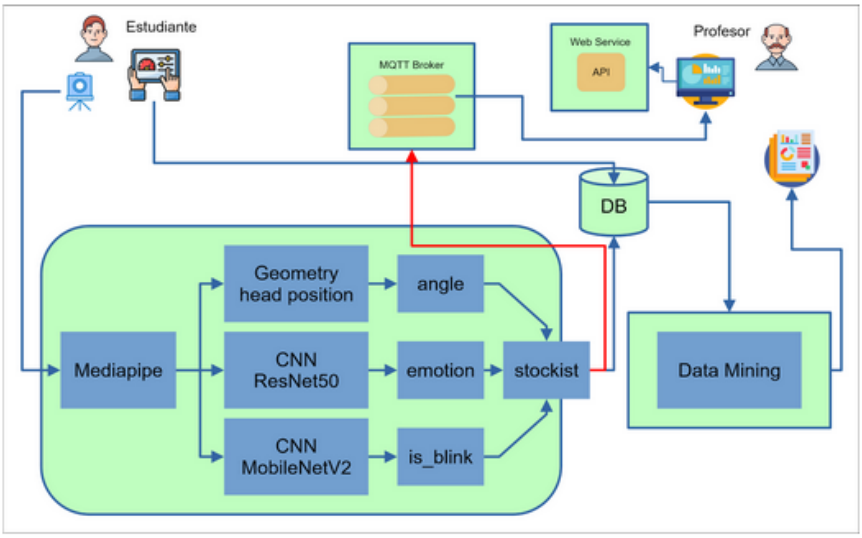


Fig. 6.2: Componentes

6.3.1 Conjunto de datos

FER 2013. Es un conjunto de imágenes etiquetadas en 7 clases según [5]: enojo, disgusto, susto, neutral, felicidad, tristeza y sorpresa. Las imágenes fueron extraídas de internet con un tamaño de 48x48 en escala de grises. Open and close eyes. Es un conjunto de imágenes etiquetadas en 2 clases: abierto y cerrado. Estas imágenes fueron recolectadas de un experimento para detectar síntomas de sueño en personas que conducen vehículos en la noche.

6.3.2 Preparación e implementación de modelos

Primero se adecuó un ambiente en una máquina virtual de Kaggle que ofrece semanalmente 30 horas de cómputo con una capacidad de 32GB de RAM y una GPU TP-100. Para la implementación de los modelos se empleó el framework TensorFlow2, que contiene una amplia gama de arquitecturas de redes neuronales pre-entrenadas. Se crearon y entrenaron dos modelos CNN con transferencia de aprendizaje (ver Fig. 14.3), uno para la clasificación de emociones (ResNet50) y otro para la detección de parpadeos (MobileNetV2). Cada modelo tuvo un rendimiento del 63% y 80%, respectivamente. En ambos casos se podaron las capas de los modelos debido a dos razones:

1. Las imágenes de los conjuntos de datos tienen una dimensión relativamente pequeña y al pasar por todos los pasos de convolución originales pierden información hasta incluso obtener matrices de 1×1.
2. El tamaño del modelo según pruebas de ensayo y error, no deben sobrepasar los 20MG para obtener una frecuencia aceptable de fotogramas por segundo (fps).

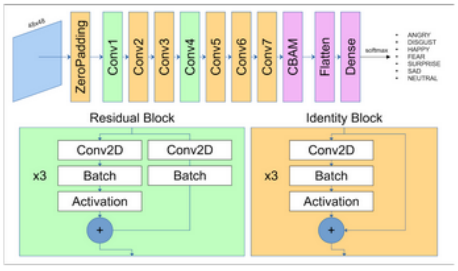


Fig. 6.3: ResNet50

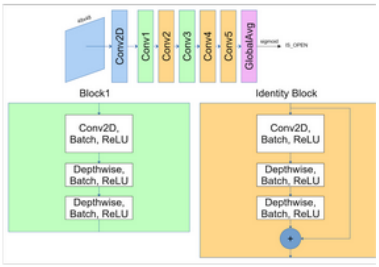


Fig. 6.4: MobileNetV2

Fig. 6.5: Arquitecturas CNN

6.3.3 Adecuación del entorno

En la Institución Educativa Técnica en Sistemas Santiago de la zona rural del suroccidente de Colombia se adecuó un aula de clases para evaluar una serie de REA por alrededor de 105 estudiantes del grado 1 hasta 11, elegidos aleatoriamente, quienes tenían rasgos mestizos e indígenas. A pesar

de la anonimización de los datos[1], por razones éticas contempladas en el Marco Ético para la IA en Colombia[2], se decidió obtener el consentimiento informado para proteger la integridad del menor. Dichos documentos están en poder de los investigadores y el formato se puede ver en el repositorio de Google Drive[3]. Los REA fueron extraídos de la página Colombia Aprende[4] y revisados junto con docentes para ver la viabilidad de su aplicación. Una vez se eligieron 3 REA de las áreas de lenguaje, matemáticas y ciencias por cada grado académico para un total de 33 elementos, se les agrego una capa de código en JavaScript para capturar las interacciones de los estudiantes con el material.

6.3.4 Recolección de datos

Los estudiantes fueron monitoreados mediante una aplicación web (ver Fig. 14.4). Cada sesión de aproximadamente 15 minutos almacenaba datos cada segundo. Tras explicarles la dinámica del recurso, se les permitía interactuar libremente con el material (ver Fig. 14.5). Durante la sesión, el investigador y un docente estaban disponibles para resolver dudas rápidamente. Los datos recolectados se describen brevemente en la Tabla 6.1.

Table 6.1: Descripción de las variables observables.

Nombre	**Tipo**	**Descripción**
sesion_id	Entero	Identificación única de la sesión
ang	Decimal	Índice de enojo
dis	Decimal	Índice de disgusto
fea	Decimal	Índice de susto
hap	Decimal	Índice de felicidad
neu	Decimal	Índice de neutralidad
sad	Decimal	Índice de tristeza
sur	Decimal	Índice de sorpresa
rate_open_eye	Decimal	Tasa de apertura
head_position	Entero	Posición de la cabeza
gender	Entero	Genero biológico del participante
old_year	Entero	Edad del participante
school_year	Entero	Grado académico del participante
rea_level	Entero	Nivel del REA
rea_pos	Carácter	Posición en el REA I: introducción, A: actividad, R: resumen, T: tarea
time_relative	Entero	Tiempo contado desde el inicio

Para identificar fallos antes de la etapa de recolección, se planificaron pruebas con 5 estudiantes. Se describió que el Jetson Nano no soporta una tasa de 9 fps debido a errores de volcado de memoria, por lo que se redujo a 5 fps. También se observó que el modelo detector de parpadeos generaba más falsos positivos en personas con ojos rasgados. Para corregir esto, se tomó una pequeña muestra de imágenes de los ojos de la población objetivo y se realizó un ajuste fino, mejorando así la precisión del modelo.

[1] CEPAL – Datos anonimizados

[2] Marco Ético para la IA en Colombia

[3] Formato consentimiento informado

[4] Portal web – Colombia Aprende – Recursos para aprender

6.4 Resultados

6.4.1 Minería de datos

A fin de encontrar preliminarmente patrones o tendencias en el conjunto de datos, este se importó a un software de minería llamado RapidMiner. En él se crearon y modificaron algunas variables observables como se presenta en la Tabla 8.2.

Table 6.2: Descripción de las variables latentes.

Nombre	**Tipo**	**Descripción**
is_positive	Entero	Indica si la emoción es positiva o negativa
score_emotion	Decimal	Puntaje para el nivel de aprendizaje
change	Decimal	Tasa de cambio emocional
prob	Decimal	Probabilidad de cambio emocional

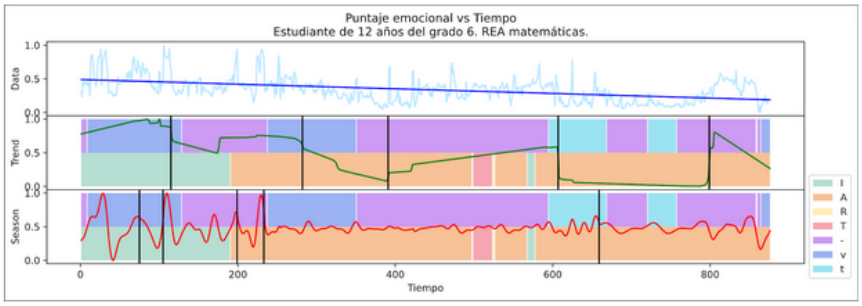


Fig. 6.6: Puntos de cambio en la descomposición de la serie de tiempo en gráficas de tendencia y estacionalidad para un estudiante. Las etiquetas corresponden a la posición del REA

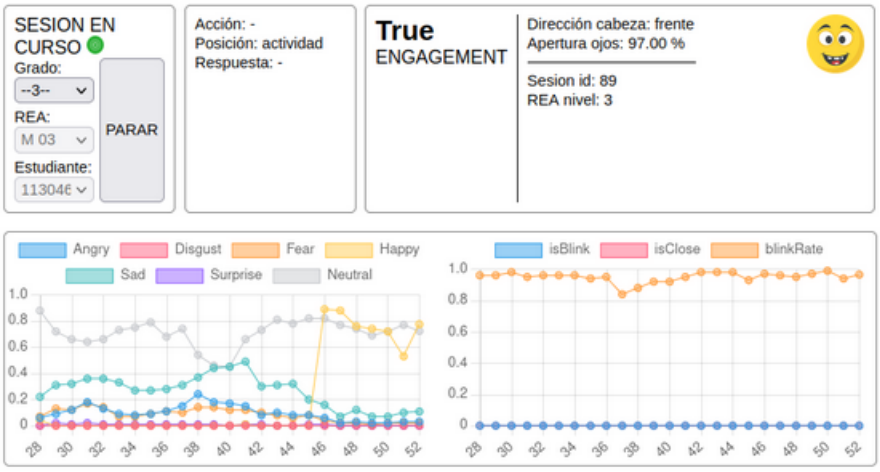


Fig. 6.7: Aplicación para monitorear la recolección de emociones.

La variable "score_emotion" se calculó a partir de la ecuación 6.1 presentada en [14].

$$se = neu + 2 * hap + 3 * sur - (fea + dis + 2 * sad + 2 * ang) \tag{6.1}$$

Fig. 6.8: Montaje. Lado izquierdo vista del Jetson Nano y componentes. Lado derecho un estudiante durante una evaluación con un REA.

Para las variables "change" y "prob" se determinaron a partir del cálculo de cambio de puntos bayesianos empleando la librería de Rbeast en Python[5]. En la Fig. 14.5 se representa de manera gráfica los momentos. De la Fig. 6.6 se observa que existe una caída de emociones positivas en la curva azul claro a medida que transcurre la sesión de evaluación. Esta tendencia se ve reflejada

[5] PIP Rbeast https://pypi.org/project/Rbeast/

en una de las gráficas de [16]. Contrario a la Fig. 6.9 se evidencia un aumento de las emociones positivas entre mayor sea el grado académico, pasando de un 30% hasta un 80%.

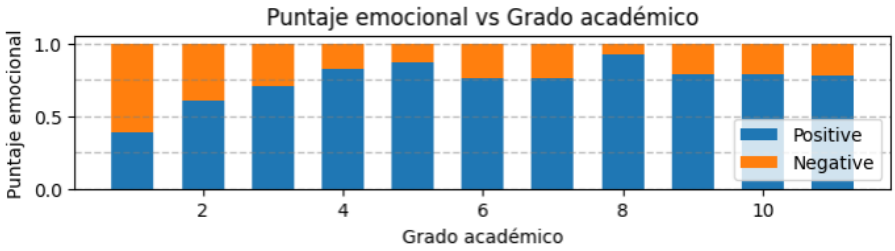


Fig. 6.9: Puntos de cambio en la descomposición de la serie de tiempo en gráficas de tendencia y estacionalidad para un estudiante. Las etiquetas corresponden a la posición del REA.

6.5 Conclusiones y trabajo futuros

Se concluye que la guía proporciona una serie de pasos junto con herramientas de diseño e implementación, para capturar de manera eficiente y precisa datos emocionales en un ambiente real de aula de clases. Se tomaron las medidas necesarias para proteger la integridad del menor, siguiendo el Marco Ético para la IA en Colombia. El proceso de minería de datos en sus fases tempranas de compresión de la información revela relaciones y patrones que pueden ser explotados para la generación de modelos de agrupamiento y predicción de series de tiempo. Para enriquecer las variables observables en el futuro, sería interesante formar un perfil psicológico más detallado de cada participante. Un ejemplo es la técnica PIAR[6], que se enfoca en identificar algún tipo de déficit cognitivo en estudiantes.

Agradecimientos. Este estudio fue financiado por el grupo de investigación GICO de la Universidad del Cauca. Los algoritmos, modelos de IA y conjuntos de datos están disponibles en el repositorio público de GitHub[7].

Referencias

1. D. E. Atkins. “A Review of the Open Educational Resources (OER) Movement: Achievements, Challenges, and”. En: Review Literature And Arts Of The Americas (2007).Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016)
2. A. Tlili y col. “Are open educational resources (OER) and practices (OEP) effective in improving learning achievement? A meta-analysis and research synthesis”. En: International

[6] Secretaría de Educación de Bogotá – PIAR

[7] GitHub – mdquilindo/rea_emotion

Journal of Educational Technology in Higher Education 20.1 (2023). Publisher: Springer International Publishing. issn: 23659440. doi: 10.1186/s41239-023-00424-3.
3. J. M. Spector y col. “Handbook of research on educational communications and technology: Fourth edition”. En: Handbook of Research on Educational Communications and Technology: Fourth Edition (2014). ISBN: 9781461431855, págs. 1-1005. doi: 10.1007/978-1-4614-3185-5.
4. M. U. Abdullah y A. Alkan. “A Comparative Approach for Facial Expression Recognition in Higher Education Using Hybrid-Deep Learning from Students’ Facial Images”. En: Traitement du Signal 39.6 (2022). Type: Article, págs. 1929-1941. doi: 10.18280/ts.390605.
5. O. Sumer y col. “Multimodal Engagement Analysis From Facial Videos in the Classroom”. En: IEEE TRANSACTIONS ON AFFECTIVE COMPU- TING 14.2 (jun. de 2023), págs. 1012-1027. issn: 1949-3045. doi: 10.1109/ TAFFC.2021.3127692.
6. W. E. Villegas-Ch, J. Garcia-Ortiz y S. Sanchez-Viteri. “Identification of Emotions From Facial Gestures in a Teaching Environment With the Use of Machine Learning Techniques”. En: IEEE ACCESS 11 (2023), págs. 38010-38022. issn: 2169-3536. doi: 10.1109/ACCESS.2023.3267007.
7. S. Gupta, P. Kumar y R. Tekchandani. “An optimized deep convolutional neural network for adaptive learning using feature fusion in multimodal data”. En: Decision Analytics Journal 8 (2023). Type: Article. doi: 10 . 1016/j.dajour.2023.100277.
8. Y. Tian, T. Kanade y J. F. Cohn. Handbook of Face Recognition. January 2017. Publication Title: Handbook of Face Recognition. 2011. isbn: 978-0- 85729-932-1. doi: 10.1007/978-0-85729-932-1.
9. S.-Y. Lin y col. “Continuous Facial Emotion Recognition Method Based on Deep Learning of Academic Emotions”. En: SENSORS AND MATERIALS 32.10, SI (2020), págs. 3243-3259. issn: 0914-4935. doi: 10.18494/SAM. 2020.2863.
10. A. Abedi. “Affect-driven ordinal engagement measurement from video”. En: Multimedia Tools and Applications (2023). Type: Article. doi: 10.1007/s11042-023-16345-2.
11. J. Fernández Herrero, F. Gómez Donoso y R. Roig Vila. “The first steps for adapting an artificial intelligence emotion expression recognition software for emotional management in the educational context”. En: British Journal of Educational Technology 54.6 (2023). Type: Article, págs. 1939-1963. doi: 10.1111/bjet.13326.
12. W. Zhuang y col. An Integrated Model for On-Site Teaching Quality Evaluation Based on Deep Learning. En: Wireless Communications and Mobile Computing 2022 (2022). Type: Article. doi: 10.1155/2022/9027907.
13. A. Behera y col. “Associating Facial Expressions and Upper-Body Gestures with Learning Tasks for Enhancing Intelligent Tutoring Systems”. En: International Journal of Artificial Intelligence in Education 30.2 (2020). Type: Article, págs. 236-270. doi: 10.1007/s40593-020-00195-2.
14. J. Pei y P. Shan. “A Micro-expression Recognition Algorithm for Students in Classroom Learning Based on Convolutional Neural Network”. En: Traitement du Signal 36.6 (2019). Type: Article, págs. 557-563. doi: 10.18280/ ts.360611.
15. D. Zhang. “Affective Cognition of Students’ Autonomous Learning in College English Teaching Based on Deep Learning”. En: Frontiers in Psychology 12 (2022). Type: Article. doi: 10.3389/fpsyg.2021.808434.
16. Z. Trabelsi y col. “Real-Time Attention Monitoring System for Classroom: A Deep Learning Approach for Student’s Behavior Recognition”. En: Big Data and Cognitive Computing 7.1 (2023). Type: Article. doi: 10.3390/ bdcc7010048.

7
Abdominal decompression protocol defined by therapeutic elements of ABDOPRE: 1st version

D. Méndez, C. Arámbulo][1][0009-0003-1091-6282]
C. Arámbulo[2][0009-0005-9528-1611]
F. Pracca[3][0000-0002-0956-4561]
F.Simini[4][0000-0003-0241-8737]

Abstract The intra-abdominal hypertension and the abdominal compartment syndrome are severe complications which affect patients in intensive care units (ICUs), increasing length of stay and mortality. The therapeutic use protocol of the ABDOPRE device, is designed to reduce intra-abdominal pressure noninvasively through the application of external negative pressure. The pathophysiology of intra-abdominal hypertension and abdominal compartment syndrome is discussed, along with the limitations of current treatments, and the evolution of the ABDOPRE device is presented. Preliminary clinical results in critical patients suggest that ABDOPRE is effective in reducing intra-abdominal pressure and is well tolerated. The flow chart resulted in this work illustrates the translation of the medical concept to a detailed specification, ready to be included in the software programming of the ABDOPRE microcontroller. ABDOPRE is thus a promising biomedical device for reducing intra-abdominal pressure, aiming to prevent abdominal compartment syndrome without resorting to surgical decompression. Implementing this in daily practice is a goal that aspires to be achieved after introducing improvements suggested by future clinical studies and technological developments.

Keywords Intra-abdominal Pressure, Intra-abdominal Hypertension, Abdominal Compartment Syndrome, Treatment, ABDOPRE, Biomedical Equipment.

7.1 Introduction

During the second half of the 19th century, intra-abdominal pressure (IAP) began to be measured to correlate its increase with physiological alterations affecting organ function in critically ill patients. In 1984, the concept of abdominal compartment syndrome (ACS) was introduced, and indications for abdominal decompression were established [3],[4]. The 2004 International Consensus Conference in Noosa, Australia, unified the diagnostic and treatment criteria for intra-abdominal hypertension (IAH) and ACS, thus facilitating the comparison of results between studies [3],[4]. IAH is defined as a sustained or repeated increase in IAP above 12 mmHg, while ACS is characterized by an IAP exceeding 20 mmHg associated by multiple organ dysfunction or failure [3],[4].

[1,2,4]Núcleo de Ingeniería Biomédica de las Facultades de Medicina e Ingeniería, Universidad de la República, Montevideo, Uruguay.
[2]Departamento de Medicina Intensiva, Hospital de Clínicas,Montevideo, Uruguay.
e-mail: dmendez@fing.edu.uy

Both conditions are common in critically ill patients, particularly those with sepsis, abdominal trauma, acute pancreatitis, major burns, complicated post-surgery, among others [4]. The elevation of IAP can have devastating effects on multiple organ systems, including the cardiovascular, respiratory, renal, and gastrointestinal systems, significantly increasing mortality [4]. Currently, the management of IAH and ACS includes medical and surgical interventions depending on the degree of intra-abdominal pressure the patient presents. Medical strategies focus on the evacuation of intraluminal contents and the administration of diuretics, while surgical interventions, such as laparotomic decompression, are reserved for severe cases [3]. However, these maneuvers can be invasive and risky [3],[4]. The ABDOPRE device was developed as a non-invasive solution for reducing IAP [1],[2],[5],[6]. This research aims to specify the behavior of ABDOPRE Vaccum Bell (Fig. 7.1 during intra-abdominal pressure (IAP) reduction therapy [1],[2],[5],[6]. The therapeutic elements of ABDOPRE are defined as i) the reduction of IAP by a given amount of pressure within a specified time, and ii) the maintenance of IAP for a specified duration. The specification is written in formal language using these therapeutic elements of ABDOPRE, represented through a flowchart. Preliminary clinical results from its use in ICU patients at the Dr. Manuel Quintela Hospital de Clínicas in Uruguay will be described.

Fig. 7.1: ABDOPRE Vacuum bell

7.2 Materials and Methods

7.2.1 Inclusion of ABDOPRE in the Intra-abdominal Hypertension Treatment Protocol

A literature search was conducted on the published background regarding ABDOPRE [1],[2], as well as on the guidelines and complementary information for the treatment of intra-abdominal hypertension and abdominal compartment syndrome [3],[4]. The literature search was performed on portals such as Timbó Foco, PubMed, and IEEE Xplore Digital Library, among others. Once the relevant literature provided by the bibliographic sources was reviewed, steps were established for the therapeutic protocol of IAH. The indication for the use of ABDOPRE consists of conditions of IAH with intra-abdominal pressures between 12 mmHg and 20 mmHg (IAH grade 1 and 2). The application protocol includes the following steps:

1. Place ABDOPRE on the patient's abdomen for one hour at a negative pressure of 35 mmHg [1], while monitoring vital signs.
2. After one hour, measure the IAP, intracranial pressure, respiratory, hemodynamic, and renal parameters.
3. Evaluate if the IAP decreased to values below 12 mmHg: i) If IAP ¡12 mmHg, remove the device. ii) If IAP ¿=12 mmHg, activate ABDOPRE for one more hour with a negative pressure of -35 mmHg.
4. At the end of the next hour, for patients without an initial therapeutic response, thoroughly evaluate the patient for other causes of IAP decompensation.
5. Evaluate if the IAP decreased to values below 12 mmHg: i) If IAP ¡12 mmHg, remove the device. ii) If IAP ¿=12 mmHg, the treating physician will decide whether to activate ABDOPRE again for one more hour with a negative pressure of -35 mmHg.

Subsequently, a flowchart specifying this protocol is developed. This diagram facilitates the understanding and monitoring of the activities to be followed, allowing the identification of key points and possible areas for improvement. Using standardized symbols, a coherent and precise representation of the therapy is ensured, ensuring that all involved parties have a shared vision of the procedure.

7.2.2 Results

The result of our work is the flow chart of Fig. 7.2 as a detailed specification of a new treatment, as per its verbal specifications of the interdisciplinary team. ABDOPRE therefore represents an innovation in the field of intensive care medicine and biomedical engineering. The flow chart of Fig. 7.2 illustrates the translation of the medical concept to a detailed specification, ready to be included in the software programming of the ABDOPRE microcontroller. This software embodies a new clinical treatment, which was suggested empirically for the first time when ABDOPRE was developed [1]. It should be considered that IAP was previously always treated with a surgical procedure, hopefully turned unnecessary in some cases by ABDOPRE. By operating a pump and a bell affixed on the abdomen, for the first time the IAP is reduced to levels below 12 mmHg as a consequence of applying a pressure reduction protocol by ABDOPRE. The detailed specification in the diagram ensures a successful implementation based on the understanding of the underlying processes.

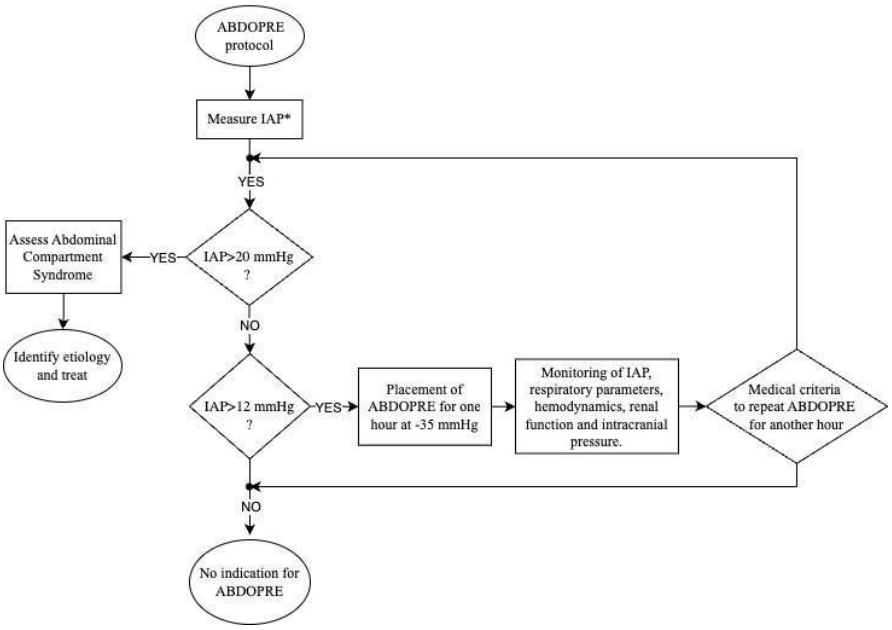


Fig. 7.2: Flowchart of the ABDOPRE therapeutic protocol. *IAP: Intraabdominal pressure.

7.3 Discussion

First use of ABDOPRE was conducted on four IAH Grade 1 and Grade 2 patients in the ICU of the Hospital de Clínicas, Uruguay. The selected patients had IAP between 12 and 15 mmHg. The ABDOPRE Fig 2 protocol was applied [1]. In three out of four patients, a reduction in IAP was observed: 5 mmHg, 2 mmHg, and 3 mmHg, respectively from 14, 12 and 12 mmHg, well into the safe pressure zone. In the fourth patient, who had severe obesity, IAP increased due to a misalignment in the vacuum bell geometry, highlighting the need to adapt the device design for different body morphologies [1], geometric modification performed shortly after [7]. These preliminary results indicate that ABDOPRE is effective in reducing IAP, with good patient tolerance and no adverse effects. However, several areas for improvement are identified:

- Device Adaptability: The efficacy of the device may be compromised in case of severe obesity. This suggests the need to develop vacuum bells of different shapes and sizes to better adapt to the patient's anatomy [5].
- Durability and Maintenance: The continuous use of the device in a clinical setting requires it to be robust and easy to maintain. The use of better materials and a modular design could improve the device's durability [5].
- Software Optimization: Although the software allows for precise IAP control, additional improvements can be made to the user interface to ease its use by medical personnel [6].
- Long-Term Studies: Preliminary results are promising, but long-term studies with a large number of patients are required to confirm the efficacy and safety of ABDOPRE. These studies should include diverse clinical scenarios to fully evaluate the device's potential, in addition to the first protocol presented here in Fig. 7.2.

7.4 Conclusion

The ABDOPRE biomedical equipment represents an innovative and less invasive alternative for managing IAH in critically ill patients. Preliminary results suggest that it is effective and well tolerated, although improvements in its design and therapeutic algorithm functionality are needed to optimize its use in a variety of clinical settings. Future larger and long-term clinical studies will be essential to validate these findings and determine the impact of ABDOPRE on clinical practice. Knowledge and monitoring of IAH are crucial for early diagnosis and appropriate treatment, as untreated ACS can result in multiorgan dysfunction and high mortality. It is essential to monitor parameters such as lung mechanics, hemodynamic parameters, diuresis, IAP, and pH in critically ill patients susceptible to developing ACS.

7.5 Future work

Given that ABDOPRE is still an experimental treatment, a thorough evaluation of potential short-, medium-, and long-term side effects is recommended. Among these, consideration should be given to dermatological effects, as well as any other repercussions that could affect the patient's overall health. This involves establishing specific protocols that consider individual variations in treatment response, ensuring that effectiveness is maximized while potential risks are minimized. Personalizing these durations of use is vital to tailor the treatment to each patient's particular needs, thus ensuring more effective and safer management of abdominal hypertension. Considerable effort is being made to determine the protocols to be applied in other clinical conditions. This work is particularly relevant in the context of the ABDOPRE third bell, designed for use in obese patients. This approach also aims to optimize therapeutic outcomes and improve the quality of life for individuals affected by obesity. Future work on ABDOPRE will consist of refinements of Fig 7.2 -and the protocol it specifies- which will be applied to patients, laying the groundwork for future technological improvements and adaptations. The first ABDOPRE protocol implemented as in Fig 7.2 was limited to adult patients (age ¿ 18) with IAP between 12 and 20 mmHg (IAH stage 1 & 2) and under critical care [1].

References

1. Schandy A, Pracca F, Simini F. Dispositivo de aplicación externa para reducción de presión intraabdominal: experiencia clínica preliminar. Revista Uruguaya de Medicina Intensiva. 2020;35(2):123-129.
2. González Díaz CA, Chapa González C, Laciar Leber E, Vélez HA, Puente NP, Flores DL, Andrade AO, Galván HA, Martínez F, García R, Trujillo CJ, Mejía AR, eds. Treatment of Abdominal Hypertension: Development of an Original Non-invasive Device ABDOPRE. VIII Latin American Conference on Biomedical Engineering and XLII National Conference on Biomedical Engineering Proceedings of CLAIB-CNIB 2019, October 2-5, 2019, Cancún, México. IFMBE Proceedings. 2020; 75:567-574.
3. The Pediatric Guidelines Sub-Committee for the World Society of the Abdominal Compartment Syndrome, Kirkpatrick AW, Roberts DJ, De Waele J, Jaeschke R, Malbrain MLNG, et al. Intra-abdominal hypertension and the abdominal compartment syndrome: updated consensus definitions and clinical practice guidelines from the World Society of the Abdominal Compartment Syndrome. Intensive Care Med. July 2013;39(7):1190-206.

4. Vidal-Cortés P, Díaz-Carrasco MS, Aguilar Alonso E. La hipertensión intraabdominal y el síndrome compartimental abdominal: ¿qué debe saber y cómo debe tratarlos el cirujano? Cirugía Española 2008;83(5):247-257.
5. Maassardjian G, González M, Sánchez P, Pracca F, Simini F. Airtight Hypoallergenic Gasket Design for ABDOPRE Vacuum Bell. XXIV Congreso Argentino de Bioingeniería y XIII Jornadas de Ingeniería Clínica SABI 2023, Buenos Aires, Argentina. 2023; http://www.nib.fmed.edu.uy/sitio_nib/BibliotecaNIB/PublNIB388.pdf
6. González M, Sánchez P, Pracca F, Simini F. Optimización de un circuito de adquisición de sensores de presión del sistema de reducción de la hipertensión abdominal basado en campana de vacío ABDOPRE. XXIV Congreso Argentino de Bioingeniería y XIII Jornadas de Ingeniería Clínica SABI 2023, Buenos Aires, Argentina. 2023; http://www.nib.fmed.edu.uy/sitio_nib/BibliotecaNIB/PublNIB397.pdf
7. Urruty. L, Simini. F, Pereyra. M, David. M y Pracca. F, "Non-invasive Treatment of Abdominal Hypertension: a new vacuum chamber," en 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA: IEEE, 2014; doi:10.1088/1742-6596/90/1/012035.

8
Análisis comparativo de la clasificación de Dengue y Chikungunya mediante técnicas de Machine Learning.

Wilson Arrubla-Hoyos[1]
Jorge Gomez Gomez [2]
Khalid Saeed[3]
Emiro De-la-Hoz-Franc[4]

Resumen The study addresses the problem of differential diagnosis between dengue and chikungunya in endemic areas, where both diseases present similar symptoms. It compares classical techniques like decision trees with ensemble methods, using a dataset that collects information from patients between 2015 and 2020 in Brazil. The results show that ensemble algorithms such as Random Forest, Stacking, and Hard Voting outperform classical techniques like kNN and Boosting, achieving 79% in all evaluated metrics. Although decision trees perform slightly lower, with 78% in all metrics, they remain a viable option due to their ease of interpretation, especially in clinical settings. Additionally, it is identified that arthralgia is the most important variable for distinguishing between chikungunya and dengue, supporting the importance of considering specific symptoms in the early classification of these diseases. These results have the potential to significantly impact the early classification of these diseases, supporting medical decision-making, especially in areas with a shortage of doctors and epidemiology specialists.

Palabras clave Chikungunya; Dengue; Differential classification; Ensemble Methods; Machine Learning.

8.1 Introducción

Según la Organización Mundial de la Salud (OMS), la cocirculación de los virus del dengue y chikunguña es un problema de salud pública en toda la región de las Américas [1, 2, 3, 4]. El dengue es endémico y tiene ciclos con picos de brotes cada tres a cinco años [1], siendo esta enfermedad la más grave debido a que genera la mayor cantidad de casos fatales [4, 5, 6].

El diagnóstico diferencial de estas enfermedades es crucial en áreas endémicas, ya que los dos virus presentan manifestaciones clínicas similares [7]. Algunas investigaciones han identificado características clínicas que pueden ayudar a diferenciarlas, como el dolor de cabeza, la hemorragia y la

[1]Facultad de Ingeniería, Universidad Nacional Abierta y a Distancia, Sincelejo 700002
[2]Facultad de ingeniería, Universidad de Córdoba, Montería 230001
[3]Faculty of Computer Science, Bialystok University of Technology, Bialystok, Poland
[4]Departamento de computación y electrónica, Facultad de Ingeniería, Universidad de la Costa, Barranquilla 080002, Colombia
e-mail: Wilson.arrubla@unad.edu.co, jeliecergomez@correo.unicordoba.edu.co,
e-mail: k.saeed@pb.edu.pl, edelahoz@cuc.edu.co

hepatomegalia en el dengue, y la combinación de fiebre y artralgia en las infecciones por chikunguya [8, 9]. Sin embargo, el informe más reciente de la Organización Panamericana de la Salud (OPS) ha establecido una síntesis de evidencia por expertos que emana directrices para el diagnóstico del dengue, Zika y chikunguña en la región de las Américas. Dicha guía presenta los signos y síntomas que diferencian estas tres enfermedades, que tienen un cuadro clínico similar y presentan dificultades para su diagnóstico temprano [10]. Frente a esta situación, algunos investigadores han propuesto soluciones basadas en machine Learning (ML) como apoyo clínico al diagnóstico temprano del dengue y otras enfermedades similares [11, 12, 13, 14]. Por ejemplo, en [15] presentaron modelos ML basados en datos recolectados en Brasil entre 2015 y 2020. Estos autores evaluaron diferentes técnicas clásicas como *k*NN y Bayes, y ensambles como Random Forest (RF) y XGBoost entre otras más; para clasificar el dengue, chikunguña y otra enfermedad. Asimismo, exploraron la búsqueda de los mejores hiperparámetros obteniendo valores aceptables en las cuatro métricas de calidad. En la misma línea y con el mismo dataset, en [16] propusieron un estudio basado en redes convolucionales CNN, donde convirtieron los datos en imágenes y obtuvieron buenas métricas, con un 82% de exactitud y precisión. Por otro lado, [17] propone un algoritmo híbrido para la clasificación de ocho enfermedades transmitidas por mosquitos, incluyendo chikunguña y dengue. En el estudio se comparan siete técnicas de clasificación, destacándose redes neuronales artificiales (ANN), máquinas de soporte vectorial (SVM), el algoritmo J48 y técnicas computacionales más complejas como random forest y AdaBoost. Aunque la propuesta híbrida (HML) de clasificación, que combina aprendizaje basado en refuerzo en redes neuronales recurrentes (RNN) con clasificadores convencionales o más complejos, alcanzó un desempeño superior con una precisión del 98.76%, los algoritmos de árboles J48 y random forest presentaron desempeños similares.

Este artículo analiza y compara modelos de aprendizaje automático (ML) convencionales y de ensambles para la predicción de dengue y chikunguña, abordando la diferenciación entre ambas enfermedades. Dado que ambas presentan problemáticas en el diagnóstico temprano debido a la similitud de sus síntomas, se seleccionaron técnicas basadas en la literatura previa. Algunas técnicas convencionales han demostrado un buen rendimiento en la clasificación, mientras que los modelos de ensambles tienden a ofrecer un mejor desempeño, aunque con un mayor costo computacional durante el entrenamiento. Además, se destacan técnicas clásicas de fácil interpretación, como los árboles de decisión, que son especialmente útiles en el ámbito médico, ya que los profesionales de la salud están familiarizados con estos métodos. El resto del documento está organizado de la siguiente manera: la sección 2 discute los materiales y métodos utilizados en la experimentación, la sección 3 presenta los resultados y la discusión, y la sección 4 concluye el artículo.

8.2 Materiales y métodos

El experimento se desarrolló utilizando los datos públicos disponibles en *Clinical cases of Dengue and Chikungunya*[1], que contiene 17,172 registros y 27 atributos de pacientes con Dengue y Chikunguña, así como casos descartados, recopilados entre 2015 y 2020 en Amazonas y Recife, Brasil. Se empleó la plataforma Jupyter Notebook en Google Colab, utilizando el lenguaje Python y las librerías pandas, numpy y scikit-learn. Para desarrollar los modelos se siguieron los siguientes pasos:

[1] https://data.mendeley.com/datasets/bv26kznkjs/1

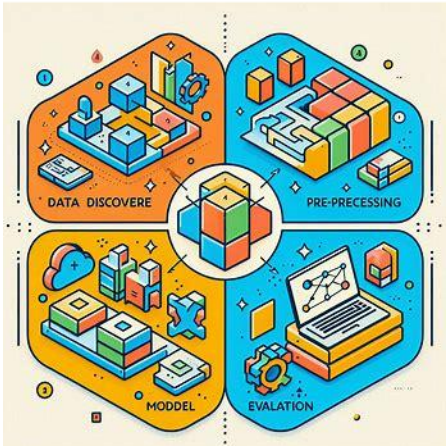


Fig. 8.1: Pasos del proceso de experimentación.

8.2.1 Entendimiento de los datos

El conjunto de datos depurado contiene un total de 11, 448 registros, divididos equitativamente entre las categorías de DENGUE y CHIKUNGUNYA, con 5, 724 registros cada una. Se incluyeron 27 variables en el análisis. La categoría OTROS fue excluida para enfocarse exclusivamente en estas dos enfermedades, dada la similitud de sus síntomas, lo que dificulta su clasificación. La Tabla 8.1 proporciona una descripción detallada de las variables incluidas en el dataset.

Table 8.1: Descripción del dataset

#	variable	Descripción	#	variable	Descripción
1	EDAD	Edad del paciente	14	ARTRITE	Síntoma artritis- SI, NO
2	SEXO	Sexo del paciente, hombre, mujer	15	ARTRALGIA	Síntoma artralgia- SI, NO
3	GESTANT	Edad gestacional del paciente(cuartil) en caso de sexo femenino-1er trimestre, 2do trimestre, 3er cuarto, Se ignora la edad gestacional, No, No aplicable, Ignorado	16	PETEQUIA_N	Síntoma petequias- SI, NO
4	CS_RACA	Raza del paciente-blanco, negro, amarillo, marrón, Indígena, Ignorado	17	LACO	Prueba torniquete positivo- SI, NO
5	CS_ZONA	Área de residencia- urbano, zonas rurales, Periurbano, Ignorado	18	DOR_RETRO	Dolor retro ocular- SI, NO
6	FEBRE	Síntoma fiebre- SI, NO	19	DIABETES	Enfermedad preexistente - diabetes - SI, NO
7	MIALGIA	Síntoma mialgia - SI, NO	20	HEMATOLOG	Enfermedad preexistente - Enfermedades hematológicas hematológica- SI, NO
8	CEFALEIA	Síntoma cefalea- SI, NO	21	HEPATOPAT	Enfermedad preexistente - Enfermedades del hígado- SI, NO
9	EXANTEMA	Síntoma exantema- SI, NO	22	RENAL	Enfermedad preexistente - Enfermedad renal- SI, NO
10	VOMITO	Síntoma vomito- SI, NO	23	HIPERTENSA	Enfermedad preexistente - Hipertensión- SI, NO
11	NAUSEA	Síntoma nausea- SI, NO	24	ACIDO_PEPT	Enfermedad preexistente - Enfermedad del ácido péptico- SI, NO
12	DOR_COSTAS	Síntoma dolor de espalda- SI, NO	25	AUTO_IMUNE	Enfermedad preexistente - Enfermedad autoinmune- SI, NO
13	CONJUNTVIT	Síntoma conjuntivitis- SI, NO	26	DIAS_SYMPTOMS	Número de días que el paciente experimenta síntomas- SI, NO
			27	CLASSI_FIN	Clasificación final de pacientes (DENGUE, CHIKUNGUÑA)

8.2.2 Preprocesamiento de los datos

Se llevaron a cabo diversas etapas para garantizar la coherencia de los datos y prepararlos adecuadamente para el entrenamiento del modelo. En primer lugar, se revisaron los descriptivos, los valores atípicos y los datos faltantes. Se ajustó la variable “días de síntoma” para la enfermedad del Dengue, considerando que el ciclo de la enfermedad suele ser mayor a 9 días, tomando como referencia la media. Para la enfermedad del Chikungunya, se realizó un ajuste para los valores

mayores de 12 días, basándose en información proporcionada por el protocolo de la enfermedad de la OMS. Además, se trató el problema de los datos faltantes en las variables nominales aplicando una técnica de imputación avanzada basada en k-vecinos más cercanos (kNN).

8.2.3 Creación del modelo

Se emplearon técnicas de aprendizaje automático utilizando las bibliotecas Scikit-learn, Matplotlib, Numpy y Pandas en Google Colab, con Python como lenguaje de programación. El entrenamiento se realizó utilizando el 70% de los datos y se evaluó con el 30% restante. Se utilizaron diversas técnicas basadas en una revisión de la literatura previa, incluyendo Árboles de Decisión (DT), Redes Neuronales Artificiales (ANN), Máquinas de Soporte Vectorial (SVM), y k-Vecinos más Cercanos (kNN). Además, se aplicaron métodos de conjunto como Bosques Aleatorios (RF), Boosting, Hard Voting, Soft Voting, Stacking y Bagging.

8.2.4 Evaluación del modelo

Se empleó la matriz de confusión para calcular métricas como exactitud, precisión, recuperación y puntuación F1, lo que permitió comparar el rendimiento de los modelos clásicos con los métodos de conjunto.

8.3 Resultados y discusión

La tabla 8.2 muestra los resultados del proceso experimental, destacando las diversas técnicas empleadas.

Table 8.2: Métricas de calidad de los modelos

ML technique	**accuracy**	**precision**	**recall**	**F1- Score**
Decision Tree	78%	78%	78%	79%
Knn	68%	68%	68%	68%
Neural Network	77%	77%	77%	77%
RF	79%	79%	79%	79%
Baggin	77%	77%	77%	77%
Boosting	71%	71%	71%	71%
Hard-Voting	79%	79%	79%	79%
Soft- voting	77%	78%	77%	77%
Stacking	79%	79%	79%	79%

Los resultados revelan una superioridad de los algoritmos de ensambles sobre las técnicas clásicas empleadas en el experimento. Sin embargo, los árboles de decisión alcanzaron un rendimiento del 78%, muy similar al de técnicas más complejas como Random Forest RF, Stacking y Hard Voting, que lograron un 79%. Estas técnicas requieren una mayor capacidad de procesamiento

para desarrollar los modelos y la mayoría trabaja como caja negra, a excepción de Random Forest. En la Fig. 8.2 se presenta la importancia de las variables en la clasificación del RF, destacando que la artralgia es la variable más importante para esta tarea.

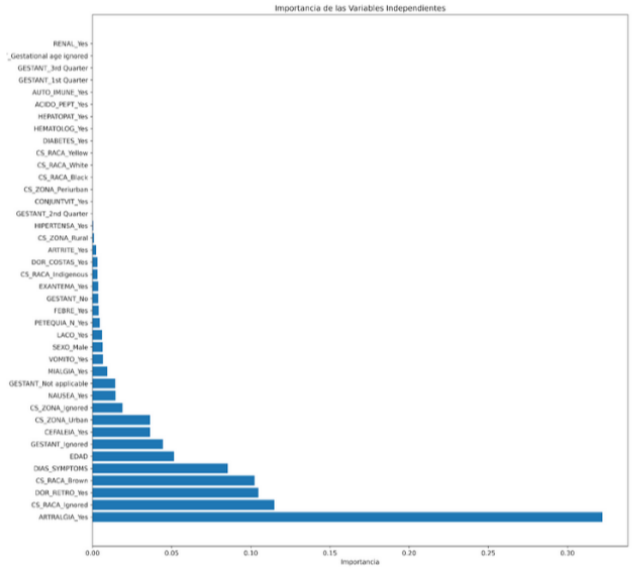


Fig. 8.2: Importancia de las variables del modelo Random Forest

Por otro lado, los resultados revelan la dificultad que presentan las redes neuronales artificiales (ANN) para clasificar las enfermedades en este conjunto de datos. Esta limitación se hizo evidente en este estudio, así como en investigaciones previas, como la realizada por [15] con el mismo conjunto de datos, y en [17], que empleó un conjunto de datos distinto. Estas dificultades podrían deberse a la similitud en los síntomas de las enfermedades, o a la cantidad de los datos como señala [18] lo que complica la diferenciación precisa. Esta observación subraya la necesidad de explorar técnicas adicionales o enfoques híbridos para mejorar la precisión en la clasificación de enfermedades similares.

La Fig. 14.3 muestra la matriz de confusión de RF, que se utiliza para derivar las métricas de calidad utilizadas para evaluar el rendimiento de los modelos. De los 1282 datos evaluados, se clasificaron 1048 como verdaderos positivos y 1282 como verdaderos negativos. Hubo 436 falsos positivos y 669 falsos negativos. Además, este algoritmo tiene una precisión del 76% en la clasificación del Dengue y del 81% en la clasificación de la chikunguña. Por otro lado, es importante destacar que el algoritmo de árbol de decisión es más fácil de interpretar para respaldar la toma de decisiones en área clínica. La comunidad médica está familiarizada con este algoritmo, ya que permite visualizar de manera gráfica el proceso de toma de decisiones [19], lo que facilita la comprensión de cómo los síntomas pueden considerarse en el proceso de clasificación, como se muestra en la Fig. 8.4 .

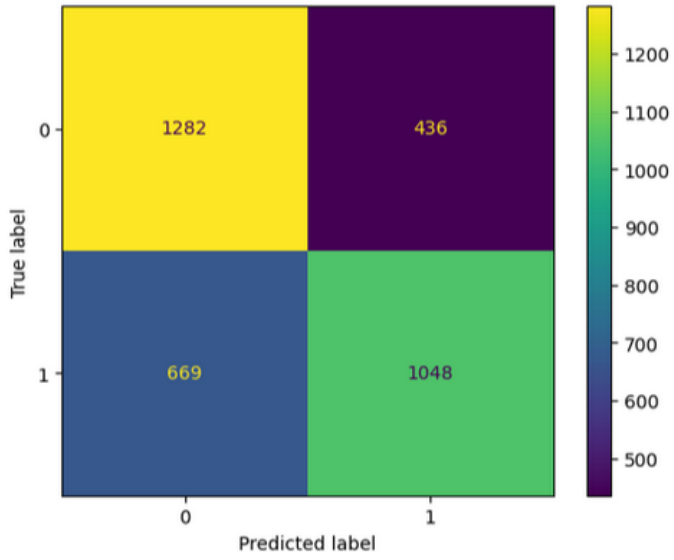


Fig. 8.3: Matriz de confusión del modelo random forest

Además, la Fig. 8.4 muestra que la artralgia es la variable más importante para diferenciar entre el Chikunguña del dengue, lo cual está en línea con los hallazgos clínicos establecidos por la Organización panamericana de la salud OPS [10], con una alta certeza en la evidencia. Además, el dolor retro ocular es una variable importante para distinguir entre ambas enfermedades, con una certeza clínica moderada [10].

En cuanto a las métricas de este algoritmo, según la matriz de confusión, de un total de 1411 datos, se clasificaron 1255 como verdaderos positivos y 1411 como verdaderos negativos. Hubo 462 falsos negativos y 307 falsos positivos. Además, se obtuvo una precisión del 80% en la clasificación del dengue y un recall del 82% en la clasificación del Chikunguña, lo que indica un buen rendimiento en la clasificación de ambas enfermedades. Además, los resultados obtenidos en las cuatro métricas de calidad, con un 78%, muestran un mejor rendimiento que los obtenidos por [15], quienes trabajaron con el mismo dataset y obtuvieron un rendimiento inferior en las cuatro métricas. Es importante tener en cuenta que su estudio implicó un proceso de multiclasificación en la variable objetivo, mientras que en este trabajo solo se consideraron el Dengue y el Chikunguña, que son enfermedades con un cuadro clínico similar.

Por otra parte, los resultados obtenidos por los árboles de decisión, aunque son más bajos que los obtenidos por [16], quienes, al igual que los autores anteriores, llevaron a cabo un proceso de multiclasificación con redes neuronales convolucionales (CNN) logrando un 82% en todas sus métricas, es mucho más fácil de interpretar. Esto se debe a que los árboles de decisión presentan las variables más importantes para realizar la clasificación, en comparación con las CNN, que funcionan como cajas negras y solo se obtiene el resultado.

Estos resultados respaldan la clasificación temprana de enfermedades como el dengue y el Chikunguña, que presentan síntomas clínicos similares. Esta clasificación temprana podría ser crucial para salvar vidas, especialmente en áreas donde ambas enfermedades coexisten, lo que dificulta su diagnóstico precoz. Esto es especialmente relevante en lugares de difícil acceso donde la escasez de médicos puede agravar la situación.

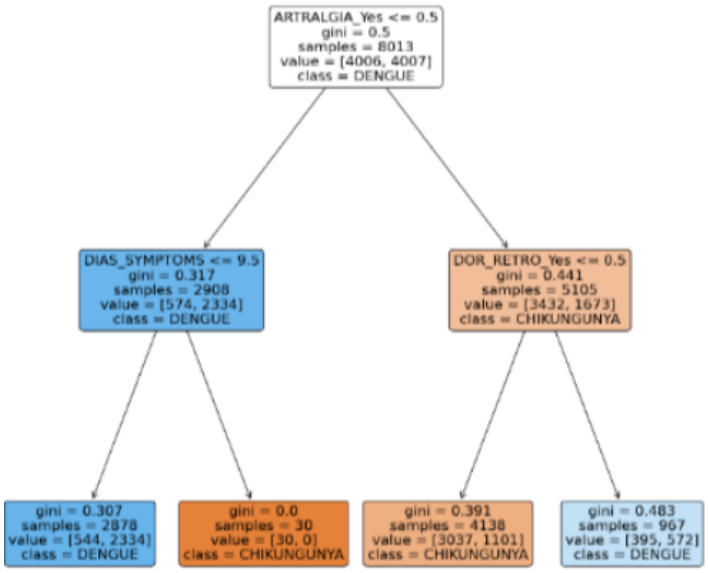


Fig. 8.4: Árbol de decisión para la clasificación de la enfermedad.

8.4 Conclusiones

Los algoritmos de ensambles, como Random Forest, Stacking y Hard Voting, mostraron un rendimiento superior en comparación con las técnicas clásicas como *k*NN y Boosting, alcanzando un 79% en todas las métricas de calidad evaluadas. Aunque los árboles de decisión obtuvieron un rendimiento ligeramente inferior al de los algoritmos de ensamblaje, con un 78% en todas las métricas, siguen siendo una opción viable debido a su facilidad de interpretación, lo que los hace más útiles en entornos clínicos donde se requiere una comprensión detallada de las decisiones del modelo.

Los resultados del estudio destacan que la artralgia es la variable más significativa para distinguir entre Chikunguña y Dengue, coincidiendo con las recomendaciones más recientes emitidas por la Organización Panamericana de la Salud. No obstante, se identificaron otras variables críticas, como el dolor retroocular, la duración de los síntomas y la raza del paciente, que también juegan un papel importante en la diferenciación de estas enfermedades. Estos hallazgos enfatizan la necesidad de una evaluación integral de estos síntomas específicos para mejorar la precisión en el diagnóstico temprano de estas afecciones, las cuales presentan un cuadro clínico notablemente similar. La consideración de múltiples factores puede, por lo tanto, facilitar un diagnóstico más preciso y oportuno, lo que es fundamental para el manejo adecuado y la prevención de complicaciones en los pacientes afectados.

Finalmente, los resultados de este trabajo pueden tener un impacto significativo al apoyar la clasificación temprana de estas enfermedades, lo que a su vez respaldaría la toma de decisiones médicas. Este impacto sería aún mayor en áreas apartadas de las zonas urbanas, donde hay escasez de médicos y especialistas en epidemiología.

Referencias

1. OMS Expansión geográfica de los casos de dengue y chikungunya más allá de las áreas históricas de transmisión en la Región de las Américas Available online: https://www.who.int/es/emergencies/disease-outbreak-news/item/2023-DON448 (accessed on 22 February 2024).
2. Espinal, M.A.; Andrus, J.K.; Jauregui, B.; Waterman, S.H.; Morens, D.M.; Santos, J.I.; Horstick, O.; Francis, L.A.; Olson, D. Emerging and Reemerging Aedes-Transmitted Arbovirus Infections in the Region of the Americas: Implications for Health Policy. Am J Public Health 2019, 109, 387–392, doi:10.2105/AJPH.2018.304849.
3. Vidanapathirana, M. Dengue Haemorrhagic Fever in Chronic Kidney Disease and Heart Failure: Challenges in Fluid Management. Tropical Medicine and Health 2024, 52, doi:10.1186/s41182-024-00600-9.
4. Kuo, C.-Y.; Yang, W.-W.; Su, E.C.-Y. Improving Dengue Fever Predictions in Taiwan Based on Feature Selection and Random Forests. BMC Infect Dis 2024, 24, 1–12, doi:10.1186/s12879-024-09220-4.
5. Schaefer, T.J.; Panda, P.K.; Wolford, R.W. Dengue Fever. 2017.
6. Effler, P.V.; Pang, L.; Kitsutani, P.; Vorndam, V.; Nakata, M.; Ayers, T.; Elm, J.; Tom, T.; Reiter, P.; Rigau-Perez, J.G.; et al. Dengue Fever, Hawaii, 2001–2002. Emerging infectious diseases 2005, 11, 742.
7. Villamil-Gómez, W.E.; Rodríguez-Morales, A.J.; Uribe-García, A.M.; González-Arismendy, E.; Castellanos, J.E.; Calvo, E.P.; Álvarez-Mon, M.; Musso, D. Zika, Dengue, and Chikungunya Co-Infection in a Pregnant Woman from Colombia. International Journal of Infectious Diseases 2016, 51, 135–138, doi:10.1016/j.ijid.2016.07.017.
8. Kharwadkar, S.; Herath, N. Clinical Manifestations of Dengue, Zika and Chikungunya in the Pacific Islands: A Systematic Review and Meta-Analysis. Reviews in Medical Virology 2024, 34, doi:10.1002/rmv.2521.
9. Paniz-Mondolfi, A.E.; Rodriguez-Morales, A.J.; Blohm, G.; Marquez, M.; Villamil-Gomez, W.E. ChikDenMaZika Syndrome: The Challenge of Diagnosing Arboviral Infections in the Midst of Concurrent Epidemics. Ann Clin Microbiol Antimicrob 2016, 15, 42, s12941-016-0157–x, doi:10.1186/s12941-016-0157-x.
10. PAHO Síntesis de evidencia: Directrices para el diagnóstico y el tratamiento del dengue, el chikunguña y el zika en la Región de las Américas. Revista Panamericana de Salud Pública 2022, 46, 1, doi:10.26633/RPSP.2022.82.
11. Abdualgalil, B.; Abraham, S.; Ismael, W.M. Early Diagnosis for Dengue Disease Prediction Using Efficient Machine Learning Techniques Based on Clinical Data. J. Robot. Control. 2022, 3, 257–268, doi:10.18196/jrc.v3i3.14387.
12. Aleixo, R.; Kon, F.; Rocha, R.; Camargo, M.S.; De Camargo, R.Y. Predicting Dengue Outbreaks with Explainable Machine Learning. In Proceedings of the Proceedings - 22nd IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing, CCGrid 2022; Fazio M., V.M., Panda D.K.,. Prodan R.,. Cardellini V.,. Kantarci B.,. Rana O., Ed.; Institute of Electrical and Electronics Engineers Inc., 2022; pp. 940–947.
13. Appice, A.; Gel, Y.R.; Iliev, I.; Lyubchich, V.; Malerba, D. A Multi-Stage Machine Learning Approach to Predict Dengue Incidence: A Case Study in Mexico. IEEE Access 2020, 8, 52713–52725, doi:10.1109/ACCESS.2020.2980634.
14. Biswas, N.; Uddin, K.M.M.; Rikta, S.T.; Dey, S.K. A Comparative Analysis of Machine Learning Classifiers for Stroke Prediction: A Predictive Analytics Approach. Healthcare Analytics 2022, 2, 100116, doi:https://doi.org/10.1016/j.health.2022.100116.
15. Tabosa de Oliveira, T.; da Silva Neto, S.R.; Teixeira, I.V.; Aguiar de Oliveira, S.B.; de Almeida Rodrigues, M.G.; Sampaio, V.S.; Endo, P.T. A Comparative Study of Machine Learning Techniques for Multi-Class Classification of Arboviral Diseases. Frontiers in Tropical Diseases 2022, 2.
16. Medeiros Neto, L.; Rogerio da Silva Neto, S.; Endo, P.T. A Comparative Analysis of Converters of Tabular Data into Image for the Classification of Arboviruses Using Convolutional Neural Networks. PLoS One 2023, 18, e0295598, doi:10.1371/journal.pone.0295598.

17. Shaikh, Salim Gulab, et al. ”Original Research Article Hybrid machine learning method for classification and recommendation of vector-borne disease.” Journal of Autonomous Intelligence 7.2 (2024).
18. Emeršič, Žiga, et al. ”Training convolutional neural networks with limited training data for ear recognition in the wild.” arXiv preprint arXiv:1711.09952 (2017).
19. Martínez, R.E.B.; Ramírez, N.C.; Mesa, H.G.A.; Suárez, I.R.; León, P.P.; Morales, S.L.B. Árboles de decisión como herramienta en el diagnóstico médico. 2009, 6.

9

Implementación de un Sistema de Gestión de Seguridad de la Información conforme a ISO/IEC 27001:2022 en la Universidad de la Amazonia

Kevin Alejandro Benavides Nuñez[1][0009-0007-4972-2548]
Heriberto Fernando Vargas Losada[2][0000-0002-8561-0793]

Resumen Esta investigación se centró en la implementación de un Sistema de Gestión de Seguridad de la Información (SGSI) en la oficina de tecnologías de la información (OTI) de la Universidad de la Amazonia, adhiriéndose a las normas ISO/IEC 27001:2022. El estudio empleó la metodología MAGERIT v.3 para el análisis y evaluación de riesgos, complementada por el ciclo PDCA para garantizar la mejora continua. Mediante métodos exhaustivos de recopilación de datos, como encuestas, entrevistas y análisis de documentos, se realizó un inventario detallado y una clasificación de los activos de información de la universidad. La identificación de amenazas, vulnerabilidades y riesgos potenciales reveló áreas críticas que necesitan una mayor protección. El estudio destaca la importancia de la formación continua, las revisiones periódicas y la inversión tecnológica para gestionar eficazmente los riesgos de seguridad de la información y garantizar la integridad, confidencialidad y disponibilidad de los datos.

Palabras clave Sistema de gestión de la seguridad de la información (SGSI), ISO/IEC 27001:2022, Metodología MAGERIT v.3, Análisis de riesgos, Protección de datos.

9.1 Introducción

En la actualidad, la protección de datos sensibles y críticos es más crucial que nunca, y la tecnología desempeña un papel central en la gestión y el tratamiento de la información. El desarrollo de nuevos modelos y procesos en la gestión empresarial se ha visto impulsado por los avances de la tecnología informática. Sin embargo, estos avances también han dado lugar a nuevas amenazas para los activos internos y las fuentes externas de información. Por lo tanto, es esencial que las organizaciones adopten modelos y patrones que mantengan la información protegida y organizada [1]. En esta era digital, los ciberdelincuentes han incrementado significativamente los ataques a usuarios y empresas. Se estima que se han producido aproximadamente 467.000 ciberataques en el mundo, causando considerables pérdidas económicas y de información. Esta situación pone de manifiesto la urgente necesidad de mantener la confidencialidad, integridad, disponibilidad y privacidad de la información para garantizar su protección efectiva [2].

Para lograr una implementación efectiva de los procesos de seguridad de la información, es necesario dedicar esfuerzos constantes y seguir estrategias sólidas, planes operativos precisos

Universidad de la Amazonia, Florencia, Colombia
e-mail: \{ke.benavides,heri.vargas\}@udla.edu.co

y procedimientos adecuados. Además, deben aplicarse eficazmente las etapas del ciclo PDCA (Planificar, Hacer, Comprobar, Actuar) [3]. En este contexto, la Universidad de la Amazonia reconoce su responsabilidad de garantizar la salvaguarda y privacidad de la información. Para cumplir con este com- promiso, se propone implementar un SGSI basado en los principios de la norma ISO 27001:2022. Este enfoque integral abarca desde la detección de riesgos hasta la aplicación de medidas preventivas y correctoras.

9.2 Metodología

Esta investigación adoptó un enfoque mixto, combinando métodos cualitativos y cuantitativos para obtener una visión integral y detallada del problema. Este enfoque metodológico se fundamenta en la necesidad de capturar tanto la profundidad de las experiencias individuales y contextuales (métodos cualitativos) como la amplitud y generalización de los datos estadísticos (métodos cuantitativos). La combinación de ambos métodos permitió una comprensión profunda y completa de la situación actual y de los requisitos necesarios para implantar un SGSI eficaz.

Para la ejecución de este proyecto, se adoptó la metodología MAGERIT v.3, reconocida por su robustez en el análisis y evaluación de los riesgos inherentes a los activos de información. MAGERIT ofrece un marco sistemático y estructurado para identificar, valorar y gestionar riesgos, lo que es fundamental para asegurar una protección adecuada de los activos informáticos. Además, esta metodología se complementó con la norma NTC-ISO/IEC 27001:2022 [4], que proporciona un enfoque de gestión de la seguridad de la información basado en el ciclo PDCA (Planear, Hacer, Verificar, Actuar). Este ciclo garantiza la mejora continua, asegurando que el SGSI evolucione y se adapte a los cambios en el entorno y las amenazas. La recolección de datos se llevó a cabo mediante una serie de técnicas complementarias:

- Encuestas.
- Entrevistas estructuradas y no estructuradas.
- Observación simple.
- Análisis de documentos.

9.3 Resultados

9.3.1 Ejecución del SGSI:

Se utilizó la metodología PHVA (Planificar, Hacer, Verificar, Actuar) para gestionar el SGSI en la Oficina de Tecnología de la Información (OTI) de la Universidad de la Amazonia, con un enfoque en la mejora continua y adaptación.

Fase Planificar.

- **Análisis de Seguridad:** Se evaluó el nivel de seguridad según ISO/IEC 27001:2022.
- **Gestión de Activos:** Se realizó un inventario detallado y clasificación de los activos de información, como bases de datos, correos electrónicos, hardware, y software.
- **Valoración de Activos:** Se asignaron valores de confidencialidad, integridad y disponibilidad para priorizar la seguridad de cada activo.

Fase Hacer.

- **Identificación de Amenazas y Vulnerabilidades:** Se identificaron y clasificaron las amenazas que podrían afectar los activos, como acciones no autorizadas, fenómenos naturales, problemas técnicos, y actos violentos.

Fase Verificar.

- **Identificación de Riesgos:** A partir de las amenazas y vulnerabilidades, se evaluaron los riesgos en términos de impacto y probabilidad.
- **Matriz de Riesgo:** Los riesgos se categorizaron para priorizar las acciones de mitigación necesarias.

Fase Actuar.

- **Medidas de Protección:** Se propusieron controles específicos para mitigar riesgos, como autenticación multifactorial, planes de contingencia, instalación de software antivirus, capacitación del personal, y estrategias de recuperación ante desastres.

9.4 Discussion

El análisis realizado en la Universidad de la Amazonia mostró que ya se han implemen- tado medidas de seguridad de la información alineadas con la norma ISO/IEC 27001:2022, pero aún hay áreas de mejora. La metodología MAGERIT v.3 y el ciclo PDCA permitieron un análisis riguroso de riesgos, destacando la necesidad de controles adicionales para proteger los activos críticos. Se enfatizó la importancia de mantener actualizado el inventario de activos y de realizar evaluaciones periódicas de amenazas y vulnerabilidades. A pesar de los beneficios esperados de un SGSI basado en ISO/IEC 27001:2022, como mayor protección de datos y cumplimiento normativo, se identificaron desafíos como la resistencia al cambio, la necesidad de inversión en tecnología y formación, y la falta de personal capacitado y recursos.

9.5 Recomendaciones

Para superar estos retos, se recomienda:

- Formación continua: Implantar programas de formación y concienciación para todo el personal sobre la importancia de la seguridad de la información y los nuevos procedimientos.
- Revisión y actualización periódica: Realizar evaluaciones y actualizaciones periódicas del SGSI para garantizar su eficacia y pertinencia ante nuevas amenazas y cambios en la infraestructura.
- Inversión en tecnología: Invertir en herramientas y tecnologías de seguridad avanzadas para reforzar la protección de los activos de información.
- Compromiso de la alta dirección: Garantizar el apoyo y compromiso continuos de la alta dirección para la implantación y mantenimiento del SGSI.

- Implementar un sistema de IA: Detectar comportamientos anómalos en el tráfico de red y automatice la respuesta a incidentes de seguridad.

Referencias

1. L. Li, “Data Security Technology in Electronic Commerce System Develop- ment,” International Conference on Applied Intelligence and Sustainable Com- puting, ICAISC 2023, 2023, doi: 10.1109/ICAISC58445.2023.10200851.
2. J. L. Gamboa Suárez, “IMPORTANCIA DE LA SEGURIDAD INFORMÁTICA Y CIBERSEGURIDAD EN EL MUNDO ACTUAL,” 2020.
3. M. Nicho, “A process model for implementing information systems security governance,” Information and Computer Security, vol. 26, no. 1, pp. 10–38, 2018, doi: 10.1108/ICS-07-2016-0061.
4. E. Vicente, A. Mateos, and A. Jiménez-Martín, “Risk analysis in information systems: A fuzzification of the MAGERIT methodology,” Knowl Based Syst, vol. 66, pp. 1–12, Aug. 2014, doi: 10.1016/J.KNOSYS.2014.02.018.

10
Desarrollo de propuesta de seguridad de la información basado en la Norma ISO/IEC 27002 en área tecnológica para una organización

Madeline Daniela Narváez[1][0009-0003-9078-7379]

Resumen La seguridad informática ha contribuido en la actualidad para cumplir diversas funciones empresariales, las cuales ha permitido el buen funcionamiento de sus servicios, sin embargo, en la actualidad estas han sido constantemente vulneradas y atacadas, perjudicando su confiabilidad, privacidad e integridad, por lo que se han implementado diferentes políticas, prácticas y protocolos para aplicar seguridad en los diversos tipos de sistemas informáticos, las cuales son de vital importancia en las empresas, universidades, entre otros sectores. En la actualidad, las redes de comunicaciones son un componente crítico en cualquier organización, y las universidades no son la excepción. En una reconocida Universidad de la ciudad de Pasto, la seguridad de la información es parte fundamental para la circulación de comunicación entre las diversas áreas y usuarios de la universidad, es por ello que la presente investigación se enfoca en la adopción de métodos orientadores a la gestión de las vulnerabilidades, relacionados con los procesos de la información, bajo la norma ISO/IEC 27002 [5], adicionalmente con el respaldo previo de la estimación de riesgo gracias a emplear el método MAGERIT v3 y la herramienta eMarisma, finalmente gracias a este estudio, se permitió establecer una revisión estructurada de la situación en cuanto a seguridad de la organización. Este documento está dirigido a ingenieros, administradores de sistemas y cualquier persona encargada de la dirección y el buen mantenimiento de los sistemas de información de Instituciones de Educación Superior, especialmente al área de tecnología.

Palabras clave Seguridad de la Información, Marisma, Magerit v3, ISO 27002.

10.1 Introducción

Los sistemas de información tienen un impacto significativo en la forma en que una empresa, una industria o un negocio deben adaptarse al entorno digital y a las nuevas corrientes de datos, en el contexto de las redes informáticas, son esenciales para llevar a cabo las actividades planificadas en cualquier organización. Estos sistemas desempeñan un papel fundamental en la transformación digital y en la gestión eficaz de la información en el ámbito empresarial. (Fernández et al, 2019). La ejecución de un estándar para fortalecer la seguridad de la información en las organizaciones definidas en la ISO 27002[5] para el aseguramiento de la confidencialidad, disponibilidad e integridad de la información, es una obligación para las Instituciones de educación superior en Colombia

[1]Universidad Mariana, Pasto, Colombia.
e-mail: madnarvaez@umariana.edu.co

exigido por el ente encargado MinTIC (Ministerio de Tecnologías de la Información y Comunicaciones de Colombia), se evidencia que actualmente las universidades alcanzan promedios de calificación bajas, debido a faltas de identificación desde su organización, impidiendo evaluar los controles a ser implantados en vista de las necesidades, por consiguiente, resultan importantes resaltar los aspectos del tratamiento de tal información en relación a sus procesos. En la actualidad las empresas están en la búsqueda de metodologías y procesos que les permitan adaptarse a las nuevas tecnologías, procurando contar con sistemas de in- formación automatizados, acceso usable y disponible, logrando con ello una comunicación idónea, como puede constatar en [1]. Las metodologías y técnicas empleadas por diversas empresas del sector de educación superior se han centrado en el control de la seguridad, estas normalmente se encuentran sustentadas en estudios de tipo subjetivo para determinar y analizar las vulnerabilidades y riesgos en tales campos.

Una reconocida Universidad de la ciudad de Pasto cuenta con una oferta de programas de pregrado y posgrado diversificada y renovada, en modalidad presencial, virtual, dual y de registro único. La Universidad cuenta con infraestructura tecnológica de conectividad para las sedes y ampliaciones, además de incluir el uso de plataformas tecnológicas para la gestión analítica de datos y toma de decisiones, cuenta con las certificaciones ISO mínimas en ciberseguridad, para la analítica de datos, gestión y toma de decisiones académicas y administrativas a través de desarrollo de software propio y licenciamiento de terceros, por lo cual se hace pertinente analizar estrategias para im- plementar dentro de la institución políticas basados en la norma ISO/TEC 27002 [5], que permita conservar la integridad de la información.

Para la institución superior el activo más importante se deriva en preservar la información, mayor aun en la actualidad donde los avances tecnológicos han llevado a cambios radicales en el campo educativo, las practicas inadecuadas por parte de los usuarios, virus propagados por diversos medios, ataques de hackers, influyen en la continuidad de los servicios, causando impactos negativos en la utilización de los sistemas de información [3], por lo cual surge la necesidad de realizar procedimientos que mitiguen ataques a la seguridad de la información donde se disipe la confidencialidad e integridad de los mismos.

10.2 Metodologías de análisis y gestión del riesgo

Para la investigación, fue necesario comparar metodologías de análisis y gestión del riesgo, lo que permitió seleccionar dos de ellas para el proyecto. A continuación, en la Tabla 10.1 se presenta información que facilita la decisión sobre cuál de estas metodologías es la más factible para la propuesta.

Para el contexto específico de la universidad, se optó por una combinación de ambas metodologías, utilizando los puntos fuertes de cada una para complementar las debilidades de la otra. La elección se justifica de la siguiente manera:

- La flexibilidad de Marisma permite una rápida adaptación a las necesidades específicas de la universidad, facilitando una implementación más ágil en comparación con la rigidez y exhaustividad de Magerit.
- Magerit aporta un enfoque más detallado y completo para la identificación y análisis de riesgos, lo cual es crucial para una institución que maneja una amplia variedad de datos sensibles y confidenciales.
- La combinación permite equilibrar la profundidad de análisis con la capacidad de implementación eficiente, asegurando que los controles y medidas de seguridad sean adecuados y efectivos sin ser prohibitivamente costosos o complejos.
- Utilizar elementos de Magerit proporciona un respaldo adicional en términos de reconocimiento y validez de la metodología, especialmente importante para auditorías y cumplimiento normativo.

Table 10.1: Contraste de Metodologías de Análisis y Gestión del Riesgo

	Magerit	**Marisma**	**Octave**	**Mehari**
Fases	1. Caracterización de los activos valoración, dimensionamiento. 2. Caracterización de las amenazas relación con activos, valoración (degradación y frecuencia). 3. Caracterización de las salvaguardas: tipología, valoración (nivel eficacia). 4. Estimación del estado deriesgo: impacto Y riesgo (intrínseco- residual).	1. Selección de controles. 2. Selección de criterios de riesgo. 3. Definición de relaciones entre Tipos de Activos por Amenazas por criterios de Riesgo. 4. Establecimiento de relaciones entre amenazas por controles.	1. Construcción de perfiles activo-menaza,identificación de contramedidas actuales. 2. Análisis de vulnerabilidades en la infraestructura. 3. Definición de riesgos prioritarios y sus correspondientes planes de mitigación.	1.Fase preparatoria: evaluación del contexto técnico y estructural de la organización. 2.Fase operacional: análisis de activos, evaluación la calidad de los servicios de seguridad y evaluación del riesgo. 3.Fase de tratamiento: priorización del riesgo, selec-ción de medidas e indicadores.
Ventajas	- Promueve la sensibilización a los responsables, sobre la existencia del riesgo. - Analiza los activos en varias dimensiones. - Correcta documentación en cuanto a tipos de activos de información y amenazas.	- Permite el mayor nivel posible de automatización y reutilización con la mínima cantidad de información, recogida en un tiempo muy reducido.	- Compromete a todos los actores de la organización. - Crea un ambiente colaborativo. - Involucra los procesos y aspectos organizacionales como elementos del modelo.	-Análisis de riesgos cuantitativa en base a fórmulas. - Permite identificar vulnerabilidades mediante auditorías.
Desventajas	No incluye los procesos, vulnerabilidades del modelo y aspectos organizacionales como elementos.	Se centra en la gestión de activos de información.	Dimensionamiento limitado.	Se centra específicamente en mostrar las debilidades de un sistema de información

La elección de estas metodologías para este proyecto está justificada por su adecuación al contexto específico de la universidad, su simplicidad y eficiencia en la gestión de riesgos, y su enfoque específico en la protección de activos educativos críticos.

10.3 Resultados

En este capítulo se presentará la parte esencial de la propuesta en cuestión, que consiste en la implementación del proceso de análisis de riesgos, este proceso se llevará a cabo mediante la utilización de la guía MAGERIT V3 y la metodología Marisma, con el objetivo de identificar y comprender los activos involucrados en la organización. Además, se abordará el desarrollo de las Políticas de Seguridad de la Información, haciendo referencia a los controles establecidos en la norma ISO/IEC 27002 [5]. Estas actividades se llevarán a cabo en el contexto de una reconocida universidad de la ciudad de Pasto, ubicada en Nariño, Colombia.

10.3.1 Análisis

Análisis de riesgos y amenazas del sistema de información, empleando la normativa 27001 para la gestión de la seguridad, cabe aclarar que la información presentada a continuación fue proporcionada gracias al centro de servicios informáticos de la universidad, con valores aproximados para su confidencialidad. Consecutivamente, se generan los 10 dominios y los 35 objetivos de control, seguidamente se desglosan los 114 controles para los siguientes objetivos de control definiendo su alcance. Señalado por el marco de la Norma ISO 270001, es importante categorizar los activos de información a mayor claridad, a continuación, en la tabla 2 se pretende establecer las categorías identificadas, para realizar la estimación del valor de la información contenida en este sistema se toma la referencia del costo de ficha de inventario por los 165.000 bienes que se albergan, de manera sintetizada los activos se resumen así:

Table 10.2: Activos por categoría

Nombre	Categoría	Valor
Datos u Información	Data	25’682.650,000
Equipo Auxiliar	Auxiliar	6’236.459,000
Hardware	HW	2’459.205,000
Instalaciones físicas	L	2’450.000,000
Servicios	S	5’000.000,000
Software	SW	2’000.000,000
TOTAL		18’145.664,000

10.3.1.1 Análisis y gestión del Riesgo

La organización dispone de una matriz dentro de las cuales se replica la estructura organizacional de esta manera se consolida su jerarquía respectiva. Por medio de un chekclist se agrupa los activos de la organización, para proceder con la respectiva eva- luación. Las 56 amenazas del contexto están

predefinidas por defecto, pero son susceptibles de cambiar debido a la naturaleza de la posibilidad de ocurrencia y porcentaje de degradación para este caso de estudio. De acuerdo con el análisis respectivo se identifican amenazas de gravedad, a continuación se visualiza un resumen.

10.3.1.2 Análisis y gestión del Riesgo

La organización dispone de una matriz dentro de las cuales se replica la estructura organizacional de esta manera se consolida su jerarquía respectiva. Por medio de un chekclist se agrupa los activos de la organización, para proceder con la respectiva eva- luación. Las 56 amenazas del contexto están predefinidas por defecto, pero son susceptibles de cambiar debido a la naturaleza de la posibilidad de ocurrencia y porcentaje de degradación para este caso de estudio. De acuerdo con el análisis respectivo se identifican ame nazas de gravedad, a continuación se visualiza un resumen.

Table 10.3: Categorización amenazas

Tipo	Código Amenaza	Nombre Amenaza	Probabilidad Ocurrencia	Porcentaje Degradación
Ataques intencionados	[A.25]	Robo	Alto (80.0%)	Alto (80.0%)
Errores y fallos no intencionados	[E.20]	Vulnerabilidades de los programas (software)	Medio (60.0%)	Alto (80.0%)
Ataques intencionados	[A.11]	Acceso no autorizado	Medio (60.0%)	Alto (80.0%)
Ataques intencionados	[A.12]	Análisis de tráfico	Alto (80.0%)	Alto (80.0%)
Errores y fallos no intencionados	[E.7]	Deficiencias en la organización	Alto (80.0%)	Medio (60.0%)

De manera similar, los 869 ítems referentes a los activos por amenazas resultantes, que representan la probabilidad de ocurrencia de las diversas amenazas con respecto a cada uno de los activos. Seguidamente se realiza una breve descripción de las amenazas más propensas si se persiste la ausencia de controles, la tabla 14 condensa la información previamente descrita.

La relación entre los activos y las amenazas se contrastan con la evaluación de los controles respectivos de cada objetivo de control que define el patrón dado por eMarisma.

En este punto es donde intervienen los principales actores de la Universidad en fin de que los conocimientos de cada área en función de sus competencias sean integrada y considerada en el análisis de los 114 controles, a continuación, se visualiza un resumen de como cada control mitiga una amenaza especifica relacionada con un activo particular. De manera planificada y previamente organizada, se desglosan las directrices que implican el contenido de cada control, gracias a la herramienta es posible realizar comentarios que permitan una inspección exhaustiva de cada gestión que se va a analizar.

10.4 Evaluación

Gracias a la herramienta es posible realizar una evaluación, en donde se obtiene una visión específica del análisis mediante las gráficas ilustrativas, en donde se indican porcentajes del cumplimiento de la Norma ISO 2700, el resultado general establece un 76% de cumplimiento.

Table 10.4: Amenaza por activo, descripción

Nombre de amenaza	Nombre Activo	Descripción
Acceso no autorizado	Matriculas estudiantes	Dado que es un recurso esencial que contiene información sensible sobrelos estudiantes de la universidad, la persistencia de esta amenaza podríacomprometer la disponibilidad y la integridad de los datos. La ausencia de políticas que regulan estos accesorios agrava la situación.
Análisis de tráfico	Red inalámbrica	En la organización, si no se implementan controles efectivos de monitoreo de tráfico, existe la posibilidad de que terceros realicen análisis detallados de origen, destino, volumen y frecuencia de intercambio de datos. Esto hace que las redes de comunicación sean vulnerables.
Robo	Aulas de informática	A pesar de que la universidad cuenta con sistemas de vigilancia en sus alrededores, no existen políticas de seguridad física definidas, aprobadas y comunicadas. Esto deja a la universidad vulnerable ante el robo de equipos de laboratorio.
Vulnerabilidades de los programas (software)	Webmaster	En la actualidad, el sistema carece de un mantenimiento regular por parte del equipo de TI, lo que lo exponen a Múltiples vulnerabilidades y lo hace susceptible a posibles ataques.
Deficiencias en la organización	Soporte técnico	Si la incorrecta definición de roles y las lagunas en las responsabilidades de los usuarios persisten y no se ajustan a las descripciones de trabajo correspondientes, es probable que se produzcan errores recurrentes en la información y en las operaciones de la organización.

Table 10.5: Mitigación de amenaza por control

Activo	Amenaza identificada	Control (ISO/IEC 27002)
Datos académicos	Acceso no autorizado	Control A.9.1: Gestión de accesos
Datos Financieros	Pérdida de datos	Control A.12.3: Respaldo de la información
Datos de Investigación	Modificación no autorizada	Control A.10.1: Criptografía

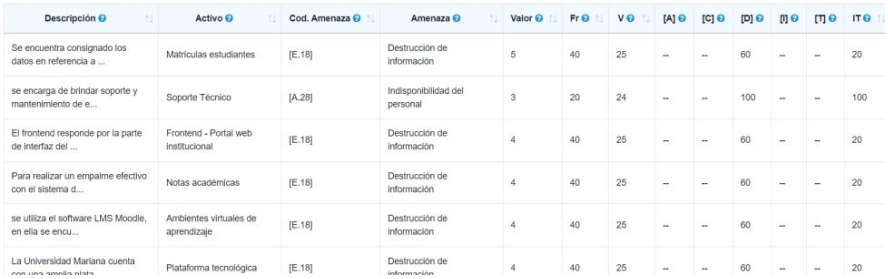
10.4.0.1 Análisis de Riesgos

Dado el análisis identificado anteriormente, se presenta a continuación en cuanto a riesgos, los hallazgos primarios, la utilidad de los riesgos oscilan dada su naturaleza con las instancias analizadas. La herramienta eMarisma permite por medio de tres fases valorar donde se registran la criticidad y riesgo de los activos, en mayor, menor y medio; a continuación, se muestra como la universidad requiere un tratamiento considerable; tomando gran importancia en el sistema de información que este presenta. Seguidamente, se presenta un resumen de la declaración de aplicabilidad, que incluyen los controles seleccionados de la norma ISO/IEC 27002 [5], la justificación de su selección y su relación con los riesgos identificados.

Posteriormente se identifica como la indisponibilidad del personal, la destrucción de información, puede repercutir de manera significativa un riesgo considerable, es importante mencionar que un valor considerable como lo es las políticas de acceso no tienen políticas de acceso estip-

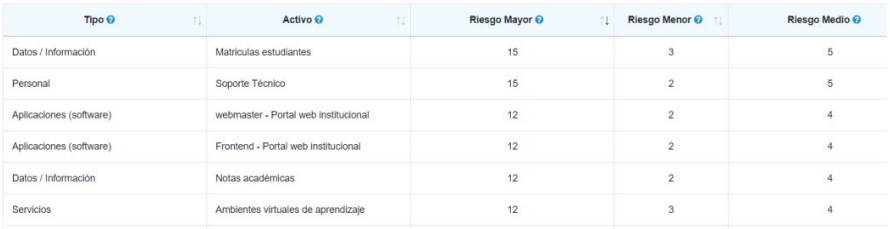
Table 10.6: Dominios con porcentaje

Dominio	**Porcentaje cumplimiento**
[A.5] - Políticas de Seguridad de la información	75%
[A.6] - Organización de la seguridad de la información	54%
[A.7] - Seguridad ligada a los recursos humanos	70%
[A.8] - Gestión de activos	82%
[A.9] - Control de acceso	90%
[A.10] - Criptografía	63%
[A.11] - Seguridad física y del entorno	85%
[A.12] - Seguridad de las operaciones	87%
[A.13] - Seguridad de las comunicaciones	75%
[A.14] - Adquisición, desarrollo y mantenimiento de los sistemas de información	92%
[A.15] - Relación con proveedores	61%
[A.16] - Gestión de incidentes de seguridad de la información	77%
[A.17] - Aspectos de seguridad de la información para la gestión de la continuidad del negocio	87%
[A.18] - Cumplimiento	83%



Descripción	Activo	Cod. Amenaza	Amenaza	Valor	Fr	V	[A]	[C]	[D]	[I]	[T]	IT
Se encuentra consignado los datos en referencia a ...	Matrículas estudiantes	[E.18]	Destrucción de información	5	40	25	--	--	60	--	--	20
se encarga de brindar soporte y mantenimiento de e...	Soporte Técnico	[A.28]	Indisponibilidad del personal	3	20	24	--	--	100	--	--	100
El frontend responde por la parte de interfaz del ...	Frontend - Portal web institucional	[E.18]	Destrucción de información	4	40	25	--	--	60	--	--	20
Para realizar un empalme efectivo con el sistema d...	Notas académicas	[E.18]	Destrucción de información	4	40	25	--	--	60	--	--	20
se utiliza el software LMS Moodle, en ella se encu...	Ambientes virtuales de aprendizaje	[E.18]	Destrucción de información	4	40	25	--	--	60	--	--	20
La Universidad Mariana cuenta con una amplia plata...	Plataforma tecnológica	[E.18]	Destrucción de información	4	40	25	--	--	60	--	--	20

Fig. 10.1: Grafica generada por e Marisma - instancias análisis de riesgos



Tipo	Activo	Riesgo Mayor	Riesgo Menor	Riesgo Medio
Datos / Información	Matrículas estudiantes	15	3	5
Personal	Soporte Técnico	15	2	5
Aplicaciones (software)	webmaster - Portal web institucional	12	2	4
Aplicaciones (software)	Frontend - Portal web institucional	12	2	4
Datos / Información	Notas académicas	12	2	4
Servicios	Ambientes virtuales de aprendizaje	12	3	4

Fig. 10.2: Grafica generada por e Marisma - instancias análisis de riesgos

Table 10.7: Declaración de aplicabilidad

ISO/IEC 27002	Descripción	Justificación de la Selección	Relación con Riesgos Identificados
A.9.1	Gestión de accesos	Necesario para proteger datos sensi bles de accesos no autorizados.	Amenaza de acceso no autorizado a datos académicos y financieros.
A.12.3	Respaldo de la información	Garantiza la disponibilidad y recuperación de datos en caso de pérdida.	Amenaza de pérdida de datos financieros.
A.10.1	Criptografía Protege la integridad y confidencialidad de la información sensible.	Amenaza de modificación no autori zada de datos de investigación.	

uladas, adicionalmente, la perdida accidental de infor-mación, significa una amenazas sobre los datos en general, dado que la información se encuentra en un soporte informativo significando amenazas específicas, la herramienta proporciona mapas de calor, identificando los riesgos con mayor criticidad en el que la universidad debe realizar un debido tratamiento, adicionalmente se logra apreciar la estructura de cada acción en su desempeño diario, con respecto al nivel de debilidad, es importante mencionar que la institución muestra niveles de riesgos altos, esto debido a la falta de controles que se han realizado. Seguidamente se procede a llevar a cabo el plan de tratamiento de riesgos, durante el proceso de análisis de riesgos, se identifican valores nominales significativamente al-tos, que oscilaron entre 6 y 5. En vista de esta situación, se confirma la utilidad

Orden	Grupo	Código	Control	Cobertura	R. Residual Bloque Inicial	R. Residual Bloque Final	R. Residual Inicial	R. Residual Final	VR	Responsable	F. Previst
18	Administración CHKL_1	[A.5.1.1]	Políticas para la seguridad de la información	0.75	50.428	22.028	50.428	50.0816	6	Centro de servicios informaticos	28/06/202 17:00:00
19	Administración CHKL_1	[A.5.1.2]	Revisión de la política de seguridad de la información	0.75	50.0816	3.42857	50.0816	49.3026	6	Centro de servicios informaticos	28/06/202 17:00:00

Fig. 10.3: Grafica generada por e Marisma - instancias análisis de riesgos

de la herramienta eMarisma, se su-giere que se aspira a alcanzar un nivel de riesgo deseado de 5, y se generan recomen-daciones para mejorar los controles pertinentes. El análisis del estado de riesgo en el sistema de información del ámbito de estudio reveló la existencia de posibles riesgos vinculados directamente a la ejecución de tareas por parte de empleados y administradores. Estos riesgos incluyen el manejo inadecuado de información confidencial, el uso incorrecto de equipos y servicios de red, errores en la configuración de recursos de hardware y software, entre otros. Estas amenazas fueron identificadas y documentadas a través de la Metodología MAGERIT. Además, se ob-servará un aumento en las solicitudes e incidentes relacionados con la mala gestión de activos, los cuales quedaron registrados en el sistema de seguimiento de tickets de la organización. Dadas estas circunstancias desfavorables, se hace necesaria la implemen-tación de regulaciones adecuadas para las actividades y comportamientos del personal, con el objetivo de fomentar una cultura organizacional comprometida y consciente de la importancia de la seguridad de la información. En este contexto, se procede a generar las políticas se Seguridad de la Información en base a normativa ISO/IEC 27002 [5] para la Universidad el cual consta de un documento de 35 hojas donde

se proponen políticas de seguridad de la información para el área de sistemas de la universidad mariana.

10.5 Conclusiones

Las políticas de seguridad implementadas, basadas en los controles de la ISO/IEC 27002 [5], han tenido un impacto significativo en varios aspectos clave de la seguridad de la información en la universidad:

- La implementación de controles como la gestión de accesos (Control A.9) y la criptografía (Control A.10) ha fortalecido la protección de datos académicos, fi-nancieros y de investigación. Esto ha reducido significativamente la probabilidad de accesos no autorizados y de la manipulación no autorizada de información crítica.
- La aplicación de controles relacionados con la seguridad física y ambiental (Con-trol A.11) ha mitigado riesgos asociados con desastres naturales y acceso físico no autorizado a los sistemas de información. Esto ha mejorado la resiliencia de la infraestructura tecnológica de la universidad.
- La implementación de programas de concienciación en seguridad de la informa-ción (Control A.7) ha incrementado la conciencia entre el personal y los estudian-tes sobre la importancia de la seguridad de la información, reduciendo el riesgo de incidentes causados. por errores humanos.

La implementación de políticas de seguridad basadas en la ISO/IEC 27002 [5] ha tenido un impacto significativo en la mejora de la seguridad de la información en la univer-sidad. Mediante la articulación clara entre amenazas y controles, la evaluación conti-nua de riesgos y la adopción de prácticas avanzadas de seguridad, la universidad ha fortalecido su infraestructura de seguridad, aumentado la confianza de sus usuarios y asegurado el cumplimiento normativo. Con las recomendaciones propuestas para la mejora continua, la universidad puede seguir avanzando hacia un entorno más seguro y resiliente.

Referencias

1. Anchundia, C. (2017). Ciberseguridad en los sistemas de información de las universidades. Dominio de Las Ciencias, 3(3), 200–217. https://doi.org/10.23857/dom.cien.pocaip.2017.3.mono1.ago.200-217 Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016)
2. Mena, N. (2018). Redes sociales y gestión de la información: un enfoque desde la teoría de grafos. Ciencias de la información y tecnología, 43, 10.Author, F.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010)
3. Murillo, S. (2019). La transferencia de los riesgos cibernéticos en empresas internacionales con alto nivel de capitalización bursátil. Revista de Pensamiento Estratégico y Seguridad CISDE, 3(1), 67-90.N
4. Crespo, E. y Cordero, G. (2018). Estudio comparativo entre las Metodologías CRAMM y MAGERIT para la gestión de riesgo de ti en las MPYMES. UDA AKADEM, (1).http://udaakadem.uazuay.edu.ec/article/view/129
5. International Organization for Standardization (ISO). (2013). Information technology — Security techniques — Code of practice for information security controls (estándar ISO/IEC 27002:2013). https://www.iso.org/standard/54533.html

11
Caracterización y procesamiento de señales de acelerómetros y giroscopios para evaluar la capacidad funcional del adulto mayor

Agustina Nappa[1]
Natalia Garay[2][0000-0002-3580-2412]
Franco Simini[3][0000-0003-0241-8737]

Abstract Aging can lead to a deterioration in quality of life, making it a significant challenge today. It is essential to have devices that can accurately and effectively assess the functional capacity of the elderly. Accelerometers and gyroscopes are analyzed as they are innovative tools for this evaluation. These non-invasive devices allow for continuous and precise measurements of movement and balance, which are crucial for detecting changes in a person's daily physical activity. For full-body analysis, a triaxial accelerometer is used near the center of mass, at the pelvis. The results show how these devices can provide precise data on body dynamics and stability, which is vital information for identifying movement patterns and potential fall risks in older adults

Keywords Accelerometry, inertial sensors, functional capacity, gait, balance, aging.

Resumen El envejecimiento puede llevar a un deterioro en la calidad de vida, por lo que es un gran desafío actualmente. Es necesario contar con dispositivos que evalúen de manera precisa y eficaz la capacidad funcional del adulto mayor. Se analizan acelerómetros y giroscopios ya que son herramientas innovadoras para esta evaluación. Son dispositivos no invasivos y permiten mediciones continuas y precisas del movimiento y equilibrio, esenciales para detectar cambios en la actividad física diaria de la persona. Para el análisis del cuerpo completo, se utiliza un acelerómetro triaxial cerca del centro de masas, en la pelvis. Los resultados muestran cómo estos dispositivos pueden proporcionar datos precisos sobre la dinámica y estabilidad corporal, que es una información muy importante para identificar patrones de movimiento y posibles riesgos de caídas en los adultos mayores.

Palabras clave: Acelerometría, sensores inerciales, capacidad funcional, marcha, equilibrio, envejecimiento.

Universidad de la República
Facultad de Ingeniería
[1]Estudiante del XXXIII Seminario de Ingeniería Biomédica 2024
[2,3]Docentes del XXXIII Seminario de Ingeniería Biomédica 2024.
e-mail: ngarayb@gmail.com

11.1 Introducción

El envejecimiento en las personas lleva consigo una serie de cambios físicos, cognitivos y sociales que pueden afectar su capacidad funcional y calidad de vida[1]. Mantener una buena capacidad funcional en la persona mayor está asociado con una mejor calidad de vida, menor riesgo de discapacidad y menor dependencia de cuidadores y sistemas de atención médica. En este contexto, la evaluación de la capacidad funcional en los adultos mayores es uno de los principales desafíos del siglo XXI[2]. Esta evaluación es crucial para proporcionar a los pacientes la atención que necesitan según su diagnóstico. Además, es necesario que el dispositivo utilizado para la evaluación sea cómodo tanto para el paciente como para el personal de salud que lo utilice. Una de las herramientas más prometedoras para realizar esta evaluación son los acelerómetros y giroscopios, sensores portátiles que permiten medir el movimiento y la actividad física de manera objetiva y precisa[3]. Los acelerómetros y giroscopios ofrecen varias ventajas para la evaluación de la capacidad funcional en adultos mayores. Estos dispositivos son pequeños, ligeros y no invasivos, lo que los convierte en el método más eficiente en la actualidad. Además, pueden proporcionar mediciones continuas y a largo plazo del movimiento, lo que permite detectar cambios sutiles en la actividad física y el equilibrio a lo largo del tiempo[4]. Sin la incorporación de tecnologías como los acelerómetros y giroscopios, la evaluación de la capacidad funcional en adultos mayores suele depender en gran medida de métodos subjetivos y observacionales. Los profesionales de la salud utilizan escalas de evaluación basadas en la observación directa del paciente durante actividades cotidianas, cuestionarios autoadministrados o pruebas específicas de movilidad y equilibrio. Sin embargo, estos métodos presentan limitaciones en términos de objetividad, fiabilidad y precisión[5]. En esta monografía se analiza el funcionamiento de estos sensores, en particular para obtener señales de las mediciones de la evaluación del equilibrio, ritmo de marcha y control al levantarse-sentarse.

11.2 Metodología

Para realizar este trabajo se llevó a cabo una revisión bibliográfica dirigida, utilizando las palabras claves ‘giroscopio’, ‘acelerómetro’, ‘adulto mayor’. Para buscar el material se utilizaron los portales Timbó Foco y Google Scholar.

11.3 Principios de funcionamiento de acelerómetros y giroscopios

La capacidad funcional se define como la aptitud de un individuo para realizar tareas cotidianas. En el adulto mayor, esta capacidad se ve afectada por una disfunción cognitiva principalmente física, que generalmente afecta la memoria operativa, flexibilidad cognitiva y atención dividida o selectiva. Es importante tener un método preciso y certero para medir la capacidad funcional, ya que esto permite tener un diagnóstico y que el personal de salud brinde las herramientas necesarias para poder mejorar la calidad de vida del adulto mayor [6]. En esta sección se describen los dispositivos inerciales utilizados para la evaluación de la capacidad funcional en los adultos mayores cuando realizan distintas tareas: mantener el equilibrio, caminar y sentarse y levantarse de una silla. Los acelerómetros y giroscopios son dispositivos denominados IMU (Unidad de medición inercial). Estos sensores se utilizan en la captura de movimiento, y para ello se componen de tres ejes perpendiculares en los que cada uno de los ejes contiene un acelerómetro y un giroscopio. De esta forma, los acelerómetros miden la aceleración lineal en tres dimensiones, mientras que los giroscopios miden la velocidad angular. En conjunto, estos sensores se utilizan para rastrear la posición y la orientación de un objeto o individuo en relación con un punto de inicio y la velocidad

[7]. Además, para una mayor precisión se pueden incluir circuitos electrónicos para acondicionar y procesar las señales generadas por los sensores, así como algoritmos que a partir de las medidas de aceleración y velocidad angular obtengan valores de la orientación del dispositivo.

11.3.1 Acelerómetros

Los acelerómetros son sensores capaces de registrar la aceleración de los objetos o individuos. Estos sensores son portátiles, cómodos y fáciles de fijarse al cuerpo del individuo. Pueden ser de tres tipos dependiendo del número de ejes en los que se registren las aceleraciones: uniaxiales, que registran los movimientos en un único eje (generalmente vertical), los biaxiales, sensibles a movimientos en dos ejes (vertical y antero-posterior) y triaxiales, sensibles a movimientos en los tres ejes (vertical, antero-posterior y lateral). En la actualidad, los acelerómetros comerciales son triaxiales, debido a que estos dispositivos proporcionan más información para ser utilizada en los modelos de estimación de la capacidad funcional [7]. Los acelerómetros basan su funcionamiento en la Segunda Ley de Newton. Proporcionan una medida de la segunda derivada de la posición (ecuación de segundo orden), es decir de la aceleración. Al resolver la ecuación de segundo orden con sus respectivas condiciones iniciales, es posible obtener la velocidad con la cual se mueve el objeto o individuo, así como la posición en la cual se encuentra, tomando como punto de referencia (o condición inicial), la lectura de cuando acelerómetro se encuentra en reposo, en base a su inclinación con respecto a la horizontal [8]. Existen diversos tipos de acelerómetros. Los más utilizados actualmente son los acelerómetros piezoeléctricos, capacitivos, los cuales pertenecen a un grupo denominado sistemas micro-electro-mecánicos (MEMS) y piezorresistivos. Cada tipo de acelerómetro tiene sus propias ventajas y es seleccionado según la aplicación específica, las condiciones de operación y los requisitos de precisión y costo. Los que analizan de forma más precisa el movimiento humano, y por lo tanto los más utilizados para esto en la actualidad son los acelerómetros capacitivos. Dichos acelerómetros contienen dos placas conductoras en su interior: una estática, utilizada como punto de referencia, y otra dinámica, que se mueve (ya que está sostenida por pequeños resortes) junto con la masa cuando esta es afectada por alguna aceleración[9]. La capacitancia está dada por la ecuación 11.3.1:

$$C = \frac{\varepsilon_0 A}{d} \tag{11.1}$$

donde ε es la constante dieléctrica (en este caso del aire), A es el área de las placas, y d es la distancia entre las placas. Por lo tanto, la capacitancia únicamente varía en función de la distancia entre las placas. Si se parte de un estado de reposo con una cantidad estable de cargas eléctricas acumuladas en las placas del capacitor, y debido a una aceleración estas se comienzan a juntar, la capacitancia aumentará. Al ocurrir esto, una mayor cantidad de portadores de carga se moverán hacia las placas, generando una corriente eléctrica que puede ser medida. Lo contrario ocurre cuando la masa se comienza a alejar, la capacitancia disminuirá, los portadores de carga se alejarán de las placas, y se generará una corriente en el sentido opuesto. A partir de estas corrientes se obtiene el voltaje, el cual es pasado por filtros para eliminar las componentes del ruido no deseado de vibraciones o del ambiente, y de esta forma obtener una señal de voltaje que varía suavemente con el tiempo, representando la aceleración [10]. Este tipo de acelerómetros se realizan en tamaños muy pequeños, de forma que puedan ser colocados en dispositivos electrónicos. Sin embargo, la desventaja que presentan es que al utilizar el aire como constante dieléctrica del capacitor, cuando esta cambia (por ejemplo, al aumentar la humedad en el ambiente), la precisión del sensor se puede ver afectada.

11.3.2 Giroscopios

Los giroscopios son dispositivos que pueden cumplir dos funciones: entregar información sobre la variación de la orientación de un sistema con respecto a un eje de referencia, o entregar información sobre la velocidad con que varía la orientación de un sistema cuando se encuentra rotando, es decir, su velocidad angular[9]. Existen tres tipos de giroscopios: Mecánicos, vibratorios de efecto Coriolis, y ópticos. Los que analizan de forma más precisa el movimiento humano son los vibratorios, que al igual que los acelerómetros capacitivos, pertenecen al grupo de sensores MEMS, debido a su tamaño reducido para poder integrarse en pulseras u otros dispositivos electrónicos cómodos para el paciente. Los giroscopios vibratorios basan su funcionamiento en la aceleración de Coriolis. Esta aceleración es siempre perpendicular al eje de rotación del sistema y a las componentes radial y tangencial de la velocidad del cuerpo. Utilizando la Segunda Ley de Newton, la velocidad angular del sistema se puede expresar como (ver ecuación 11.3.2):

$$Velocidadangular = \frac{Fuerza}{Masa \times (-2 \times Veloc.Radial} \tag{11.2}$$

Basándose en la ecuación anterior, el giroscopio contiene en su interior una masa, la cual es forzada a oscilar con una frecuencia determinada. Debido a esto, cuando el sistema se ha rotado, la masa oscilante experimentará una frecuencia de Coriolis que la desplazará hacia la izquierda o la derecha dependiendo de la dirección de la vibración. De manera similar a los acelerómetros, este desplazamiento puede ser utilizado para calcular la fuerza que experimentó la masa, y de esta forma obtener la velocidad angular del sistema [11].

11.4 Medición de algunos movimientos del adulto mayor: Marchar, mantener el equilibrio, y sentarse y levantarse de una silla.

El cuerpo humano realiza numerosos movimientos medibles, y dependiendo del objetivo del estudio, se colocan acelerómetros en ubicaciones específicas. En estudios que analizan el cuerpo completo, se utiliza un acelerómetro triaxial cerca del centro de masas, en la pelvis. Además, se deben filtrar y minimizar las aceleraciones no deseadas, como las causadas por movimientos de tejidos blandos o vibraciones externas. A continuación, se presentan las gráficas obtenidas en el estudio, mostrando los movimientos medidos con acelerómetros y giroscopios en distintas partes del cuerpo.

11.4.1 Marcha

La Figura 11.1 Muestra las componentes de la aceleración en los ejes X, Y y Z obtenidas mediante un acelerómetro triaxial en la zona lumbar de un sujeto durante la marcha. La aceleración en el eje X refleja el balanceo lateral del tronco, mientras que la aceleración en el eje Y representa movimientos anteroposteriores, típicos de una marcha estable. La componente más relevante es la aceleración en el eje Z, que indica los impactos de los pies con el suelo durante el apoyo. Los picos en la aceleración Z corresponden a eventos de apoyo, y su consistencia refleja la regularidad de la marcha. Variaciones significativas en la altura y el espaciado de estos picos pueden indicar problemas de equilibrio o coordinación. La gráfica también permite identificar diferencias en los patrones de aceleración entre individuos con y sin predisposición a caídas, proporcionando información

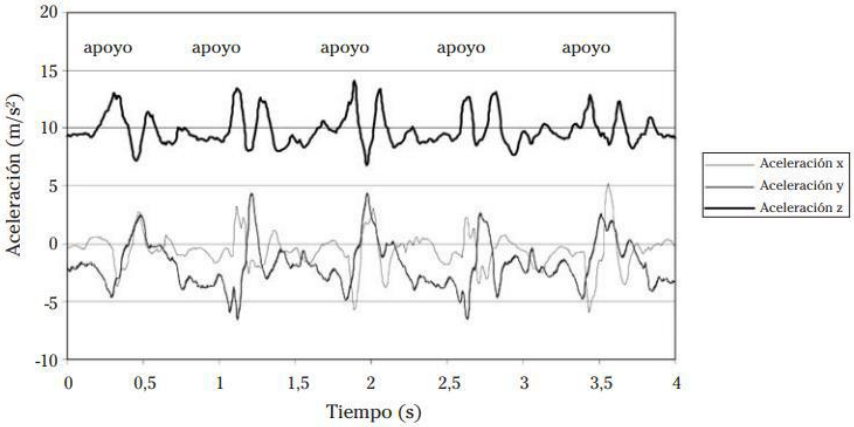


Fig. 11.1: Aceleraciones obtenidas durante la marcha con un acelerómetro triaxial colocado en la zona lumbar del sujeto. Obtenida de: [12].

valiosa para intervenciones preventivas. En resumen, esta gráfica ofrece una visión detallada del comportamiento del tronco durante la marcha, siendo esencial para evaluar la estabilidad y la regularidad del movimiento, especialmente en estudios de prevención de caídas en adultos mayores.

11.4.2 Sentarse y levantarse de una silla

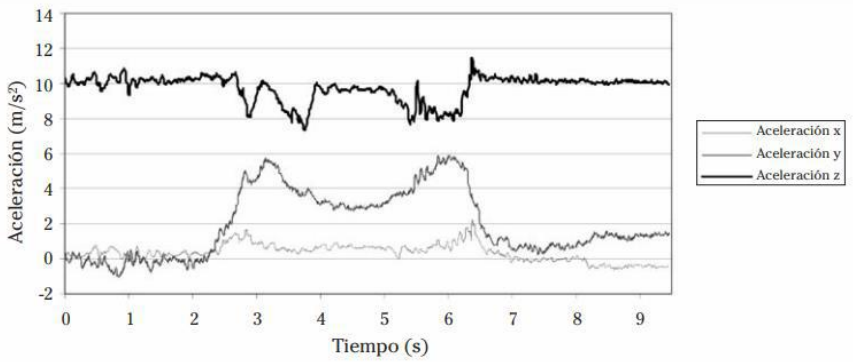


Fig. 11.2: Aceleraciones obtenidas durante el movimiento de sentarse y levantarse con un acelerómetro triaxial colocado en la zona lumbar. Obtenida de: [12].

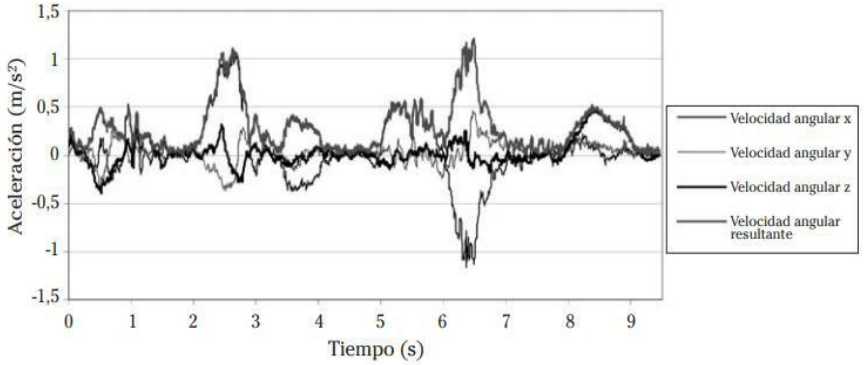


Fig. 11.3: Velocidades angulares obtenidas durante el movimiento de sentarse y levantarse con un giroscopio triaxial colocado en la zona lumbar. Obtenida de: [12].

La Figura 11.2 muestra las aceleraciones en los ejes X, Y y Z durante los ejercicios de levantarse y sentarse, registradas con un acelerómetro triaxial en la zona lumbar. Las máximas aceleraciones coinciden con los momentos de impulso. El eje X refleja movimientos laterales, el eje Y muestra movimientos anteroposteriores pronunciados y el eje Z picos claros en los impulsos, destacando el esfuerzo vertical. Estas aceleraciones ayudan a entender la dinámica del movimiento y evaluar la fuerza y estabilidad del sujeto. Por otro lado, la Figura 11.3 presenta las velocidades angulares en los ejes X, Y y Z, medidas con un giroscopio triaxial en la misma ubicación. La velocidad angular en el eje X indica la rotación lateral del tronco, en el eje Y la rotación anteroposterior y en el eje Z la rotación alrededor del eje vertical, crucial para el equilibrio. Estas velocidades angulares evalúan la agilidad y control del movimiento, siendo mayores en sujetos con mejor estado físico y coordinación.

Las gráficas combinadas de aceleraciones y velocidades angulares ofrecen una visión completa de los movimientos al sentarse y levantarse, permitiendo identificar patrones y variaciones indicativos de la funcionalidad y posibles problemas de movilidad. Estas medidas son esenciales para evaluar la capacidad funcional en adultos mayores y diseñar intervenciones para mejorar su estabilidad y prevenir caídas.

11.4.3 Equilibrio

La evaluación del equilibrio en adultos mayores es fundamental para identificar riesgos de caídas y problemas relacionados con la estabilidad postural, es decir, la habilidad para mantener una postura, sin necesidad de mover la base de apoyo. Esta tarea implica la integración de la información sensorial del cuerpo, abarcando tanto la percepción de la posición corporal en relación con el entorno como la capacidad para generar las fuerzas necesarias para controlar el movimiento. La señal obtenida de un acelerómetro en el sacro durante la evaluación del equilibrio postural ofrece dos parámetros clave: la amplitud y la frecuencia. Estos parámetros reflejan los movimientos del centro de masa del cuerpo, y su análisis puede revelar signos de inestabilidad postural. Las amplitudes grandes en la señal del acelerómetro indican mayores desplazamientos del cuerpo en respuesta a pequeños desequilibrios, lo cual es un claro signo de inestabilidad. Por otro lado, las frecuencias altas en la señal son indicativas de movimientos rápidos y erráticos del cuerpo, sugiriendo una menor capacidad para

mantener el equilibrio. Comparar la amplitud y frecuencia de las señales del acelerómetro bajo distintas condiciones permite determinar cómo varía el control corporal del adulto mayor bajo distintos desafíos. Por ejemplo, evaluar el equilibrio de pie en una superficie firme proporciona una línea base de control postural, mientras que estar de pie en una superficie blanda aumenta la dificultad del control postural y puede revelar problemas de equilibrio que no se manifiestan en superficies firmes. Asimismo, la postura con los pies juntos reduce la base de apoyo y aumenta la dificultad para mantener el equilibrio, mientras que con los pies separados se ofrece una base de apoyo más amplia y generalmente más fácil de mantener. Por otro lado, la evaluación con los ojos abiertos proporciona una referencia básica, ya que la visión ofrece información crucial para el equilibrio. Sin embargo, cerrar los ojos desafía el sistema de equilibrio y revela la dependencia en los otros sistemas sensoriales, como los propioceptivos y vestibulares.

11.5 Discusión y conclusion

Los acelerómetros y giroscopios, al ser pequeños, ligeros y no invasivos, pueden integrarse fácilmente en dispositivos portátiles que los pacientes pueden usar durante sus actividades cotidianas. Esto facilita la recopilación de datos en tiempo real y a largo plazo, proporcionando una visión más completa y precisa del estado funcional del paciente. La implementación de acelerómetros y giroscopios en la evaluación de la capacidad funcional de los adultos mayores ofrece una ventaja significativa sobre los métodos tradicionales. Estos dispositivos permiten mediciones objetivas y continuas, lo cual es esencial para detectar cambios sutiles en la movilidad y el equilibrio que no podrían identificarse con métodos observacionales y subjetivos. Además de ser herramientas eficaces y fiables, tienen un bajo costo económico: con un solo acelerómetro triaxial en la cintura, se pueden medir movimientos, actividad física, gasto energético, equilibrio y otras funciones esenciales. La combinación con giroscopios mejora aún más la precisión al añadir datos sobre orientación y cambios de posición.

Referencias

1. Biology of Aging, National Institute on Aging. Accedido: 21 de junio de 2024. [En línea]. Disponible en: https://www.nia.nih.gov/about/budget/biology-aging-3
2. Changes That Occur to the Aging Brain — Columbia Mailman, Columbia University Mailman School of Public Health. Accedido: 21 de junio de 2024. [En línea]. Disponible en: https://www.publichealth.columbia.edu/news/changes-occur-aging-brain-what-happens-when-we-get-older
3. Neuromusculoskeletal Research Laboratory. Accedido: 21 de junio de 2024. [En línea]. Disponible en: https://research.ohio.edu/clark-brian/publications/impaired-dopaminergic-function-associated-mobility-capacity-older-adults
4. F. Porciuncula et al., "Wearable Movement Sensors for Rehabilitation: A Focused Review of Technological and Clinical Advances", PM&R, vol. 10, n.o 9S2, sep. 2018, doi: 10.1016/j.pmrj.2018.06.013.
5. M. C. Vargas-Del-Valle et al., "Implementaciones tecnológicas en la prueba de valoración funcional y desempeño corto Short Physical Performance Battery (SPPB), para el adulto mayor", Rev. Tecnol. En Marcha, vol. 35, n.o 2, pp. 125-138, jun. 2022, doi: 10.18845/tm.v35i2.5206.
6. Z. E. Leitón Espinoza et al., "Cognición y capacidad funcional en el adulto mayor", Rev. Salud Uninorte, vol. 36, n.o 1, pp. 124-139, abr. 2020, doi:10.14482/sun.36.1.618.97.

7. M. Filho y J. Alves, “Nuevas perspectivas en la metodología de validación de acelerómetros para la estimación de la actividad física en adultos mayores”, 2015, Accedido: 19 de mayo de 2024. [En línea]. Disponible en: http://hdl.handle.net/10550/50825
8. D. Octavio et al., DETECCIÓN DE MOVIMIENTOS ESPECÍFICOS DE LA CABEZA CON EL ACELERÓMETRO Y GIROSCOPIO, MEDIANTE CORRELACIÓN CRUZADA. 2013.
9. J. F. G. García, “Tesis para cubrir parcialmente los requisitos necesarios para obtener el grado de Maestro en Ciencias”.
10. M. J. M. Ayora, “Análisis Experimental de la Repuesta de un Acelerómetro MEMS”.
11. D. F. P. Espín y N. Sotomayor, “DISEÑO Y CONSTRUCCIÓN DE UNA PLATAFORMA DIDÁCTICA PARA MEDIR ÁNGULOS DE INCLINACIÓN USANDO SENSORES INERCIALES COMO ACELERÓMETRO Y GIROSCOPIO”.
12. M. Izquierdo, A. Martínez-Ramírez, J. L. Larrión, M. Irujo-Espinosa, y M. Gómez, “Valoración de la capacidad funcional en el ámbito domiciliario y en la clínica: Nuevas posibilidades de aplicación de la acelerometría para la valoración de la marcha, equilibrio y potencia muscular en personas mayores”, An. Sist. Sanit. Navar., vol. 31, n.o 2, ago. 2008, doi:10.4321/S1137-66272008000300006.

12
Propuesta de integración tecnológica basada en el modelo SAMR para educación básica y media secundaria en la Institución Educativa Supía.

Marla Alejandra López Cerón[1][0009-0004-2957-9340]]
Sergio Rodríguez Jerez[2][0000-0002-3521- 0206]

Resumen En este artículo se propone la integración tecnológica en la Institución Educativa Supía, para articular los procesos de enseñanza-aprendizaje en básica y media secundaria. Lo anterior como producto de las reflexiones y exigencias resultantes del periodo de educación virtual/a distancia dado por la crisis del COVID-19. Se inició con la descripción del modelo pedagógico empleado en la institución, y el contenido de las guías de auto instrucción para su adaptación a experiencias de aprendizaje virtuales o híbridas; luego, se realizó un diagnóstico de habilidades digitales orientada a docentes y estudiantes, para finalmente desarrollar la propuesta a partir de estrategias didácticas, con los recursos disponibles de acuerdo con el contexto del municipio de Supía, Caldas. Se llevó a cabo una prueba piloto con los estudiantes y docentes de la Institución, mostrándose a favor con la propuesta. Se espera a futuro su implementación por completo en la Institución Educativa.

Palabras clave Educación mediada por TIC, Modelo SAMR, Didáctica digital, Learning design, Clases virtuales.

12.1 Introducción

Al decretarse el aislamiento preventivo en Colombia a causa del COVID-19, las instituciones educativas recurrieron a otras modalidades para dar continuidad al servicio educativo. Según el Boletín Estadístico de la Secretaría de Educación de Caldas [1] la tasa de cobertura en educación básica y media secundaria de los colegios oficiales era del 89,91%. Para reducir la tasa de deserción, se les dio autonomía a las instituciones educativas para garantizar un servicio flexible durante la cuarentena del 2020. En este proceso se percibieron debilidades en la calidad educativa. El secretario departamental de educación, Fabio Hernando Arias Orozco, expresó que: En Colombia no se puede hablar de educación virtual, sino a distancia, pues las condiciones de conectividad no son las mejores, específicamente en las zonas rurales. Por esta razón, docentes y rectores han trabajado arduamente para dar contenidos pedagógicos en material físico y hacer el acompañamiento en un 100% por medio de llamadas y WhatsApp [2]. Así fue el caso de la Institución Educativa Supía, del departamento de Caldas, caracterizada por tener población flotante y donde el 30% de los

[1]Institución Educativa Supía, Supía, Colombia
e-mail: maalopezce@unal.edu.co
[2] Universidad Sergio Arboleda, Bogotá, Colombia
e-mail: sergio.rodriguez@usa.edu.co

estudiantes de esta institución, residen en zona rural con diversas condiciones de acceso a recursos tecnológicos. Lo anterior de acuerdo con la información expuesta en la audiencia de rendición de cuentas de la Institución Educativa Supía del año 2022. El colegio utiliza como instrumento estándar, guías de auto instrucción que permiten al estudiante apoyarse del contenido en caso de no poder asistir a la institución, situación que es frecuente por problemas viales y la falta de transporte escolar por temporadas. Sin pautas claras de educación virtual, se asignaron las mismas guías durante el confinamiento para evitar la brecha de acceso. pero, en su mayoría manifestaron frustración por no comprender los contenidos ante la ausencia del maestro, hasta el punto de desertar de forma voluntaria hasta que se aprobaran las clases presenciales. Al cabo de un mes, se añadieron clases con diferentes plataformas; pero, los estudiantes que lograron participar estas clases manifestaron una frustración similar, esta vez, por la falta de adaptación. Si bien en Colombia existe un "Marco de competencias digitales docentes". No existe un protocolo de estrategias didácticas digitales o un marco de transformación digital educativo, que permita la integración paulatina de tecnologías al aula de clase. Así surge la pregunta de investigación: ¿Cómo articular por medio de un modelo de integración tecnológica la transición entre educación presencial y virtual en la Institución Educativa Supía?

12.2 Trabajos relacionados

En España [3] se creó una política de transformación digital educativa en ámbitos no universitarios, ejecutándose a partir de lo que denominan Plan de Actuación Digital (PAD) un plan de acción diseñado para cada plantel educativo, que incluya estrategias de mejora para ser "competente" en el ámbito digital. Por su parte en México [4a] deciden abordar la perspectiva docente de la transformación digital educativa, analizando modelos existentes como TAM, TPACK Y SAMR en conjunto, para indicar qué elementos tener en cuenta a la hora de implementar un sistema de gestión de aprendizaje en educación superior, exponiendo que es necesario que las autoridades institucionales visibilicen y emprendan acciones para integrar tecnologías en el aula. Desde Brasil, el artículo de Brunetti [5] complementa el estudio anterior, con el objetivo de comprender las perspectivas metodológicas que influyen en el docente del siglo XXI, concluyendo que primero se debe conocer el nivel de apropiación del docente en herramientas tecnológicas antes de seleccionar un modelo de integración TIC en el aula. Al final incluye el marco europeo para la competencia digital de los docentes como propuesta de exploración, instrucción e integración. Mientras tanto en Suramérica, los estudios relacionados se centran en las experiencias obtenidas tras el cambio abrupto de educación presencial a virtual tras el asilamiento preventivo. En Ecuador [6] plantearon estrategias de enseñanza virtual por medio de programas de radio, televisión, plataformas educativas asíncronas, esto como resultado de la falta de preparación previa de docentes y padres de familia para afrontar el sistema apresurado de educación en casa. En zona rural de Argentina [7] crearon escuelas con modalidad virtual-rural, donde las clases eran virtuales o en línea, y se contaba con un coordinador para mediar el proceso de los estudiantes, facilitando el acceso a educación primaria, secundaria, multigrado y adulto en sus territorios, ayudando a evitar el" desarraigo" y fortalecer los sectores productivos (agrícola y turístico) de su entorno. Para Colombia en el año 2013, el Ministerio de Educación Nacional (MEN), crea un documento llamado "Competencias TIC para el desarrollo profesional docente" con el fin de "orientar los procesos formativos haciendo uso de las TIC", [8]. El archivo ilustra 5 competencias (Tecnológica, comunicativa, pedagógica, de gestión, e investigativa), que permiten conocer al docente en qué nivel de apropiación se encuentra; además, proporciona orientaciones mínimas para el docente y el directivo docente. Sin embargo, las orientaciones proporcionadas apuntan una lista de chequeo que el docente debe completar por su propia cuenta. Después de la pandemia, el MEN reconoce a través de su portal "Colombia Aprende" que no cuenta con un marco estándar de competencias digitales docentes para Latinoamérica, tal como presenta la Unión Europea con su "Marco común de competencias digitales docentes" y el "Marco de competencias docentes en materia de TIC" de la UNESCO, debido al nivel de

desigualdad en materia de acceso y políticas de inclusión digital. Por lo anterior, es notable la falta de estudios enfocados en experiencias de integración digital educativa orientada a los colegios públicos; si bien, el enfoque en Instituciones de Educación Superior puede ser un punto de partida, es necesario el enfoque en educación básica y media secundaria, con el planteamiento de estrategias concretas en el aspecto pedagógico e instrumental a través de un plan de acción o modelo de integración con el cual se puedan cumplir con las competencias digitales docentes pertinentes para el adecuado proceso de enseñanza-aprendizaje en post pandemia.

12.3 Marco teórico

Dada la necesidad de una propuesta que contenga reglas definidas y pautas es importante conocer qué modelo de integración tecnológica se puede adaptar al modelo pedagógico de la Institución Educativa Supía. Para esto se realizó una comparación de diferentes modelos de integración digital a nivel educativo, entre estos los modelos TAM3, TPACK, SAMR, y la consulta de estrategias didácticas como el aprendizaje autorregulado y Microlearning que permita a los estudiantes y docentes a estar preparados para fortalecer los procesos de enseñanza-aprendizaje.

12.3.1 Modelo de Aceptación Tecnológica TAM

Un modelo que considera la transición hacia la virtualidad o aprovechamiento de los recursos tecnológicos en el proceso de enseñanza-aprendizaje, es el modelo TAM o Modelo de Aceptación Digital, creado por Davis en 1989 [4b]. Para su versión TAM3, aborda una serie de criterios para determinar qué tan útil será para el usuario final, y otro grupo de variables que determinarán la percepción de facilidad de uso que tiene. Sin embargo, es importante analizar si los resultados de este modelo se realizan bajo la perspectiva del estudiante, del docente, o de la comunidad en general.

12.3.2 Modelo TPACK

Por otro lado, existe el modelo TPACK que pretende recopilar 3 elementos fundamentales para la integración tecnológica en la educación: El conocimiento en su área, de las herramientas tecnológicas a emplear y conocimiento pleno del contenido. [9] Este proceso tiene la desventaja de que podría tomar bastante tiempo si se tiene en cuenta que la Institución debe contar con docentes altamente cualificados en el dominio tecnológico y manejo de contenidos. Si bien es útil en la educación superior [10] es más teórico que práctico, por lo cual exigiría un estudio más profundo si se quiere implementar en secundaria.

12.3.3 Modelo SAMR

El modelo SAMR (Sustituir, Aumentar, Modificar, Redefinir) es un modelo propuesto por Rubén Puentedura, el cual plantea cuatro niveles de integración de la tecnología en los ámbitos educativos, desde la digitalización de procesos hasta la transformación de actividades que no podían concebirse

en la escuela tradicional y que son posibles gracias a la incorporación de recursos TIC [11]. El ascenso entre cada nivel es flexible, permitiendo la integración hasta generar aprendizajes que habría sido difíciles de alcanzar en el aula tradicional. Se acomoda al contexto, dado que no exige para su ejecución inicial cambios en infraestructura tecnológica y que cada docente puede iniciar desde la fase que su habilidad y recursos tecnológicos inmediatos le permitan para la ejecución de sus clases. Para que este modelo se pueda aplicar desde su fase inicial (Sustituir) se debe articular con elementos de aprendizaje basado en micro contenidos (microlearning). El Micro aprendizaje o Microlearning, de acuerdo con el libro Introduction to Microlearning Course [12] es el proceso de aprender por medio de cápsulas de contenido bien planeadas en complemento con actividades, siempre y cuando estas sean de desarrollar en el corto plazo. La ventaja del microlearning, es que no se requiere el uso de dispositivos móviles para adaptar los tópicos de clase. Estos se pueden mostrar en los monitores o en los computadores de la institución.

12.4 Metodología de la Investigación

Esta investigación tiene un enfoque de carácter cualitativo basado en el paradigma hermenéutico. Ya que se busca la exploración de estrategias viables en el contexto del municipio de Supía, Caldas que permitan construir ese proceso de integración tecnológica educativa en la básica y media secundaria, en conjunto con los miembros de la Institución Educativa Supía. Se llevó a cabo la metodología de Investigación-Acción. Es adecuada para implementar acciones de cambio en el plantel educativo, ya que contribuye a la identificación del problema sin separarse del contexto y estudiarse para contribuir con la mejora o solución de la necesidad hallada [13]. Dentro del proceso se contempla la revisión y selección de modelos como punto de partida que puede contribuir al previamente señalado, entre estos modelos se encuentran los modelos TAM, TPACK y SAMR. De acuerdo con Hernández, Fernández y Baptista [14] es propicio iniciar este proceso investigación-acción para tener un diagnóstico, en este caso, del nivel de experiencia en el uso de recursos TIC en los estudiantes y docentes de la Institución Educativa Supía, y revisión documental de las prácticas educativas usadas en modalidad presencial y virtual. Lo anterior para un avance en la creación de la propuesta de solución y generar una acción de cambio a través de un modelo de integración tecnológico que pueda generar parámetros claros y pertinentes para abordar la modalidad virtual y un regreso flexible a la modalidad presencial, para que sea aplicado como un proceso continuo. Para la recolección de información se crean instrumentos asociados a una técnica específica con cuatro fases planteadas de la siguiente manera:

12.4.1 Descripción del modelo pedagógico empleado en la I.E Supía

En esta fase se describe el modelo pedagógico empleado en básica y media secundaria de la Institución Educativa Supía y el contenido de las guías de auto instrucción para su adaptación efectiva en modalidad virtual/híbrida. Se tomó en cuenta el contexto social, económico y académico de la comunidad educativa y el nivel de articulación del plan de estudios en las áreas fundamentales con los contenidos (a través de la guía). Para esto se empleó la técnica de revisión documental aplicada por medio de una matriz descriptiva de las guías de auto instrucción de 6 colegios públicos (urbanos y rurales) del municipio de Supía, con el fin de contrastar la estructura del modelo pedagógico de la Institución Educativa Supía con las demás instituciones. Como resultado todas coinciden con el modelo Escuela Activa Urbana (EAU), con 4 momentos: A(presaber), B(teoría), C(aplicación) D (complementación) por lo cual, se continuará así.

12.4.2 Diagnóstico del nivel de apropiación tecnológica en estudiantes y docentes

Se aplicó un formulario de identificación de habilidades tecnológicas a 10 estudiantes al azar de dos grupos de grado sexto, noveno y undécimo, para conseguir un grupo focal de 60 estudiantes. En el caso de los docentes, la encuesta fue aplicada a 5 profesores que fueron seleccionados teniendo en cuenta que dictaran un área fundamental en los grados pertenecientes a la muestra. El formulario utilizado es una escala de calificación basado en el estándar para las competencias del pensamiento computacional e integración tecnológica de la Sociedad Internacional de Tecnología en Educación conocido por sus siglas en inglés ISTE [15] "Los estándares ISTE son lineamientos que indican qué habilidades y competencias necesita desarrollar el estudiante para interactuar en entornos virtuales" [16]. En el formulario se plantearon 35 preguntas agrupadas en 5 categorías y se consiguieron los siguientes resultados al aplicarlo a los estudiantes (Uso básico de recursos TIC 69%, Acceso a recursos TIC 58%, Aplicación de recursos TIC en el proceso de enseñanza-aprendizaje 62%, Calidad de uso de los recursos 90%, Integración y mediación 56.16%) Aunque es una debilidad que el colegio no pueda proporcionar recursos tecnológicos de forma integral a los estudiantes, ellos desean compartir y aportar a la investigación. Selección del modelo de integración tecnológica educativa De acuerdo con las revisiones del marco teórico y la información obtenida en la fase de diagnóstico se escoge el modelo SAMR como una forma de integración tecnológica adaptable a las habilidades de los estudiantes y docentes. Las ventajas de este modelo con respecto a los comparados por la revisión documental, es que permite iniciar su implementación con poco dominio de herramientas digitales y permite el crecimiento paulatino para incluir experiencias de innovación educativa mediadas por TIC. 4.4 Aplicación de estrategias didácticas en contexto Al llegar a esta fase, se propone la integración tecnológica a través de un meso diseño de aprendizaje basado en el modelo SAMR y se evalúa empleando un cuestionario de satisfacción y/o adaptación a las estrategias didácticas aplicadas, con 4 categorías a evaluar: Usabilidad, experiencia de usuario, resultados de aprendizaje esperados, calidad percibida. El cuestionario será diligenciado por el grupo focal.

12.5 Resultados

12.5.1 Propuesta de integración digital basada en el modelo SAMR

El docente actúa como mediador, a través de los entornos virtuales de aprendizaje y por medio de la guía de auto instrucción. Entre más alto sea el nivel de uso y apropiación de los recursos tecnológicos por parte del docente, mayor será la posibilidad de pasar de una etapa a otra del modelo SAMR. El padre de familia como apoyo, el estudiante se convierte en el beneficiario, que hará uso de los contenidos que el docente le proporciona, y a medida que va desarrollando habilidades, irá también ascendiendo en la etapa del modelo SAMR. En el caso de los directivos, su contribución radica en el apoyo a este modelo a través de normas, pautas y capacitaciones que fortalezcan la difusión y aplicación del modelo en la Institución Educativa, de tal manera si ocurre una eventualidad, la comunidad educativa esté lista para tomar acción. La propuesta inició con una formación básica a los maestros en habilidades básicas como el uso de los televisores y proyectores. Estos dos recursos son con los que cuentan la mayoría de salones de clase en la Institución Educativa Supía. Los docentes participantes muestran conocimientos básicos para crear grupos de WhatsApp con sus estudiantes y compartir contenido multimedia a través del chat grupal. Para esto se llevó a cabo el desarrollo de la prueba piloto tomando como referencia el área de filosofía y artística con el apoyo de dos docentes cuyo nivel de habilidades digitales fue recolectado previamente. Diseño y estructuración Resultado de aprendizaje esperado (RAE). Se plantean los resultados esperados

de aprendizaje para el estudiante, ya sea por tema o por periodo académico. Sistema de progreso del aprendizaje (SPA). En esta unidad de aprendizaje sobre se contempla el desarrollo de clase de acuerdo con la etapa que se encuentre de transformación digital. El estudiante sigue de forma secuencial la competencia cognitiva, procedimental y actitudinal:

Cognitivo: El estudiante desarrolla el indicador de logros asignado. El alumnado usa herramientas a su alcance para describir opiniones y situaciones expositivas.
Procedimental: Es válido para el estudiante indagar por diferentes medios para aplicar su conocimiento y habilidad. A partir de recursos multimedia, crea espacios de opinión a partir de herramientas colaborativas o mensajería instantánea promoviendo la argumentación y selección de contenidos educativos pertinentes.
Actitudinal: impacto del interés del estudiante en el tema y uso mediado de las TIC.

A través del desarrollo de los saberes. El alumno crea sus propios recursos educativos digitales y los aprovecha generando espacios críticos y de realimentación.
Evaluación mediada por TIC y aseguramiento del aprendizaje. Para esto se recurre a las plataformas educativas e incluso se puede realizar la evaluación mediada por TIC con apoyo del modelo SAMR, y rúbricas claras, en este caso por la plataforma de calificaciones de la institución.
Crear un diseño de aprendizaje basado en el modelo SAMR. En este caso se observa el ejemplo de un diseño de aprendizaje aplicado al área de filosofía con el modelo SAMR. Este diseño fue creado para la actividad A (ver Figura 12.1) por medio de la herramienta Learning designer disponible en https://www.ucl.ac.uk/learning-designer/:

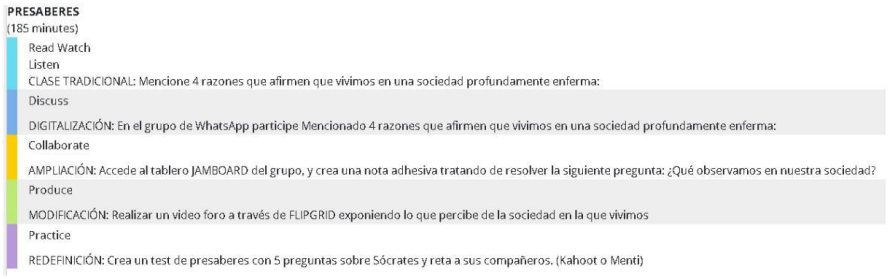


Fig. 12.1: Actividad A adaptada de la guía al modelo SAMR.

Compromiso de los directivos. Para la mejora continua, se hace necesario el apoyo con recursos disponibles y alianzas estratégicas.

12.5.2 Ejecución de prueba piloto 01: Clase de filosofía grado undécimo

Se realizó una prueba piloto con el grado undécimo. Un grupo de 38 estudiantes y la docente. Los estudiantes tuvieron acceso a una guía física y un grupo de WhatsApp. Se dio el inicio protocolario (llamado a lista, preparación de los materiales). El momento A se realizó por WhatsApp. Para el momento B no funcionó el televisor, así que se compartió el video del momento B en el grupo de WhatsApp. No hubo dudas durante la ejecución. Dado que los estudiantes mostraron tener buenas habilidades y acceso a conectividad en ese momento, se subió el nivel de la planeación de AMPLIAR a MODIFICAR, así que, con el momento C y D los estudiantes crearon sus propios contenidos relacionados con la temática por medio de la herramienta digital CANVA. La docente complementó

el tema para los estudiantes que quedaron con dudas del momento B. La clase planeada para 1 bloque de 55 minutos se extendió a 2 bloques. A través del formulario de satisfacción creado con Formularios de Google y basado en los estándares del ISTE, se evaluó la experiencia de la clase usando el modelo de trabajo propuesto, dando como resultado una favorabilidad promedio en el aspecto de Usabilidad del 92,85%, en la Relación interfaz/experiencia de usuario del 90,17%, Resultados de aprendizaje esperados de 92,26% y en Percepción de calidad 87,5%. Esto demuestra un alto nivel de aceptación y participación por parte de los estudiantes al proceso desarrollado y que es necesario el mejoramiento continuo. Se sintieron con mayor disposición a estudiar los temas de clase e incluso permitió la construcción conjunta de material educativo, por ellos y para ellos. La docente percibió que el proceso facilitó la atención personalizada. En el momento A se fomentó la participación, en el momento B la concentración, en C y D la creatividad para utilizar herramientas que facilitan el quehacer académico.

12.6 Conclusiones

Docentes y estudiantes de la I.E Supía encuentran favorable la propuesta de implementación de un modelo de integración tecnológica para la articulación de los procesos de enseñanza-aprendizaje, ya que permite la preparación y fortalecimiento del modelo pedagógico EAU que maneja actualmente la institución. El nivel de habilidades digitales docentes y estudiantil en la Institución es suficiente para llevar a cabo la integración tecnológica, aumentando o reduciendo el nivel de madurez ya que se encuentra basado en el modelo SAMR y esto facilita adaptar las actividades con el estado del aula en el aspecto de acceso a recursos y conectividad. Las estrategias didácticas que contribuyen a la implementación es la planeación basada en micro contenidos en articulación con los momentos de la guía de auto instrucción de la Institución Educativa Supía, lo anterior mediado por WhatsApp. Esto reduce el rechazo al cambio del estudiante ya que continuará aprendiendo bajo los parámetros que ya conoce. Como trabajo futuro se propone implementar por completo en la institución para analizar los cambios en el nivel de atención y desempeño disciplinario de los estudiantes que se acogen al modelo de integración tecnológica.

Referencias

1. Secretaría de Educación de Caldas. SEDCALDAS, https://educacion.caldas.gov.co/documentos/formatos/11-dependencias/19- planeacion/4222-boletin-estadistico-2019, último acceso 2022/01/14
2. Orozco, H. F., https://www.eje21.com.co/2020/05/caldasno-habla-de-educacion-virtualsino-a-distancia/, último acceso 2022/05/20
3. Junta de Andalucía, Orden de 29 de marzo de 2021, por la que se establecen los marcos de la Competencia Digital en el sistema educativo no universitario de la Comunidad Autónoma de Andalucía , 150 (2021)
4. Samperio Pacheco, V. M., & Barragán López, J. F, Análisis de la percepción de docentes, usuarios de una plataforma educativa a través de los modelos TPACK, SAMR y TAM3 en una institución de educación superior. Apertura, 10(1), 116–131, https://doi.org/10.32870/Ap.v10n1.1162, (2018a,b)
5. Brunetti Cani, J.: Proficiência digital de professores: competências necessárias para ensinar no século XXI. En: Linguagem e Ensino, 23(2), 402–428, (2020)
6. López Andrade, M. J., Sancán Pérez, E. E., & Kirenia, M. Z, Aprendizaje virtual, brecha tecnológica en la educación básica ante el COVID-19. UNESUM-Ciencias. Re-

vista Científica Multidisciplinaria. ISSN 2602-8166, 6(4), https://doi.org/10.47230/unesum-ciencias.v6.n4.2022.421, (2022)
7. Iuri, T,: Educación Secundaria Rural Virtual en Río Negro Argentina: una experiencia de más de diez años, Praxis Educativa, 26(3), 1–22, https://doi.org/10.19137/ praxiseducativa2022-260305, (2022)
8. Colombia. Ministerio de Educación Nacional, Oficina de Innovación Educativa con Uso de Nuevas Tecnologías. Competencias TIC para el desarrollo profesional docente, (2013)
9. Mishra, P., & Koehler, M. J,: Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. Teachers College Record, 108(6), 1017–1054, (2006)
10. Rodríguez Jerez, S. A,: El modelo TPACK como perspectiva de análisis en la integración de TIC para la educación: un estado del arte. En: Molina Bernal, I. A., Morales Piñero, J. C. y Rodríguez Jerez, S. A. (coord.), Importancia de las TIC en los procesos de enseñanzaaprendizaje: estudios en la educación media y superior, pp. 11- 34, https://repository.usergioarboleda.edu.co/bitstream/handle/11232/1512/El%20modelo%20TPACK%20como%20perspectiva%20de%20an%c3%a1lisis.pdf?sequence=4&isAllowed=y, (2019)
11. Puentedura, R, Transformation, Technology, and Education. http://hippasus.com/resources/tte/, (2006), último acceso: 2022/03/06
12. Commonwealth of Learning COL, Introduction to Microlearning Course. Canadá. https://oasis.col.org/colserver/api/core/bitstreams/07dd80b84-b502-4ed4-8f9f-1504d4613084/content, (2021)
13. Bell, J.: Cómo hacer tu primer trabajo de investigación. (Roc Filella Escolá, trad.), México: Gedisa, (2005)
14. Hernández, R., Fernández, C., Baptista, P.: Metodología de la investigación. México: Mc Graw Hill Educación, (2014).
15. International Society for Technology in Education. ISTE Standars, https://www.iste.org/standards, último accesso: 2022/11/20
16. Muralles, M. R.: Estándares ISTE: integración entre tecnología, educación y contexto. En: Proceedings of the Digital World Learning Conference CIEV 2019, http://biblioteca.galileo.edu/tesario/bitstream/123456789/953/1/5.pdf (2019)

13
Gestión del conocimiento (SKMS) en áreas de tecnología basado en ITIL V4

Oscar Gonzalez[1][0009-0009-2994-1617]
Marcelo Lopez[2][0000-0003-0668-1292]

Resumen La gestión del conocimiento incluye procesos de recopilación, análisis, almacenamiento y compartir datos e información que se conviertan en un pilar fundamental para transferir conocimiento que aporte valor al área de TI, y no solo en la mesa de ayuda, sino en todo el ciclo de los servicios y en las prácticas de gestión general, de servicio y de gestión técnica. La implementación de un sistema de gestiona del conocimiento basado en prácticas y recomendaciones de ITIL requiere de esfuerzos, pero trae muchos beneficios en cuanto a la mejora en la prestación de los servicios y ahorros en la optimización de los recursos tecnológicos de una organización.

Palabras clave Gestión de conocimiento, ITIL, transformación digital.

13.1 Introducción

La gestión del conocimiento basada en ITIL trae grandes beneficios, pero a su vez, enormes retos incluidos en su misma definición: crear, usar, compartir y mantener; como usar la información y que no se quede almacenada en repositorios que por muy bien diseñados y que hagan uso de tecnología de punta, no hayan sido incorporados a los procesos y a las personas; como generar el interés por compartir información que alimente no solo el proceso de tecnologías de información, sino que pueda ser aprovechada por otros procesos al interior de la organización. Existen una gran oferta de herramientas de software de apoyo a gestión de áreas de TI, generalmente de mesa de ayuda o mesa de servicio que incluyen componentes para gestionar el conocimiento, un ejemplo es Jira de la casa de software Atlassian de Australia, y su modulo Confluence, definido como un repositorio de información donde del conocimiento y la colaboración se encuentran, ofrece además de un repositorio documental una funcionalidad de páginas dinámicas en donde los integrantes de los equipos pueden generar espacios compartidos para las ideas y conceptos. Otro ejemplo es GLPI un software de gestión de servicios de código abierto, usada por varias entidades del orden nacional. Contiene funcionalidades que podrían ser usadas para una gestión del conocimiento, pero su usabilidad y alcance puede no ser suficiente para un proyecto como el objeto de este artículo.

[1] Aguas de Manizales SA ESP BIC, Caldas-Manizales, Colombia
e-mail: ing.ogonzalez@gmail.com
[2] Departamento de sistemas y computación, facultad de ingeniería, Universidad de Caldas, Caldas-Manizales, Colombia
e-mail: mlopez@ucaldas.edu.co

Estructurar el conocimiento en áreas de tecnologías de información según el marco ITIL V4, tiene dos componentes base, el primero es la gestión del conocimiento y el segundo el marco ITIL del que hablaremos más adelante. Este artículo realiza una propuesta basada en años de experiencia en diferentes organizaciones, sobre cómo implementar una base de conocimiento que el equipo de TI pueda aprovechar.

13.1.1 El problema

La tecnología y el conocimiento se han convertido en pilares fundamentales del direccionamiento estratégico y evolución de las organizaciones; en este contexto las áreas de tecnología emergen como ejecutores de habilitadores tecnológicos que impulsan este desarrollo y esta transformación, sin embargo con la cada vez más rápida evolución de las tecnologías, los responsables y equipos de TI deben buscar herramientas que faciliten sus actividades, y allí es donde aparece en un lugar importante la gestión del conocimiento tanto de la operación, como de los servicios y las herramientas.

13.1.2 Objetivo

Determinar las bases para el diseño e implementación de una herramienta de gestión del conocimiento en áreas de tecnologías de información basado en principios y prácticas de ITIL.

13.2 Marco teórico

13.2.1 Gestión del conocimiento

Para [1], el conocimiento es una mezcla de flujo de experiencias enmarcadas, valores, información contextual que proporciona un marco para valorar e incorporar nuevas experiencias e información. Se origina y se aplica en la mente de los conocedores. En las organizaciones, frecuentemente se manifiesta no sólo en los documentos, sino también en procesos, prácticas y normas. En tanto en [2], se define la gestión del conocimiento en organizaciones, como la identificación y el aprovechamiento del conocimiento colectivo para ayudar a la organización a competir. (Barragán O., 2009), señala la importancia del desarrollo de la gestión del conocimiento en el marco de una sociedad del conocimiento, integrando los actores principales y considerando los beneficios que debe recibir el individuo, como parte de una comunidad intrincada en diferentes redes que deben participar e interactuar entre sí para el bienestar común. “Según la definición de gestión del conocimiento, se trata del proceso de crear, utilizar, compartir y mantener el conocimiento de la organización para que los usuarios de la empresa puedan utilizarlo siempre que lo necesiten”. Las bondades y potencialidad de la gestión del conocimiento son variados, partiendo desde generar cambios y resultados sustentables, optimizar recursos, aprovechar el conocimiento existente, aprender continuamente, estimular la creatividad e innovación, como lo plantean diferentes investigadores en la materia, que proponen la combinación e intercambio de conocimientos en las empresas y organizaciones, con: disminución de costos, creatividad e innovación de productos, el mejoramiento organizacional, el aumento del rendimiento y de los ingresos por ventas. En este artículo, se propone un concepto de GC alineado a ITIL 4 centrado en la creación y entrega de valor, [5]. Este enfoque se basa

en los siguientes elementos clave: Creación de conocimiento: Capturar, almacenar y organizar el conocimiento de manera efectiva para que sea accesible y utilizable para todos los interesados. Compartir conocimiento: Facilitar el intercambio de conocimiento entre individuos, equipos y departamentos para fomentar la colaboración y la innovación. Aplicación del conocimiento: Utilizar el conocimiento para mejorar la toma de decisiones, resolver problemas, optimizar procesos y crear valor para los clientes y la organización. Medición del conocimiento: Evaluar el impacto del conocimiento en la organización y realizar mejoras continuas en los procesos de GC. La gestión del conocimiento para áreas de TI debe aprovechar los recursos tecnológicos, metodologías y marco ITIL para integrar la gestión del conocimiento con la gestión de servicios de TI., reconocer el valor del conocimiento de los empleados y animarlos a compartirlo, crear un ambiente que impulse la innovación y use herramientas de conocimiento para capturar y compartir, identificar, compartir y usar el conocimiento sobre riesgos para prevenirlos o reducirlos, [7],[8].

13.2.2 Metodologías para la implementación de sistemas de gestión del conocimiento

Existen diferentes metodologías para la implementación de sistemas de gestión del conocimiento (SGC) las cuales pueden adaptarse a cada caso específico u organización, según las necesidades y el contexto organizacional. A continuación, se mencionan algunas de las metodologías referencia para la implementación del caso de estudio. Cabe mencionar que las metodologías pueden adaptarse y combinarse según las necesidades del sistema de gestión del conocimiento que se quiera implementar. Modelo de Implementación de Gestión del Conocimiento (KM Implementation Model): Davenport y Prusak [1], este modelo se enfoca en identificar el conocimiento crítico, desarrollar sistemas de gestión del conocimiento y promover una cultura que apoye la creación, el intercambio y la utilización del conocimiento. Ciclo de Vida del Conocimiento (Knowledge Life Cycle): Este enfoque describe las etapas que atraviesa el conocimiento dentro de una organización, desde la creación y captura hasta la aplicación y la renovación. SECI Model (Socialization, Externalization, Combination, Internalization): Nonaka y Takeuchi, muestra cómo el conocimiento tácito, que es lo no dicho se convierte en conocimiento explícito y viceversa a través de cuatro modos de interacción: socialización (compartir experiencias), externalización (conceptualizar conocimiento), combinación (sistematizar conocimiento) e internalización (aplicar conocimiento).

13.2.3 El proceso cognitivo

Cognición, se define como la capacidad que permite crear y desarrollar conocimientos; una habilidad para relacionar y procesar datos, evaluando y sistematizando la información generada a partir de la percepción y la experiencia. Los procesos cognitivos son los componentes que permiten que procesemos la información que nos llega a través de los sentidos, que la almacenemos, manipulemos, la recuperemos e interactuemos con el mundo y que aprendamos, sobre todo que aprendamos. Las funciones cognitivas (término que habitualmente se usa de forma indistinta) son la base de nuestro conocimiento e incluye cosas tan básicas como la percepción y atención, y tan avanzadas como el pensamiento. Principales procesos cognitivos: Proceso 1 - Percepción: Captura de información través de los sentidos (visión, olfato, oído), dar significado a las señales o sensaciones que nos llegan por los órganos de los sentidos. Proceso 2 - Memoria: Almacenamiento y recuperación de los datos o información cuando se requieran. Proceso 3 - Pensamiento: Relacionar, procesar la información, generar ideas, solucionar problemas, tomar decisiones, es una forma activa de procesar la información. Proceso 4 - Lenguaje: Comunicación

13.2.4 El aprendizaje y la enseñanza

La implementación, la clasificación y calidad de los datos son muy importantes, pero la idea de aprendizaje significativo de Ausubel, *el conocimiento verdadero solo puede nacer cuando los nuevos contenidos tienen un significado a la luz de los conocimientos que ya se tienen*, le da un complemento valioso y un enfoque holístico a todo el ecosistema que rodea una herramienta de gestión del conocimiento y ofrece posibilidades de un mejor aprovechamiento y un enfoque más profundo para la organización, y quiere decir, que aprender significa que los nuevos aprendizajes conectan con los anteriores; no porque sean lo mismo, sino porque tienen que ver con estos de un modo que se crea un nuevo significado, generando así nuevo conocimiento y/o mejor apropiación del existente. El autor Delors[4], en su libro La Educación encierra un tesoro, indica cómo la educación debe ser concebida desde la integralidad de la persona, para ello, se basa en cuatro pilares fundamentales: aprender a conocer, aprender a hacer, aprender a vivir juntos, aprender a ser. Bajo este contexto en el diseño de un sistema de gestión de conocimiento para áreas de TI se deben tener en cuenta todos los componentes incluyendo a los individuos como elemento principal que además de aportar o consultar, deben ser nutridos con conocimiento que permita el aprovechamiento de las bases creadas. Las organizaciones deben buscar sistemas efectivos de transmisión de información en donde los individuos no solo repitan la información que se les da, sino que estén en capacidad de generar nuevo conocimiento a partir de los recursos dispuestos mejorando periódicamente el contenido existente.

13.2.5 ITIL

ITIL (Information Technology Infrastructure Library) es un conjunto de conceptos y buenas prácticas empleados en empresas alrededor del mundo para la gestión de servicios de TI (Tecnologías de información), que está orientada hacia la mejora de la calidad de los servicios prestados apoyando el logro de los objetivos empresariales. ITIL 4, el marco de referencia líder para la gestión de servicios de TI (ITSM), ofrece ventajas cruciales para las organizaciones: Enfoque en el valor, flexibilidad y adaptabilidad, Mejora continua, colaboración y trabajo en equipo, gestión de riesgos, experiencia del cliente, Adopción de tecnologías modernas [6].

13.3 Gestión del conocimiento en ITIL

ITIL v4 no describe procesos, describe prácticas. Axelos, la empresa encargada de mantener ITIL define una práctica de gestión como un conjunto de recursos organizacionales diseñados para realizar un trabajo o cumplir con un objetivo. Existen tres tipos de prácticas asociadas con la prestación de servicios de TI, 14 prácticas de la gestión general, 17 prácticas de gestión del servicio y 3 prácticas de gestión técnica. La gestión del conocimiento se encuentra dentro de las prácticas de gestión general y está directamente relacionada con las demás prácticas, de las cuales recibe entradas. Esta relación directa genera una sinergia entre las prácticas que permite que el conocimiento se integre para resolver problemas de maneja más eficiente. ITIL menciona en sus referencias varios componentes de gestión de conocimiento: Soporte basado en el conocimiento (KCS), conjunto de mejores prácticas para la gestión del conocimiento que recomienda la mejor forma de estructurar el conocimiento para la gestión y resolución de incidentes, buscando mejorar los SLA definidos. Los sistemas de gestión de la configuración (CMS). Contiene todos los elementos de configuración (CI) dentro de la mesa de servicio. Base de datos de gestión de la configuración (CMDB), que contiene todos los datos de configuración relacionados con la infraestructura de TI. Almacena información

sobre los componentes de hardware y software que se usan en la organización, identificando las relaciones entre ellos. Puede incluir dependencias entre los elementos (CI) y el historial de cambios. ITIL 4 define estos elementos (CI) como “todo componente que debe gestionarse para ofrecer un servicio de TI”.

13.3.1 Sistema de gestión de conocimiento del servicio (SKMS)

Se trata de un principio de la práctica de gestión del conocimiento, centraliza la información sobre los servicios definidos por el marco para una organización, busca principalmente reducir los incidentes, problemas y aumentar el tiempo de disponibilidad de los servicios. Está basado en una gran variedad de componentes relacionados con los grupos de prácticas del marco como: gestión del conocimiento en la gestión de incidentes y problemas, gestión del conocimiento en la gestión de eventos, gestión del conocimiento en la gestión de la configuración del servicio, gestión del conocimiento en la gestión de cambios y gestión del conocimiento para la mejora continua y en fuentes de datos como biblioteca de software, datos del portafolio de hardware, gestión de la configuración, inventario de servicios y productos, información de incidentes, problemas y cambios, solitudes de servicios que incluyen las soluciones registradas ante incidentes, análisis de problemas y solicitudes de cambios.

13.4 Base de conocimiento

Una base de conocimiento es una agrupación de datos generalmente disponible para los colaboradores a manera de autoservicio y en línea con información sobre los servicios, los productos, procesos, áreas u otros temas específicos. Las bases de conocimientos son un pilar básico de la práctica de gestión de conocimiento (Prácticas de gestión general), que facilitan crear, seleccionar, compartir, utilizar y gestionar el conocimiento del área de TI con principios de planeación y mejora continua. Aunque la practica da recomendaciones en cuanto a planeación y uso, el conocimiento a manera de datos puede provenir de cualquier persona, en cualquier formato, y son las personas con un mayor conocimiento de los temas, quienes se encargaran de hacer aportes a la base. Una base de conocimiento puede incluir banco de preguntas frecuentes, manuales, guías de solución de problemas, libros de ejecución, y cualquier dato que la organización o un área específica pueda necesitar o considere relevante. Aquí el área de TI se empieza a diferenciar del resto al dar un mayor uso de información como: documentación de los servicios, runbook automatizados y manuales técnicos de los productos. Hoy en día muchas bases de conocimiento están siendo implementadas en torno a la inteligencia artificial, facilitando la interacción entre el usuario y la información respondiendo las inquietudes de manera más ágil y asertiva haciendo uso de un razonamiento deductivo automatizado, que significa que cuando un usuario realiza una interacción la herramienta logra direccionarlo en menos pasos y por ende mejor tiempo, logrando un acceso más directo a la respuesta buscada; sea cual sea la implementación, la conformación y calidad de la información es pieza clave y un pilar fundamental en la construcción de la base de conocimiento.

13.5 Resultados

El primer paso para crear un SKMS es determinar qué información agregaría valor al ejercicio y daría los mejores beneficios, en esa misma línea, se deben identificar el alcance y el propósito

de la implementación, para este caso un sistema de gestión del conocimiento del servicio basado en ITIL para área de tecnologías de la información útil para gestionar de manera integral los servicios de TI. Como vimos anteriormente existen 4 procesos cognitivos principales, aunque algunos autores adicionan el proceso de la atención como un quinto componente, que para el caso de la implementación podría servir como un filtro que permite procesar solo la información que este directamente relacionada con el caso de estudio mientras se ignora información no relevante para los resultados esperados. La implementación de un sistema de gestión del conocimiento puede tener diferentes niveles de madurez que dependen de factores como el alcance del proyecto, la madurez de la organización para identificar y gestionar su conocimiento en las personas y procesos, y el uso de habilitadores tecnológicos como infraestructura, capacidad de procesamiento y uso de tecnologías de aprendizaje supervisado, ni supervisado, automatización e inteligencia artificial. La gestión del conocimiento se puede dividir en los 4 procesos principales.

Table 13.1: Procesos cognitivos, de gestión del conocimiento y SKMS

Proceso / Grupo	**Propósito**	**SKMS**	**Herramientas**
Percepción	Definir mecanismos de adquisición de datos Identificar datos relevantes		Mesa de servicio Correo electrónico
Memoria	Bases de datos Estructuras de almacenamiento Criterios de clasificación Medios de acceso a la información	Capa de datos e información, Capa de integración de la información	Tecnologías de nube Software como servicio Bases de datos documentales Mesa de servicio CMBD
Pensamiento	Identificar relaciones según clasificación Definir uso de algoritmos de procesamiento Identificación de insights	Procesamiento del conocimiento	Inteligencia artificial (IA) programación tradicional
Lenguaje	Definir niveles de acceso Definir medios de consulta	Capa de presentación	Mesa de servicio

Si bien la práctica de gestión del conocimiento no menciona el equivalente del proceso cognitivo de percepción, autores como Lopez [3] en su libro Innokit, menciona gran variedad de técnicas y herramientas para gestionar el conocimiento, en ellas incluye métodos para identificar, seleccionar y tratar la información de necesita una base de gestión de conocimiento. Una vez identificadas las bases del sistema de gestión del conocimiento, se continua con el componente tecnológico en el cual se pretendía materializar el ejercicio de investigación realizado; para ello se hace uso de tecnologías pertenecientes al ecosistema Microsoft, como lo son: Microsoft SharePoint como plataforma de colaboración y gestión de contenido que facilita el trabajo en equipo mediante la creación de sitios de trabajo compartidos y Microsoft Copilot (Microsoft Copilot Studio) usado para extraer información e interactuar con los usuarios, obteniendo los mejores resultados, en el menor tiempo posible a partir de la base de conocimiento construida en el ejercicio. Integrar estas herramientas ofreceventajas significativas en términos de colaboración y acceso a la información. El uso de la IA se aplica de manera transversal en los procesos de un modelo cognitivo, así como de diferentes capas del modelo de SKMS; sus aplicaciones son amplias, desde obtener información hasta comunicarla en lenguaje natural a los usuarios, ya sean del equipo de tecnología o clientes internos en búsqueda de respuestas. Es así como durante el proceso de implementación se pudo identificar un beneficio adicional con la creación de un BOT de soporte, el cual desde Microsoft Teams puede acceder a repositorios de la base de conocimiento y a manera de auto gestión

interactuar con los usuarios finales, evitando que se ingresen solicitudes de servicio. Los resultados esperados no buscan solo mejorar los SLA mediante la solución de incidentes y disponibilidad de los servicios, busca mediante una visión holística incluir información sobre servicios, productos, infraestructura, configuración y en general los activos de TI, que juntos puedan generar tanto una memoria institucional como nuevo conocimiento que mejore la operación y contribuya a procesos de transformación digital en las organizaciones.

13.6 Conclusiones

Adoptar un enfoque centrado en el valor permite: alineación de las iniciativas de GC con los objetivos estratégicos del negocio, demostración del impacto del conocimiento en la mejora de la eficiencia, la eficacia y la creación de valor, priorización de las actividades de GC en función de su contribución al logro de los objetivos del negocio. Se debe tener cuidado de: no alineación entre las actividades de GC y los objetivos del negocio, incapacidad para demostrar el valor del conocimiento en términos cuantitativos, evitar el enfoque en actividades de GC que no generan un impacto tangible en el negocio. Una herramienta tecnológica para gestionar el conocimiento en áreas de TI no es suficiente por si sola, debe estar acompañada de buenas prácticas y del compromiso de toda la organización, principalmente del equipo de TI que son quienes serían sus principales beneficiarios, pero de fondo toda la empresa se vería beneficiada por que además de los beneficios como optimización de procesos y en conservación del conocimiento que pueden obtenerse al inicio, su aplicación puede hacerse en otras áreas logrando así un repositorio de conocimiento a nivel organizacional. Fomentar una cultura de colaboración y aprendizaje permite a las organizaciones: creación de un entorno donde el intercambio de conocimiento es valorado y recompensado, fomento del trabajo en equipo y la colaboración entre diferentes áreas de la organización, promoción de una cultura de aprendizaje continuo y desarrollo profesional. Se debe considerar como enfrentar la persistencia de silos de información y resistencia al intercambio de conocimiento, falta de colaboración entre equipos y departamentos, cultura organizacional que no incentiva el aprendizaje y el desarrollo profesional. Un sistema de gestión del conocimiento además de servir como repositorio de datos debe promover la creación de nuevo conocimiento, y en ese mismo orden debe ser estructurado y mantenido para que evolucione con la organización. Aprovechar el marco de ITIL 4 para alinear la gestión del conocimiento con la gestión de servicios de TI posibilita: integración de las prácticas de GC con los principios y directrices de ITIL 4, alineación de los objetivos de GC con los objetivos de gestión de servicios de TI, utilización de ITIL 4 para crear una base sólida para la gestión del conocimiento en el área de TI. Ademas se debe considerar para evitar fracasos: la falta de integración entre las prácticas de GC y el marco de ITIL 4, desalineación entre los objetivos de GC y los objetivos de gestión de servicios de TI e incapacidad para aprovechar ITIL 4 para fortalecer la gestión del conocimiento en el área de TI. Si bien en la implementación relacionada con este artículo se usaron tecnologías Microsoft, las cuales tienen un costo asociado pero que son cada día más comunes en las empresas el mundo entero y en especial de Colombia; teniendo claras las bases que identificamos desde el principio (Gestión del conocimiento e Itil), se tienen alternativas validas, como hacer uso de recursos de código abierto o desarrollar los componentes. Hoy en día, Microsoft ofrece herramientas sumamente interesantes con una proyección prometedora en la mejora y adopción de tecnologías disruptivas como la Inteligencia Artificial. Lo destacable es que todas estas herramientas están integradas en un ecosistema tecnológico cada vez más fácil de usar.

Referencias

1. Davenport, T. H.; Prusak, L. (1998). Working Knowledge. Harvard University Press, (5) 2
2. Alavy, M.,; Leidner, D. (2001). Review knowledge management and knowledge management sytems conceptual foundatios and research issues. MIS Quarterly, 107-136
3. Lopez et al. (2019). Innokit, Kit de técnicas y herramientas para gestionar el conocimiento y la innovación. Universidad Nacional de Colombia
4. Delors, J. (1996). La Educación encierra un tesoro. Santillana. Ediciones Unesco
5. Khilji, N. y Nicolic, K. (marzo de 2024). La influencia de la gestión del conocimiento en la transformación digital: una visión general para la gestión del cambio y la innovación. En Conferencia sobre el futuro de la información y la comunicación (págs. 368-388). Cham: Springer Nature Suiza.
6. Junior, L. C. G., de Barros, R. M., & de Oliveira Barros, V. T. (2024, January). A model for implementing the itil 4 event management and monitoring practice in a public organization. in 20th contecsi-international conference on information systems and technology management virtual.
7. Dias, MT, Marcelo, M., Gauthier, FAO, dos Santos, GL, Bondan, G., Izidorio, G., ... & Botelho, LDLR (2024). Análisis de prácticas de gestión del conocimiento para la autogestión de organizaciones intensivas en conocimiento. Revista Abierta de Ciencias Sociales, 12 (6), 1-23.
8. Ferenhof, HA, Bonamigo, A., Rosa, LG y Vieira, TC (2024). Marco teórico de gestión del conocimiento B2B centrado en la cocreación de valor. Revista VINE de sistemas de gestión de la información y el conocimiento, 54 (2), 424-451.

14
Diseño e implementación de sistema de bajo costo para el monitoreo ambiental

Aldemir Vargas Eudor[1] [0000-0002-6656-9422]1
Juan E. Graell Alzate [2] [0009-0009-5936-5112]
Juan M. Rojas Rojas[3] [0009-0004-2113-7556]
Juan C. Escobar Naranjo[4] [0009-0008-3262-1668]
Karol Pavas Leiva[5] [0009-0006-2751-2817]

Resumen El monitoreo ambiental ha adquirido una relevancia cada vez mayor debido a la importancia en la comprensión de los fenómenos climatológicos globales. Los sistemas de monitoreo ambiental, tradicionalmente limitados por sus elevados costes, pueden implantarse ahora de forma más asequible gracias a tecnologías emergentes como IoT. Este enfoque permite la creación de amplias redes de monitorización medioambiental utilizando sensores y plataformas tecnológicas de bajo costo. En el presente trabajo se desarrolla un sistema de monitoreo ambiental de bajo costo basado en NodeMCU V3 ESP8266, que utiliza sensores para medir variables meteorológicas como temperatura, humedad y presión, transmitiendo los datos vía WiFi a un servidor local. El sistema se plantea con una arquitectura modular de microservicios, que facilita su mantenimiento y escalabilidad, permitiendo realizar ajustes y añadir módulos según sea necesario. La integración de microservicios mediante tecnologías como Docker y Kubernetes mejora la eficiencia y disponibilidad del sistema, gestionando el flujo de datos desde múltiples estaciones de monitorización a una base de datos NoSQL como MongoDB. Esta infraestructura no solo optimiza la gestión de datos en entornos IoT dinámicos, sino que también garantiza una respuesta sólida y escalable a las variaciones de la carga de trabajo.

Palabras clave Monitoreo ambiental, datos climatológicos, SGA, ETL, microservicios.

14.1 Introducción

Los sistemas de monitoreo ambiental han adquirido una importancia creciente a lo largo de los años, siendo fundamentales para comprender los fenómenos climatológicos a nivel mundial [1]. Esta relevancia radica en la generación de una mayor cantidad de datos e información esencial para respaldar la gestión de riesgos y desastres en diferentes territorios. Estos sistemas, comúnmente compuestos por sensores que recopilan datos ambientales, se integran en plataformas que facilitan la captura, almacenamiento y análisis de información, esta, puede ayudar a cuantificar e identificar los cambios ambientales y así anticipar, diagnosticar y caracterizar el problema y sus fuentes [2]. Sin embargo, la implementación de estos sistemas ha sido tradicionalmente limitada, especialmente

Universidad Nacional de Colombia - Sede
Manizales, Kilómetro 7 vía al Magdalena, Manizales, Caldas, CO, Grupo de ambientes inteligentes adaptativos GAIA.

en países en desarrollo, debido a los altos costos asociados. Además de los recursos financieros, otros obstáculos incluyen la necesidad de modificar procesos de producción e invertir en tecnología avanzada, que a menudo no está disponible localmente. Estas barreras han dificultado la creación de una red de monitoreo ambiental sólida y extensa, limitando su alcance y eficacia, especialmente para las pequeñas y medianas empresas (PYMEs) o centros de investigación públicos y privados en países como Brasil y Colombia, al enfrentar mayores desafíos en la adopción de estos sistemas [3]. Diversos estudios han demostrado la viabilidad de desarrollar sistemas de monitoreo ambiental de bajo costo utilizando tecnologías emergentes como microcontroladores, sensores económicos y plataformas IoT (Internet de las Cosas) [4, 5, 6]. Estos sistemas permiten realizar un monitoreo constante de variables ambientales clave, como temperatura, humedad, calidad del aire y precipitaciones, a través de la implementación de estaciones de monitoreo automatizadas [4]. Al reducir significativamente los costos de implementación, estos sistemas de bajo costo facilitan la creación de redes de monitoreo más extensas y accesibles. Este trabajo propone un sistema de monitoreo ambiental de bajo costo basado en IoT, iniciando desde el diseño de una arquitectura que permite su posterior escalamiento, la construcción y despliegue de estaciones que miden variables climatológicas, y el desarrollo del proceso ETL (Extracción, Transformación y Carga) que completa el sistema.

14.2 Propuesta del sistema

La recolección de datos ambientales tales como temperatura, humedad y presión atmosférica a través de sistemas de bajo costo, brinda una oportunidad al despliegue de una gran red de estaciones con estos tipos de sensores. Sin embargo, conectar una gran cantidad de dispositivos a través de IoT lleva su propio conjunto de desafíos y obstáculos, como gestionar la recolección y almacenamiento de los datos que se recopilan. En este proyecto se plantea medir variables ambientales como temperatura, humedad, y presión y cargar estos valores en un servidor local. La arquitectura propuesta para el sistema se diseñó de manera modular la cual brinda una mayor mantenibilidad al facilitar cambios aislados, una mejor reutilización de componentes, una escalabilidad mejorada al permitir agregar o modificar módulos según sea necesario y una integración facilitada al mejorar la interoperabilidad entre las partes del sistema.

14.3 Módulo IoT

Con el crecimiento de Internet, cada vez tenemos más personas conectadas entre sí. El IoT va más allá de conectar solo a las personas; da un paso adicional al conectar dispositivos electrónicos entre sí. La disminución de los costos en los dispositivos de comunicación ha permitido la interconexión de diversos dispositivos a través de una red, lo que posibilita la recolección de datos de sensores IoT [7], los cuales pueden distribuirse y procesarse en un servidor central o en un despliegue en la nube. Para el núcleo del sistema de captura de mediciones, hemos seleccionado la placa de desarrollo Node MCU V3 ESP8266, que es de bajo costo y contiene un módulo WiFi. La elección de esta placa no se basó únicamente en su precio y conectividad WiFi, sino también en otras características importantes. Entre ellas, se destacan su capacidad de procesamiento suficiente para las tareas de IoT, su facilidad de programación mediante el IDE de Arduino, y un amplio soporte de comunidad, lo que facilita el desarrollo y la resolución de problemas. Además, su bajo consumo de energía y su flexibilidad para ser utilizada en una amplia gama de aplicaciones IoT la convierten en una opción práctica y duradera para proyectos de monitoreo ambiental. La programación se realizó en código C a través del IDE de Arduino y se cargó en el ESP8266 mediante un bus de comunicación serie.

14.3.1 Metodología

Para el desarrollo del módulo IoT, se realizó una revisión de trabajos relacionados con el uso de tecnologías inalámbricas existentes para la captura de datos de sensores meteorológicos [9]. La revisión del estado del arte nos permitió comparar las ventajas y desventajas de cada una de ellas. Dado que existen diferentes tecnologías de comunicación inalámbrica que varían según el alcance de la señal. Como ejemplo tenemos: Bluetooth, Zigbee, Z-Wave, 6loWPAN, Sigfox, NFC, etc. [8], se tomó como referencia la revisión de las tecnologías de comunicación inalámbricas para IoT hecha por Sadek & Elbadawy (Ver Tabla 2) [10].

Table 14.1: Características de las tecnologías inalámbricas IoT.

Tecnología	**Estándar**	**Frecuencia**	**Rango**	**Consumo de energía**	**Costo**
Bluetooth	3GPP Release	2.4GHz	1.5 Km; 100m max for 5G	Bajo	Bajo
loRa	LoRaWan	415MHz, 868MHz,915MHz	Up to 15km	Bajo	Bajo
Wifi	IEEE802.11	2.4GHz, 5Hz	Up to 100m	Medio	Medio
ZigBee	IEEE802.15.4	2.4GHz	100m	Bajo	Medio
NFC	ISO/IEC18000-3	13.56MHz	10cm	-	-
VLC	IEEE802.15.7-2011	400 400 THz to 800THz	20m	Bajo	Medio

El resultado de la revisión nos permitió seleccionar la placa de desarrollo (NodeMCU V3 ESP8266) atendiendo los requerimientos de que el sistema sea de bajo costo, bajo consumo de energía, seguridad de las comunicaciones y facilidad de implementación.

14.3.2 Materiales

La medición de temperatura y humedad se realiza con el sensor DHT11. Este sensor tiene un termistor que sirve para medir el aire circundante (temperatura) e implementa un sensor interno capacitivo para la medición de humedad. Este dispositivo funciona mediante el uso de tres terminales, +Vcc, Gnd y DATA. A través del pin 3 (DATA) se obtiene una señal digital que es tratada en la placa de desarrollo. Para medir la presión barométrica se utiliza el sensor BMP180. Este dispositivo es de bajo consumo de energía, ofrece un rango de medición de 300 a 1100 hPa y una precisión absoluta de hasta 0.02 hPa. Finalmente se tiene el módulo RTC-DS1307 para el registro de hora y fecha de la captura del dato. Este módulo de reloj en tiempo real está basado en tecnología de Dallas Semiconductors, viene totalmente ensamblado y listo para funcionar. Posee una zocalo para una batería tipo botón de litio (CR2032) que provee la alimentación interna sino se dispone de 5V externo. Este módulo presenta una interfaz de comunicación de protocolo I2C.

Table 14.2: Descripciones materiales utilizados en el sistema. Elaboración propia.

Elemento	**Descripción**	**Costo en dólares**
NodeMCU V3 ESP8266	Placa de desarrollo para recepción de datos de los sensores y posterior envío al sistema de almacenamiento.	$4,00
BMP180	Sensor de captura de presión (PSI), altitud y temperatura	$1,00
DHT11	Sensor de captura de humedad y temperatura	$1,19
RTC-DS1307	Reloj para captura de fecha y hora	$1,25
Módulo Transceptor NRF21101 RF de 2.4GHz	Módulo opcional para la comunicación inalámbrica en caso de no disponer de redes WIFI	$1,23

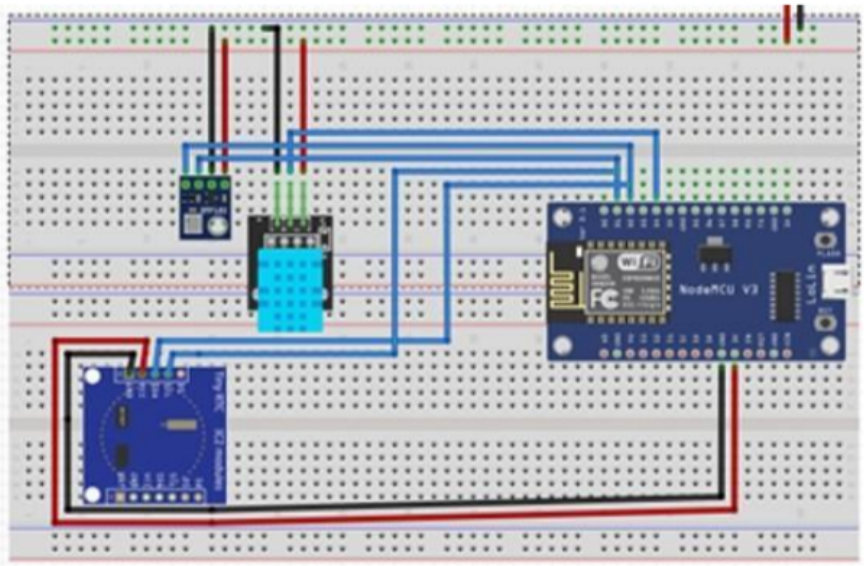


Fig. 14.1: Diagrama de conexiones. Elaboración propia.

14.4 Arquitectura

Con la selección de la placa de desarrollo y los sensores de humedad, temperatura y presión, se tiene un módulo completo de IoT para la captura y transmisión de datos. En este punto, se hace imprescindible implementar una solución tecnológica capaz de recibir, procesar y almacenar estos datos de manera eficiente y escalable. La adopción de arquitecturas basadas en microservicios ha supuesto mejoras significativas en términos de escalabilidad y disponibilidad [11] La naturaleza modular y el tamaño reducido de los microservicios permiten una rápida recuperación ante fallos o actualizaciones, asegurando altos niveles de disponibilidad. Al integrar herramientas de contener-

ización como Docker y Kubernetes, se amplifican los beneficios al permitir una gestión eficiente y escalable de los microservicios, permitiendo así mantener disponible el servicio para la ingesta de los datos producidos por las estaciones de monitoreo [12]. La arquitectura de la aplicación se basa en dos microservicios desarrollados en Go, diseñados para gestionar las solicitudes de las diferentes estaciones mediante colas. Estas, aseguran una gestión eficiente del tráfico de las peticiones, encolando las solicitudes de manera ordenada antes de redirigirlas hacia un microservicio adicional encargado de recibir estos datos desde el microservicio de encolado y escribirlos como registros en una base de datos. Esta estructura no solo asegura una gestión robusta del flujo de datos desde las estaciones hasta la base de datos, sino que también facilita la escalabilidad horizontal del sistema para manejar incrementos en la carga de trabajo de manera eficiente y adaptable a las necesidades del entorno IoT moderno [13]. El primer microservicio, implementado como un

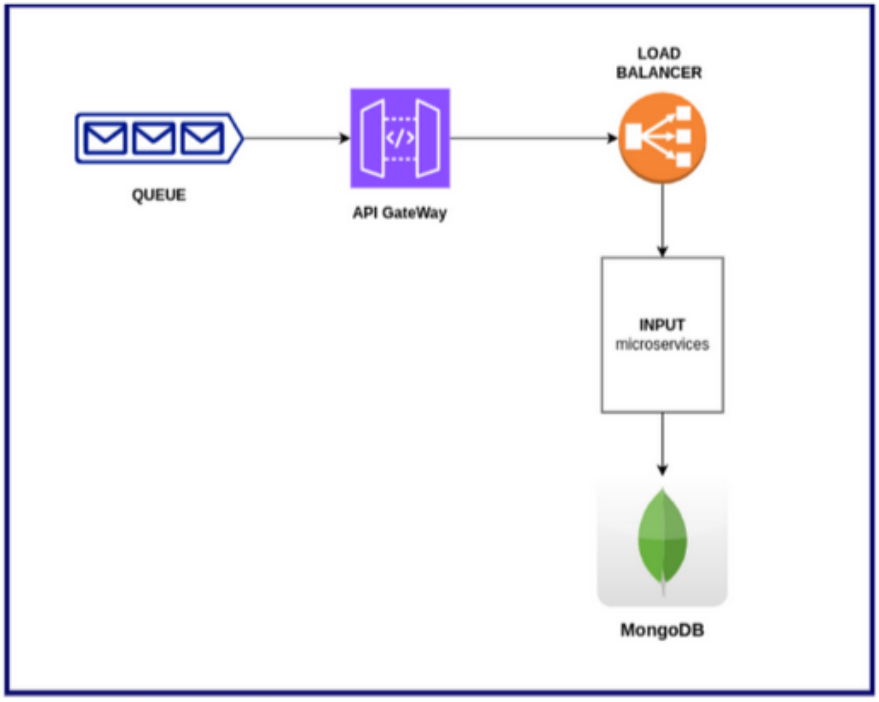


Fig. 14.2: Diagrama de conexiones. Elaboración propia.

servicio web utilizando Go Kit, funciona como el punto de entrada principal. Este componente utiliza una cola eficiente implementada con goroutines y canales de transferencia con buffer, lo cual permite manejar las solicitudes entrantes de múltiples estaciones de manera concurrente y ordenada. La utilización de concurrencia en este microservicio asegura. que el servicio de recepción de datos se libere rápidamente, optimizando el endpoint de entrada para mejorar la conexión con múltiples estaciones, y así garantizar una respuesta ágil y eficiente del sistema frente a demandas variables y picos de tráfico. Las peticiones encoladas son luego dirigidas al segundo microservicio, también implementado en Go Kit, especializado en la escritura de datos en una base de datos MongoDB. Este microservicio recibe los datos desde la cola y los procesa para su almacenamiento como registros en colecciones específicas de MongoDB. La decisión de utilizar MongoDB como

base de datos subyacente aporta beneficios clave para la aplicación. Como base de datos NoSQL, MongoDB ofrece una alta escalabilidad y flexibilidad, optimizando tanto el almacenamiento como la recuperación de datos no estructurados, lo cual es crucial en entornos IoT donde los datos suelen ser dinámicos y variados. Además, la arquitectura incluye un módulo de gestión, que se encarga de orquestar la infraestructura del sistema utilizando herramientas de infraestructura como código, como Terraform, y la orquestación de contenedores con Kubernetes. Este módulo asegura una gestión coherente y reproducible de la infraestructura, permitiendo el escalado horizontal del sistema para manejar eficientemente los picos de demanda. La combinación de microservicios en Go con MongoDB permite una arquitectura distribuida y modular, que es crucial para manejar grandes volúmenes de datos y adaptarse dinámicamente a cambios en la demanda. Esta arquitectura no solo mejora la capacidad de respuesta y la resiliencia del sistema, sino que también facilita el mantenimiento y la evolución continua de la aplicación, garantizando un rendimiento óptimo y una gestión eficaz del ciclo de vida de los datos en un entorno IoT dinámico y exigente [14].

14.4.1 Módulo de gestión

La arquitectura del sistema está implementada utilizando Terraform, una herramienta de infraestructura como código y desplegada en contenedores orquestados con Kubernetes. Esta combinación ofrece ventajas significativas en términos de escalabilidad y despliegue continuo. Terraform permite la gestión declarativa de la infraestructura, asegurando una configuración coherente y reproducible, mientras que Kubernetes facilita la orquestación y gestión de contenedores, permitiendo el escalado horizontal del sistema, aumentando la cantidad de instancias de pod, para manejar eficazmente los picos de demanda. Además, se ha integrado el despliegue automático mediante herramientas como GitHub Actions, optimizando el proceso de integración continua y entrega para asegurar la generación de la imagen con las aplicaciones listas para ser desplegadas.

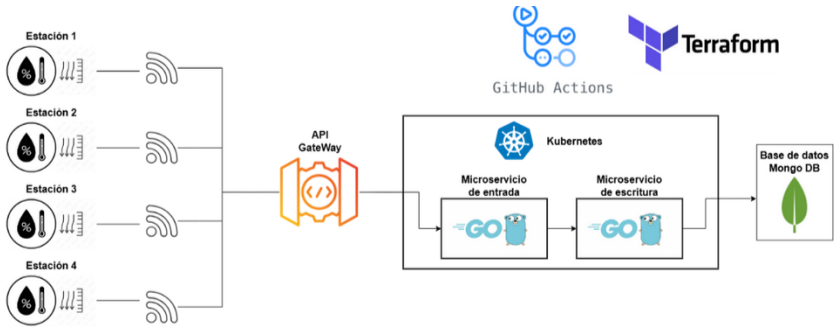


Fig. 14.3: Arquitectura del sistema completo. Elaboración propia.

14.5 Implementación y resultados

Las estaciones implementadas pueden comunicarse a través de redes wifi donde haya cobertura o a través de radio frecuencia con módulos nRF24L01 en una configuración de red multiconductora. Esta configuración permite que cada canal de RF se divide lógicamente en seis canales de datos

paralelos, permitiendo que un receptor central recoja información de seis nodos transmisores diferentes simultáneamente. El nodo central está conectado a una mini PC con Raspberry Pi, en la cual se ha implementado una aplicación para realizar peticiones POST a una API. Se puede observar en la Fig. 14.4 la estación de monitoreo finalizada en una placa de circuito impreso universal.

Fig. 14.4: Estación de monitoreo desarrollada. Elaboración propia.

Luego se procede a ubicar las estaciones en dos puntos geográficos distintos para hacer la toma de mediciones. Se instalan las estaciones en la ciudad de Manizales, Colombia. Se hace validación de que las estaciones tomen adecuadamente los registros, como se puede observar en la Fig 14.5, que gráfica las mediciones tomadas entre las 17:20:34 del 20 de junio de 2024 y las 18:23:34 del 21 de junio de 2024.

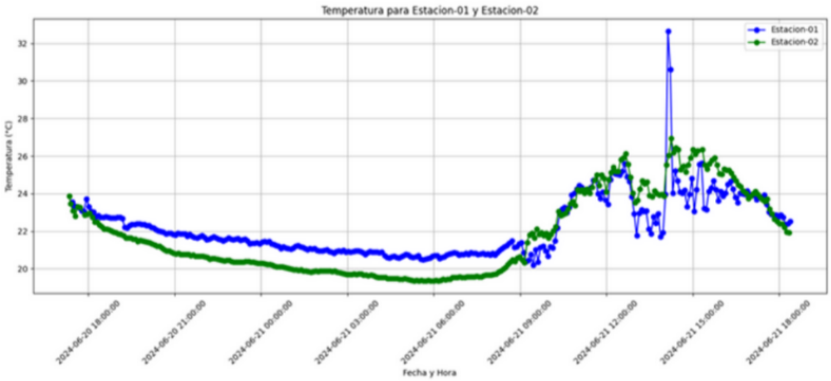


Fig. 14.5: Registros de temperatura de dos estaciones en funcionamiento. Elaboración propia.

14.6 Conclusión

El sistema desarrollado demuestra que la integración de tecnologías IoT, como la NodeMCU V3 ESP8266, es no solo técnica y funcionalmente viable, sino también una solución económicamente accesible para la creación de estaciones de monitoreo ambiental. Esta propuesta permite la recolección y transmisión de datos en tiempo real de manera remota, ofreciendo una alternativa viable para comunidades y regiones con recursos limitados, que de otro modo no podrían implementar redes de monitoreo de calidad, debido a los costos asociados con equipos comerciales. El impacto de soluciones tecnológicas, como la presentada en este trabajo, se extiende más allá de la simple reducción de costos; al establecer una red de monitoreo ambiental accesible, se habilita una mayor cobertura geográfica y una mejor capacidad para el seguimiento continuo de variables climáticas críticas. Esta capacidad es crucial en la era del cambio climático, donde la información precisa y oportuna es fundamental para la toma de decisiones informadas y la gestión de riesgos. Como trabajo futuro, está el desafío de abordar la calidad de las medidas de estos sensores realizando mediciones simultáneas entre una estación comercial y la estación propuesta, con el fin de comparar las lecturas para analizar desviaciones y proponer ajustes de calibración.

Referencias

1. X. Zhang, B. Tang, Q. Yang, W. Xu and F. Guo, ”Cost-Optimized Microservice Deployment for IoT Application in Cloud-Edge Collaborative Environment,” 2023 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Rio de Janeiro, Brazil, 2023, pp. 873-878.
2. Definición de los indicadores de la línea base ambiental de Caldas / Hábitat, 2019-2021 / Jeannette Zambrano Nájera, editora académica. – Segunda edición. –Manizales: Universidad Nacional de Colombia. Facultad de Ingeniería y Arquitec- tura; Corporación Autónoma

Regional de Caldas (Corpocaldas); Bogotá: Editorial Universidad Nacional de Colombia, 2022.
3. P. Kayal, "Kubernetes in Fog Computing: Feasibility Demonstration, Limitations and Improvement Scope : Invited Paper," 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), New Orleans, LA, USA, 2020, pp. 1-6.
4. R. K. Kodali and S. Mandal, "IoT based weather station," 2016 International Con- ference on Control, Instrumentation, Communication and Computational Tech- nologies (ICCICCT), Kumaracoil, India, 2016, pp. 680-683, doi: 10.1109/ICCI- CCT.2016.7988038.
5. Siddiqui, H., Khendek, F., & Toeroe, M. (2023). Microservices based architectures for IoT systems - State-of-the-art review. In Internet of Things (Netherlands) (Vol. 23). Elsevier B.V.
6. Walker, D., Pitt, M., & Thakur, U. J. (2007). Environmental management systems. Journal Of Facilities Management, 5(1), 49-61. https://doi.org/10.1108/14725960710726346.
7. Fernández-Ahumada, Luis Manuel, Jose Ramírez-Faz, Marcos Torres-Romero, and Rafael López-Luque. 2019. "Proposal for the Design of Monitoring and Operating Irrigation Networks Based on IoT, Cloud Computing and Free Hardware Technolo- gies" Sensors 19, no. 10: 2318.
8. S. Al-Sarawi, M. Anbar, K. Alieyan and M. Alzubaidi, "Internet of Things (IoT) communication protocols: Review," 2017 8th International Conference on Information Technology (ICIT), Amman, Jordan, 2017, pp. 685-690, doi: 10.1109/ICITECH.2017.8079928.
9. Elhadi, Sakina and Marzak, Abdelaziz and Sael, Nawal and Merzouk, Soukaina, Comparative Study of IoT Protocols (May 28, 2018). Smart Application and Data Analysis for Smart Cities (SADASC'18).
10. R. A. Sadek and H. M. Elbadawy, "Towards IoT Era with current and Future Wireless Communication Technologies: An Overview," 2022 39th National Radio Science Conference (NRSC), Cairo, Egypt, 2022, pp. 343-353.
11. X. Zhang, B. Tang, Q. Yang, W. Xu and F. Guo, "Cost-Optimized Microservice Deployment for IoT Application in Cloud-Edge Collaborative Environment," 2023 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD), Rio de Janeiro, Brazil, 2023, pp. 873-878.
12. P. Kayal, "Kubernetes in Fog Computing: Feasibility Demonstration, Limitations and Improvement Scope : Invited Paper," 2020 IEEE 6th World Forum on Internet of Things (WF-IoT), New Orleans, LA, USA, 2020, pp. 1-6.
13. Siddiqui, H., Khendek, F., & Toeroe, M. (2023). Microservices based architectures for IoT systems - State-of-the-art review. In Internet of Things (Netherlands) (Vol. 23). Elsevier B.V.
14. Christian, J., Steven, N., Kurniawan, A., & Anggreainy, M. S. (2023). Analyzing Microservices and Monolithic Systems: Key Factors in Architecture, Development, and Operations. IEEE. https://doi.org/10.1109/ic2ie60547.2023.10331155.

15
Systematic review of methods for searching learning objects

Germán A. Osorio-Zuluaga[1][0000-0002-5613-5858]
Sebastian Robledo[2][0000-0003-4357-4402]

Abstract Information retrieval is a field that has grown exponentially in recent years, finding diverse applications such as learning objects that have enriched scientific advancements. However, this surge in academic production has re-sulted in a dispersion of contributions, with few articles consolidating these con-tributions into a single document. Therefore, this scientometric review aims to identify the dynamics of scientific production on information retrieval focused on learning objects. A search was conducted in Scopus and Web of Science to identify articles related to the topic for subsequent consolidation and analysis. The results are categorized into annual production, analysis of countries, journals, and scientific collaboration among researchers. The findings indicate an overall growth rate of 8.14% and a production share of 11.53% by China, while the United States leads in impact, accounting for 18.8% of total citations. This article is crucial for researchers new to the fascinating study of learning objects, provid-ing them with a comprehensive overview of the field.

Keyword Bibliometric Analysis, Learning Objects, Information Retrieval, Repositories, e-Learning, Distance Learning.

15.1 Introduction

The rapid growth of the Web has boosted distance education. Many organizations now offer and share educational resources in various formats globally. These resources, known as learning objects (LO), are typically stored in repositories. LOs can be tagged with metadata to ease their search and retrieval, providing structured descriptions [1]. Repositories can also be federated, centralizing resources in a single portal to enhance visibility and streamline access to LOs [2]. To reuse and share LOs between repositories, standardized protocols like IEEE-LOM, Dublin Core, Can Core, and OBAA have been established. These protocols define the syntax and semantics needed to describe an LO [1]. For instance, IEEE-LOM includes nine categories and about seventy descriptive elements [2]. Searching and retrieving learning objects in repositories primarily relies on metadata [2]. Users pre-select relevant topics like author, title, and subject, and the search engine matches query terms with structured metadata. In other web areas, the use of full-text searches prevails. In this sense,

[1]Universidad Nacional de Colombia, Facultad de Administración, Manizales, Colombia
[2]Dirección Académica, Universidad Nacional de Colombia, La Paz (César), Colombia.
e-mail: \{gaosorioz,srobledog\}@unal.edu.co

full-text searches use algorithms to match query terms with words in individual documents of a repository [3]. This bibliographic review aims to analyze the papers that refer to the search and selection of learning objects in repositories, whether done by metadata or by full-text.

15.2 Methodology

The search was conducted in the Web of Science (WoS) and Scopus databases (see Table 15.1). WoS indexes 8.7 thousand high impact research journals[1], while Scopus con-tains 7+ thousand publishers, 29.2+ thousand active serial titles, 330+ thousand books, and 23.4+ million open access items[2]. WoS and Scopus are the most recommended databases for conducting systematic literature reviews due to their extensive coverage of high-quality, peer-reviewed journals across disciplines. Both apply strict selection criteria, ensuring only reputable publications are included, enhancing research credibility. Additionally, they offer advanced search tools, citation analysis, and impact met-rics, facilitating thorough and precise searches. Their prestige and wide acceptance in academia further validate their use for comprehensive literature reviews while mini-mizing redundancy in results. This approach signifies a new trend in scientometric studies [4, 5, 6]. The keywords pertained to information retrieval and learning objects, resulting in 1331 unique documents from WoS and Scopus. These results enabled a scientometric mapping of recent scientific production dynamics. The analysis revealed the most productive countries and their collaborations, the journals with the highest impact, and the collaboration net-works among researchers. This mapping offers new researchers a comprehensive over-view of global scientific dynamics in this field.

Table 15.1: Search Parameters for information retrieval.

Parameter		
Range	2003 - 2024	
Date	18 - June - 2024	
Document type	Article, Review, Conference, Data Paper, Proceeding, Papers	
Words	TITLE ((search* OR retrieve* OR extract* OR select* OR discov* OR identif* OR rank* OR classif* OR deliver* OR index*) AND (”learning object*” OR ”knowledge object*” OR elearning OR e-learning OR ”educat* resourc*” OR ”educat* mater*” OR ”educat* content*” OR ”teach* mater*”))	
	Web of Science	**Scopus**
Partial Results	448	1242
Total(WoS + Scopus)	1331	

[1] http://scientific.thomson.com/products/wos/

[2] https://www.elsevier.com/products/scopus/content

15.3 Scientometrics Analysis

15.3.1 Overall Analysis

Figure 15.1 shows the behavior of the number of publications and citations in the period from 2003 to 2023. During this time, there was an average growth of 8.14% per year. In this respect, it cannot be said that this growth has been homogeneous. Two distinct periods can be proposed: one, from 2003 to 2008; the other, from 2009 to 2021. The years 2022 and 2023 are excluded from the analysis due to the 2-3 year lag between article publication and citation, as well as journal acceptance times. These two periods are analyzed individually below.

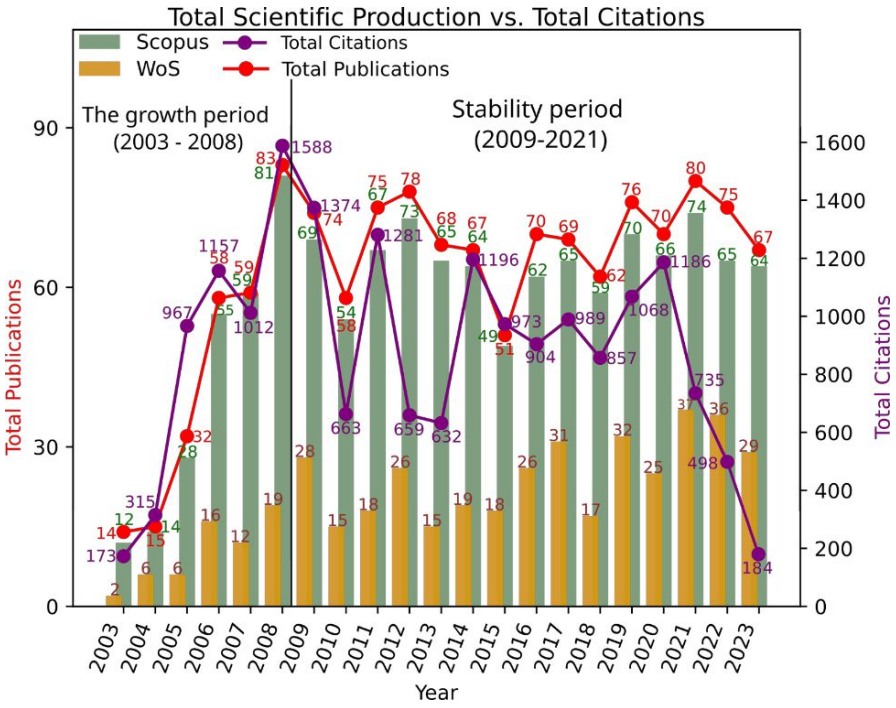


Fig. 15.1: Total Scientific Production and Citations Over Time (2003-2023)

The growth period (2003 - 2008)
In the first period, the average growth was 42.75%. In 2008, the maximum peak was seen in both publications and citations. This increase in publications could be related to the growing adoption of e-learning technologies and the need for reusable digital edu-cational resources. During these years, the academic and educational community began to recognize the importance of learning objects and digital repositories as key tools to improve teaching and learning. The increasing number of citations reflects the impact and relevance of the research published in this period. The high citation rate suggests that the published works were innovative and provided valuable solutions for the educational community.

Stability period (2009-2021) During this period, the average annual growth was -0.71%, which can be considered zero growth. Regarding the number of publications, it is observed that since 2009, the number of publications shows fluctuations, with some years of increase and others of decrease. For example, there was a notable decrease in 2010 and 2014, followed by peaks in 2011 and 2015. These fluctuations may be due to several factors, such as changes in research priorities, the emergence of new technologies and methodologies, and variations in the funding of research projects. In addition, the consolidation of standards and best practices in the use of learning objects could have led to a lower need for new publications. On the other hand, citations also show fluctuating behavior, with peaks in 2011 and 2015, and a general downward trend after 2016. The decline in citations could be related to the maturity of the field. As technologies and methodologies become more established, more recent research may not be as innovative or impactful as the initial research, resulting in fewer citations.

15.3.2 Country Analysis

Table 15.2 presents data on scientific production and citations related to learning objects in repositories, categorized by country, along with the distribution of publications across different quartiles (Q1, Q2, Q3, Q4). The analysis considers two main aspects: the volume of production and citations, and their distribution across the quartiles of the journals in which they were published. An interesting observation regarding the quality and impact (citations) of research on learning objects is that China leads in productivity, while the United States has the highest impact. This is likely due to the quality of their papers, the majority of which are published in Q1 journals.

Table 15.2: Scientific Production and Citation Distribution by Country and Quartile

country	Production		Citación		Q1	Q2	Q3	Q4
China	150	11.53%	1425	11.41%	24	16	11	17
India	127	9.76%	1031	8.25%	14	15	15	19
USA	125	9.61%	2349	18.8%	37	19	8	3
United Kingdom	73	5.61%	1058	8.47%	15	6	5	0
Spain	61	4.69%	889	7.12%	6	5	8	1
Canada	44	3.38%	205	1.64%	6	2	6	1
Italy	44	3.38%	407	3.26%	4	2	2	5
Australia	43	3.31%	449	3.59%	12	3	3	1
Germany	36	2.77%	228	1.83%	2	5	1	0
Brazil	35	2.69%	101	0.81%	2	1	4	2

China and India exhibit high publication output, yet their citation counts are relatively low compared to those of the USA. The USA, with 18.8% of the total citations, demon-strates a significantly higher impact on its publications. The United Kingdom and Spain also show notable publication and citation figures, though still below those of the USA. The USA's high proportion of publications in the Q1 quartile (37) suggests that their research is frequently published in high-impact journals. In contrast, China and India display a more balanced distribution across quartiles, with substantial representation in both Q1 and Q2. Spain and Canada, however, have a lower proportion of publications in Q1, which might indicate lesser visibility or impact of their research. Figure 15.2 illustrates the collaboration network among countries in the field of infor-mation retrieval. The primary community is led by the United Kingdom, the Nether-lands, and the United States. The "Nodes and Links Through Time" subfigure demon-strates the consolidation of relationships

within the information retrieval community. Since 2014, the percentage of collaborative links has been increasing at a higher rate than the addition of new participating countries.

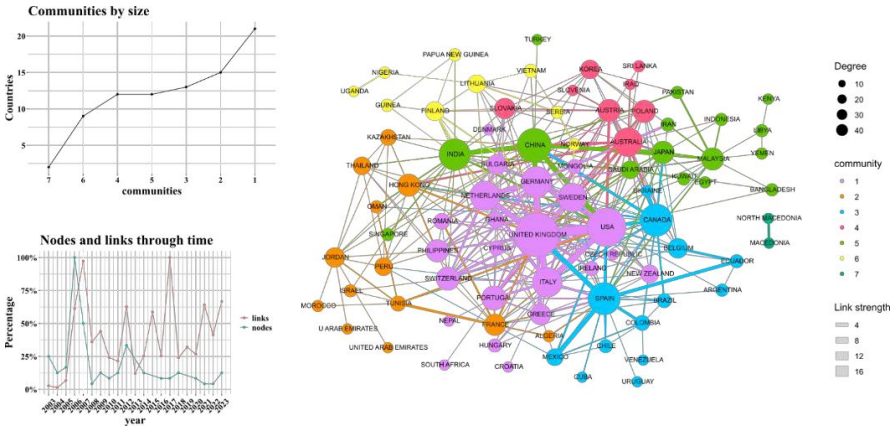


Fig. 15.2: Global Collaboration Network Among Countries in Learning Object Research

15.3.3 Author Collaboration Network

Table 15.3 presents the most productive authors based on the number of published docu-ments. Professor Yanshan Wang is the author with the most articles and holds the fifth-highest h-index. His research focuses on the application of information retrieval to elec-tronic health records and medical educational materials [7, 8]. On the other hand, professor Cristóbal Romero has the highest h-index due to his contributions to the use of data mining to facilitate online learning [9, 10]. Overall, it is evident that there are globally recognized scientists with significant contributions in their respective fields. Figure 15.3 shows the personal social networks of leading researchers in

Table 15.3: Productivity and Impact Metrics of Leading Researchers

No	Researcher	Total Arti-cles*	Scopus h-Index	Affiliation
1	Wang Y	9	22	University of Pittsburgh, United States.
2	Chen S	8	10	National Taiwan Ocean University, Keelung, Taiwan
3	De L P F	8	24	Universidad de Salamanca, Salamanca, Spain
4	Limongelli C	8	18	Università degli Studi Roma Tre, Rome, Italy
5	Nasraoui O	8	29	University of Louisville, Louisville, United States
6	Zhuhadar L	8	13	Western Kentucky University, Bowling Green, United States
7	Gil A	7	11	Universidad de Salamanca, Salamanca, Spain
8	Li Y	7	32	Queensland University of Technology, Brisbane, Australia
9	Romero C	7	41	Universidad de Córdoba, Córdoba, Spain
10	Sciarrone F	7	22	Universitas Mercatorum, Rome, Italy

text analysis and mining. Five major components emerged, each led by a prominent researcher. While they share expertise, their applications differ, leading to three distinct clusters. The field of information retrieval is mature, with growing connections among researchers.

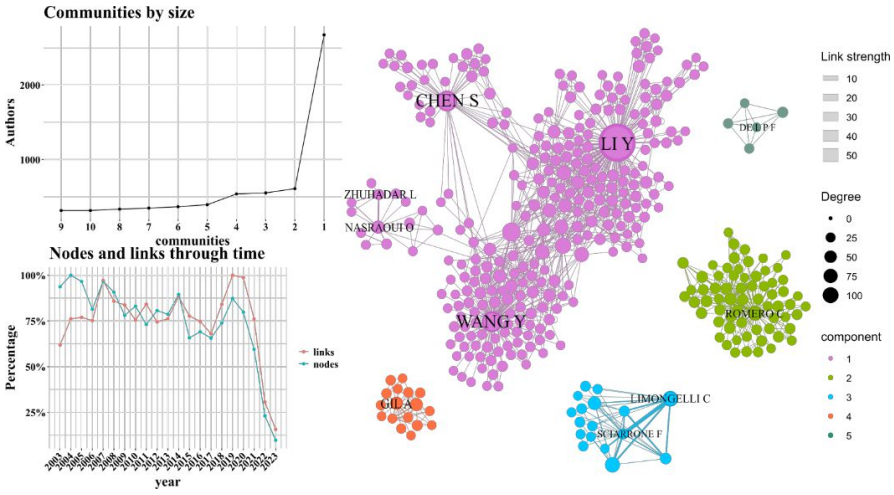


Fig. 15.3: Personal Collaboration Network of the Most Productive Authors

15.4 Conclusions

A bibliographic review was carried out on the search for learning objects in repositories based on the Scopus and Web of Science databases in the period from 2003 to 2023. For this purpose, the same search equation was used, which obtained 1331 unique documents, some of them common to both databases. The overall analysis of publications and citations over the years in the field shows an average growth of 8.14% per year between 2003 and 2023. However, this growth has not been homogeneous. Two clearly defined periods are evident: an initial period of exponential growth from 2003 to 2008, followed by a period of relative stability from 2009 to 2021. Regarding scientific production by country, China leads in the volume of publications, while the United States has the greatest impact in terms of citations. Most of the publications from the United States are concentrated in first-quartile (Q1) journals, suggesting higher quality and visibility of its research in this field. The analysis of the most productive authors in the field of learning objects in repositories highlights the significant contributions of professors Yanshan Wang and Cristóbal Romero. Yanshan Wang leads in the number of published articles, focusing on in-formation retrieval applications in electronic health records. Cristóbal Romero has the highest h-index due to his impactful research in data mining for online learning. The personal social networks of these researchers reveal five major components, led by prominent figures in the field. These networks demonstrate the maturity and well-established nature of the information retrieval field, with strengthening connections among researchers over time. The distinct clusters within the network reflect the diverse applications of text analysis and mining expertise, contributing to the field's robustness and interdisciplinary reach. In conclusion, the results of the scientometric mapping indicated that information retrieval in learning objects is a mature field with a well-established scientific community.

References

1. Ternier S, Duval E, Massart D, et al (2008) Interoperability for Searching Learning Object Repositories: The ProLearn Query Language. D-Lib Magazine 14:1
2. Duque-Méndez ND, Arturo Ovalle Demetrio, Moreno Cadavid J (2014) Objetos de Aprendizaje, Repositorios y Federaciones: Conocimiento para Todos. Universidad Nacional de Colombia, Manizales, Colombia
3. Beall J (2008) The Weaknesses of Full-Text Searching. The Journal of Academic Librarianship 34:438–444. https://doi.org/10.1016/j.acalib.2008.06.007
4. Robledo-Giraldo S, Figueroa-Camargo JG, Zuluaga-Rojas MV, et al (2023) Mapping, evolution, and application trends in co-citation analysis: a scientometric approach. Rev investig desarro innov 13:201–214. https://doi.org/10.19053/20278306.v13.n1.2023.16070
5. Ariza-Colpas PP, Piñeres-Melo MA, Morales-Ortega R-C, et al (2024) Sustainability in Hybrid Technologies for Heritage Preservation: A Scientometric Study. Sustainability 16:1991. https://doi.org/10.3390/su16051991
6. Urina-Triana MA, Piñeres-Melo MA, Mantilla-Morrón M, et al (2024) Machine Learning and AI Approaches for Analyzing Diabetic and Hypertensive Retinopathy in Ocular Images: A Literature Review. IEEE Access 12:54590–54607. https://doi.org/10.1109/ACCESS.2024.3378277
7. Zeng Y, Liu X, Wang Y, et al (2017) Recommending Education Materials for Diabetic Questions Using Information Retrieval Approaches. J Med Internet Res 19:e342. https://doi.org/10.2196/jmir.7754
8. Miranda O, Kiehl SM, Qi X, et al (2024) Enhancing post-traumatic stress disorder patient assessment: leveraging natural language processing for research of domain criteria identifi-cation using electronic medical records. BMC Med Inform Decis Mak 24:154. https://doi.org/10.1186/s12911-024-02554-8
9. Romero C, González P, Ventura S, et al (2009) Evolutionary algorithms for subgroup discovery in e-learning: A practical application using Moodle data. Expert Systems with Applications 36:1632–1644. https://doi.org/10.1016/j.eswa.2007.11.026
10. Cerezo R, Lara J-A, Azevedo R, Romero C (2024) Reviewing the differences between learning analytics and educational data mining: Towards educational data science. Computers in Human Behavior 154:108155. https://doi.org/10.1016/j.chb.2024.108155

16
Metodología para Incorporar la Inteligencia Artificial a las Compras del Estado Colombiano por medio de Recomendaciones para la Agregación de Demanda

Israel Steven Orozco Rodríguez[1][0009-0002-6373-7686]
John Corredor Franco [2][0000-0001-5121-5878]

Resumen El Estado Colombiano, tiene como objetivo alcanzar variedad, eficiencia y eficacia en la compra de los productos $bienes, servicios y obras$ que necesita para operar e implementar las políticas públicas en el país. El Estado emplea, entre otras, estrategias de agregación de demanda para alcanzar los objetivos planteados. En el presente documento, se propone una metodología basada en Inteligencia Artificial para fortalecer los esquemas de agregación de demanda para dicho fin. La metodología se compone de dos procesos de depuración y tratamiento de datos para implementar algoritmos de recomendación basados en reglas de asociación y filtros colaborativos, respectivamente, y en el marco de un proceso de experimentación con métricas de desempeño. Los resultados muestran la funcionalidad de la metodología de recomendación.

Palabras clave Acuerdos Marco de Precio, Agregación de demanda, Compras Públicas, Filtros Colaborativos, Inteligencia Artificial, Reglas de Asociación, SECOP, Sistemas de Recomendación.

16.1 Introducción

El presente artículo se enmarca en el contexto de la ejecución del gasto público por medio de las compras públicas en el país. El Estado Colombiano busca alcanzar variedad, eficiencia y eficacia en la compra de los productos (bienes, servicios y obras). La agregación de demanda es un instrumento, para que sus Entidades Estatales sumen sus necesidades y actúen en forma coordinada en el mercado y así obtener eficiencia en el gasto [1]. No obstante, las entidades estatales han realizado compras por medio de los instrumentos de agregación demanda por valores que, en promedio, no han superado el 3% de la cuantía total de las compras y contrataciones públicas al año en Colombia [2]. Adicionalmente, entre 2019 y 2022, las entidades adquirieron productos en al menos 1.280 categorías difer- entes. En este sentido, los instrumentos de agregación de demanda cuentan con un márgen ámplio de crecimiento para lo cual el presente documento propone la exploración de procesos de depuración y preparación de datos para implementar técnicas de Inteligencia Artificial con el objetivo de generar una metodología novedosa que permita crear recomendaciones de nuevos productos o clases de productos para la agregación de demanda, específicamente para los Acuerdos Marco de Precios, establecidos por la Ley 1150 de 2007 y Reglamentados por el Decreto 1082 de

[1,2]Pontificia Universidad Javeriana, Bogotá, Colombia
e-mail: isteven.orozcor@javeriana.edu.co
e-mail: corredor-johnj@javeriana.edu.co

2015. El resto del documento se encuentra estructurado de la siguiente forma. En la Sección II se realiza una revisión del estado del arte. En la sección III se discuten las principales características del Sistema de Compra Publica. En la Sección IV se describe la metodología de incorporación de Inteligencia Artificial. Finalmente, la Sección V contiene las conclusiones de la investigación.

16.2 Revisión de literatura

La búsqueda de literatura no identificó desarrollos de algoritmos para la generación de recomendaciones de agregación de demanda en las compras públicas.

16.3 El Sistema de Compra Pública

Todas las entidades del Estado, independientemente de su nivel o su ubicación, son compradores de productos y deben seguir los mismos principios: transparencia, economía y responsabilidad. Adicionalmente, las entidades deben identificar los productos que adquieren por medio del Código Estándar de Productos y Servicios de Naciones Unidas -UNSPSC. El código cuenta con una estructura jarárquica en el siguiente órden Segmento; Familia; Clase; y Producto.

16.4 Metodología para incorporar Inteligencia Artificial a las Compras del Estado Colombiano por medio de recomendaciones a la Agregación de Demanda

La Figura 16.1 presenta la metodología para incorporar Inteligencia Artificial en la generación de recomendaciones de agregación de demanda en Colombia. La metodología esta compuesta por las siguientes etapas: (0) entorno de desarrollo; (i) gestión de datos; (ii) contexto teórico; (iii) modelación y experimentación; (iv) resultados según métricas. De esta forma, la presente investigación utiliza la metodología CRISP-DM para trabajos con minería de datos [4]. Se aplican dos tipos de algoritmos para la etapa (iii) de modelación e implementación de algoritmos de recomendación: (a) reglas de asociación y (b) filtros colaborativos.

Las dos aproximaciones generan herramientas diferenciadas que utilizan para su funcionamiento estructuras de datos particulares, por lo cual se generan dos (2) procesos de depuración y tratamiento especializados. Los resultados de ambos procesos se complementan en la tarea de mejorar la variedad, eficacia y eficiencia obtenida por las entidades públicas en sus compras y contrataciones por medio de la agregación de demanda. Se resalta que los procesos implementados son solo dos (2) de los múltiples procesos posibles en esta línea de investigación de sistemas de recomendación aplicados a la agregación de demanda en las compras públicas Colombianas.

16.4.1 Etapa 0: Entorno de hardware y software

La presente investigación se desarrolló sobre un equipo Laptop-x64-based PC con 11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80GHzm 2803 Mhzm 4 Core(s) con una RAM de 16 GB. Como entorno

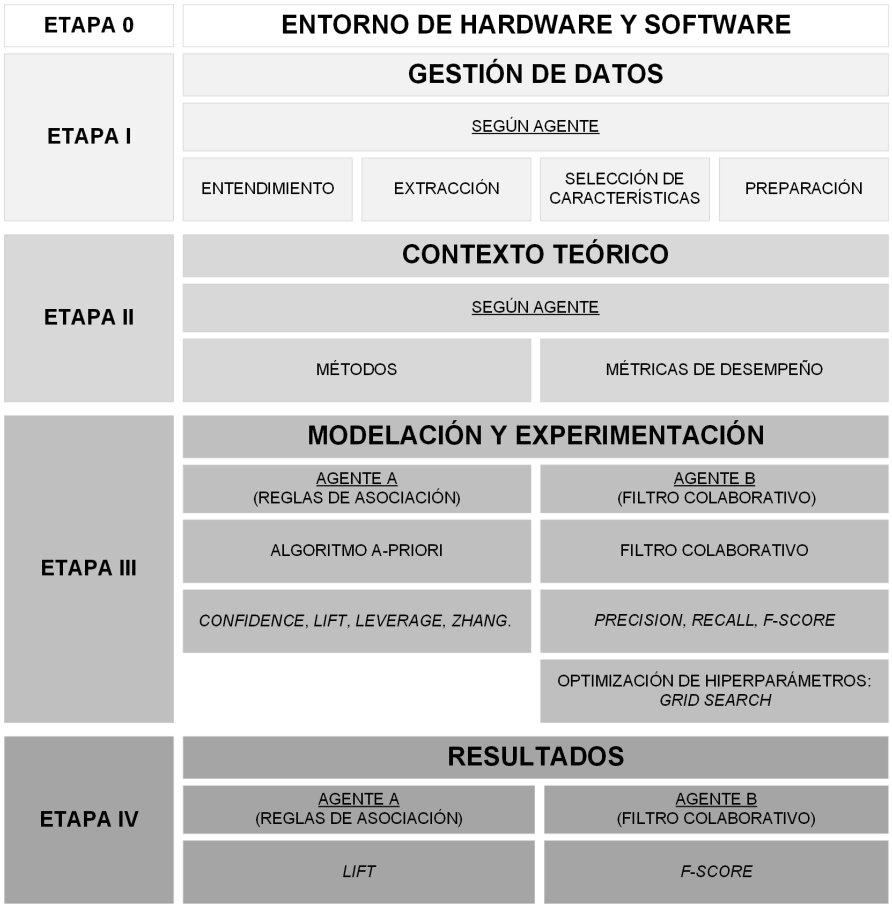


Fig. 16.1: Metodología para incoporar recomendaciones en compras públicas

de software se utilizaron códigos de Python en Colab Notebooks con el objetivo emplear el uso de recursos en nube de Google y sacar provecho de la disponibilidad de códigos y repositorios en línea [3].

16.4.2 Etapa I: Gestión de datos

Los datos de compra fueron descargados del portal de Datos Abiertos del SECOP el 31 de enero de 2024, fecha para la cual el total del conjunto presentaba un total de 5.054.825 registros (procesos de compra), cada uno con un total de 59 características o campos de información. Los filtros Implementados son los siguientes:

- Periodo desde enero de 2019 hasta diciembre de 2022: Captura los procesos de compra desde entidades con madurez en el uso del sistema transaccional (en operación desde 2015) al tiempo que

contiene información reciente que no debería presentar modificaciones por ajustes contractuales (adendas, prórrogas, modificaciones) o correcciones por calidad de datos.
- Procesos adjudicados: Contiene solo los procesos de selección que se han materializado en un contrato (compra efectiva) asociado.
- Modalidad de contratación Licitación Pública, Selección Abreviada, Mínima Cuantía.
- ID del portafolio: Es el identificador único de cada compra (transacción). Se filtra por valores únicos, de tal forma que un mismo proceso no aparezca más de una (1) vez en la base.

El número de registros se reduce a 302.761 procesos con la aplicación de los filtros descritos. Cada registro cuenta con 59 características o columnas de infor- mación. Se seleccionan las siguientes columnas de la base resultante para cada uno de los procesos (”A” y ”B”) de acuerdo con los objetivos de la investigación y los requerimientos de los algoritmos implementados: **Proceso A. Reglas de asociación**

- Código Principal de la categoría: Es el código de la categoría UNSPSC prin- cipal del objeto del contrato, es decir, el producto principal de la “canasta” adquirida por la entidad en un proceso.
- Categorías Adicionales: Categorías UNSPSC adicionales asociadas al proceso de compra, es decir, los productos adicionales (no principales) adquiridos por la entidad estatal en la “canasta” adquirida en una compra.

Proceso B. Filtros colaborativos

- NIT Entidad. Es el número de identificación de cada comprador (entidad estatal compradora).
- Código Principal de la categoría: Es el código de la categoría UNSPSC prin- cipal del objeto del contrato, es decir, el producto principal de la “canasta” adquirida por la entidad en un proceso.

Preparación de datos
Se realiza la preparación diferencial de datos según el proceso implementado de la siguiente forma:

Proceso A. Reglas de asociación
- Se descompone la columna “56. Categorías Adicionales” en los códigos UN- SPSC. Cada código identificado se incorpora en una columna individual. Como resultado, se obtienen 42 columnas con códigos UNSPSC (productos secundarios en canasta). El 42 corresponde al máximo número de productos secundarios con el que cuenta al menos una de las transacciones (contratos) contenidos en la base.
- Se toma solo la columna “Código Principal de la categoría” y el primer reg- istro de código UNSPSC derivado de la columna “Categorías Adicionales” para reducir la dimensionalidad y facilitar la implementación de las Re- glas de Asociación, De esta forma, la base de transacciones cuenta solo con “canastas” de dos (2) productos.
- Se ordena la base de códigos de menor a mayor por Código Principal UN- SPSC.
- La base resultante se incorpora en un Diccionario de Python con 56 llaves, cada una correspondiente a los Segmentos con los que cuenta el Clasifi- cador V.14 de Naciones Unidas. Esta partición se realiza con base en los dos primeros dígitos cada código UNSPSC. Como ejemplo, el Segmento del código “95141901” es “95”, el cual refiere a “Terrenos, Edificios, Estructuras y vías”.
- Se traduce cada código UNSPSC de la base, tanto principal como secundario, a nombres de producto con base en el Clasificador V.14 de Naciones Unidas a lo largo del Diccionario y sus 56 llaves. Esto significa que se cruza el Diccionario con transacciones con la base de nombres asociados a cada código estándar UNSPSC de la versión 14. Como ejemplo, el código “95141901” se transforma en la base en “Unidad médica”.
- Se crean las matrices de incidencia binarias para cada una de las 56 llaves del Diccionario de transacciones. Las columnas de cada matriz corresponden a los nombres únicos de los productos adquiridos por las entidades en las transacciones agrupadas en cada llave del Diccionario. Por cada fila (transac- ción, compra estatal) se presentan valores de “1” cuando el producto de la columna fue efectivamente adquirido en dicha transacción. El “0”, por su parte, significa que el producto de la columna no fue adquirido en la compra representada en la fila respectiva.

Proceso B. Filtros colaborativos

- Se toman únicamente las columnas, según numeración de Diccionario de Datos del conjunto "Procesos de Contratación" "NIT Entidad" y "Código Principal de la Categoría".
- Se reorganiza la la base de datos para contar, en su orden, con las columnas de Código UNSPSC y NIT de Entidad. De esta forma, la identificación de cada entidad se repite tantas veces como productos haya adquirido en las transacciones bajo análisis.
- Se toma el código UNSPSC hasta los dígitos de Clase según el código están- dar de la V.14 de Naciones Unidas.
- Se calcula el número de veces que cada entidad ha comprado o adquirido cada uno de los Códigos UNSPSC. El resultado es una preferencia revelada del comprador (entidad estatal) sobre el producto particular. A mayor resultado, mayor preferencia sobre el producto. Este conteo se incluye en una nueva columna llamada "rating".
- Para la implementación del Filtro Colaborativo, se reorganiza la base de datos en el siguiente orden de columnas ("nombre de columna en la nueva base"): Código UNSPSC ("itemID") / NIT Entidad ("userID") / Conteo de compras por producto para cada entidad ("rating").

La columna "rating" resultante cuenta con valores desde 1 hasta 4.230, por lo cual se decide dejar solo los 'ratings' que estén entre '2' y '5' . Como resultado, el número de registros totales pasan de 117.066 a 31.925. Se incorpora la base resultante con las columnas 'item id', 'user id' y 'rating' en un Diccionario de Python con 56 llaves, cada una correspondiente a los Segmentos con los que cuenta el Clasificador V.14 de Naciones Unidas. Esta partición se realiza con base en los dos primeros dígitos cada código UNSPSC. Como ejemplo, el Segmento del código "40161803" es "40", el cual refiere a "Componentes y Equipos para Distribución y Sistemas".

16.4.3 Etapa II: Contexto teórico

Las Reglas de Asociación ("affinity analysys" o "market basket análisis") y los Filtros colaborativos son métodos no supervisados utilizados con frecuencia en marketing y estrategias de venta. Las Reglas de Asociación buscan establecer "qué va bien con qué" a partir del análisis de una gran cantidad de transac- ciones. Por su parte, los Filtros Colaborativos implementados para el proceso B permiten generar recomendaciones personalizadas a un comprador con base en su información y la de compradores similares a él.

16.4.4 Etapa III: Modelación y Experimentación

La metodología desarrollada en el presente documento implementa dos (2) proce- sos basados en algoritmos de Reglas de Asociación (Proceso A) y Filtros Colabo- rativos (Proceso B), respectivamente con el objetivo de generar recomendaciones encaminadas a la agregación de demanda.

Proceso A. Reglas de asociación

Sobre las 56 matrices de incidencia binarias, correspondiente a las 56 llaves (Segmentos del código UNSPSC) del Diccionario de transacciones generado pre- viamente, se implementa el algoritmo Apriori para obtener los itemsets con un soporte mínimo de 0,01. Seguidamente se derivan las reglas de asociación con un umbral mínimo de confianza de 0,3 y con las métricas de support, confidence, lift, leverage y zhang. La figura 16.2 muestra la distribución de los lift obtenidos a lo largo de todo el diccionario de transacciones y todos los Segmentos sobre los cuales se implementó el algoritmo Apriori.

De acuerdo con la Figura 2 los resultados de lift son mayores que 0.39 en todos los casos y el 73,9% del total de los lift calculados, se encuentran en bins con centro 7.71 o superiores, reflejando una buena robustez en las reglas de asociación encontradas para todos los Segmentos de compra

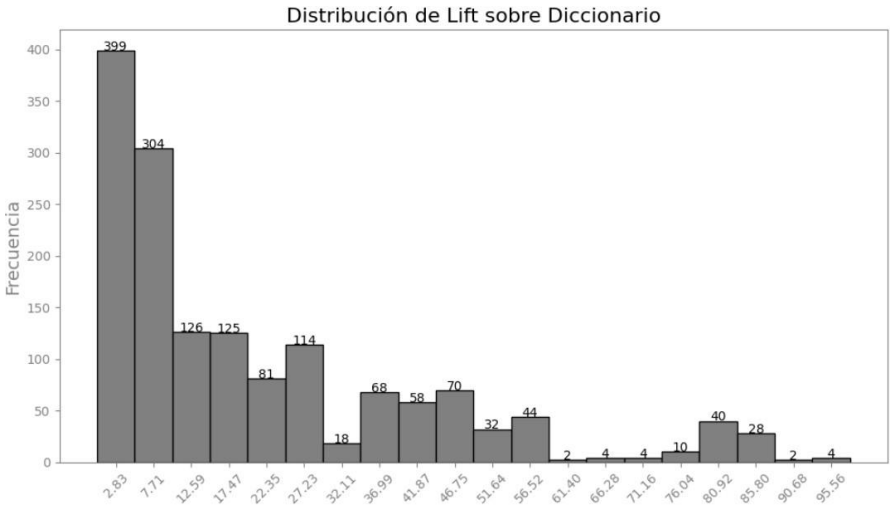


Fig. 16.2: Distribución de lift en el Diccionario de transacciones (elaboración propia).

analizados.

Proceso B. Filtros colaborativos

La implementación de los filtros colaborativos se realiza de la siguiente forma, cumpliendo las siguientes etapas: (i) entrenamiento y evaluación del sistema de recomendación; (ii) generación de recomendaciones; y (iii) evaluación de calidad de las recomendaciones:

- Se toma el Diccionario con los 56 Segmentos correspondientes a los Segmen- tos del código UNSPSC generado previamente.
- Se define una función 'get top n' que toma una lista de predicciones con un parámetro 'n' que retorna un diccionario con llaves de 'user ID' y los valores son las listas de pares de top n ítems ordenadas en orden descendente de ratings.
- Se define una función 'precision at k ' que evalúa la calidad de las recomen- daciones generadas en el top-k usando la métrica de precisión.
- Se define función 'recall at k ' que calcula la métrica de recall para filtros colaborativos.
- Se define función 'f1-score' que calcula el F-score para el filtro colaborativo.
- Se define la función 'apply recommendation system' que implementa el filtro colaborativo desde usuarios por medio de la librería Surprise en Python. La función se desagrega en los siguientes pasos: (i) Se crea un objeto 'reader' con escala de 2 a 5, con el supuesto de que ratings por debajo de 2 no se consideran relevantes; (ii) Se transforma la base en formato para la librería Surprise de Python; (iii) Se divide base en entrenamiento (80%) y test (20%); (iv) Se implementa la medida de similitud cosine; (v) Se implementa un algoritmo de KNNBasic para agrupar por vecinos cercanos; (vi) Se utiliza el método 'algo.test ' para generar predicciones en el conjunto de test; y (vii) se implementan las funciones 'get top n', 'precision at k', 'precision at k' y la función 'acuracy rmse' de la librería Suprise.

Protocolo experimental y resultados para Proceso B (Filtros Colabo- rativos)

Para incorporar un protocolo experimental, se procede a ajustar hiper- parámetros para determinar los mejores modelos a partir de los resultados de las métricas de desempeño y por tanto su utilidad relativa para la generación de alternativas de agregación de demanda o estrategias de compra para las enti- dades del Estado Colombiano. Para tal fin, y de conformidad con los algoritmos implementados, se genera una herramienta grid para emplear una batería de experimentación. Esta

batería incluye los siguientes hiper-parámetros, los cuales reflejan la naturaleza de la base y las heterogeneidades entre las entidades es- tatales que realizaron las transacciones reflejadas por la base de datos utilizada para la investigación:

1. Tamaño del vecindario (k): Para el caso de 'get top n' se tiene 'n=10', mien- tras que para 'precision at k' los valores de 'k'=5, 10, 15 y 20.
2. Métricas de similaridad: Se utiliza Consine, msd y pearson
3. Umbral de Relevancia: Para 'recall at k' el umbral inicial se fija en 3.
4. Distribución de test: Se cuenta con un tamaño del 20
5. Escala de rating: En el 'Reader ' la escala de rating se establece entre (2,5).
6. Algorimo de agrupación: Se utiliza un KNNBasic.
7. Como métricas de desempeño se calculan precisión, recall, F1-Score y RMSE.

Se utiliza F1-Score como métrica de desempeño principal por contar con una aproximación balanceada entre recall y precisión. En la Figura 8, se presenta las estadísticas de la bateria experimental. La batería experimental genera un total de 672 combinaciones, resultantes de iterar sobre las métricas de similitud, los diferentes tamaños de vecindario y las tres métricas de desempeño escogidas para cada una de las 56 llaves del Diccionario.

Para el Proceso B (Filtro Colaborativo) el ejercicio experimental permite obtener mejoras de desempeño sobre el espacio de modelos implementado, al pasar de un F1 score de 0,23 a uno de 0,35 con una mejora significativa para el caso de la partición db_fc_21 del Diccionario de transacciones. Para dicho caso, el F1 Score del filtro colaborativo es del 35,3%, lo cual representa una efectividad baja pero aceptable en cuanto a la predicción de productos que pueden ser de 'gusto' de los usuarios.

	K	Precision	Recall	F1	RMSE
count	672	672	672	672	672
mean	12.50	0.14	2.15	0.26	1.01
std	5.59	0.01	0.42	0.02	0.13
min	5.00	0.12	1.50	0.23	0.61
25%	8.75	0.13	1.89	0.24	0.96
50%	12.50	0.13	2.03	0.25	1.05
75%	16.25	0.14	2.32	0.27	1.07
max	20.00	0.19	3.57	0.35	1.29

Fig. 16.3: Estadísticas Batería de Experimentación

16.5 Conclusiones

El presente documento desarrolla una metodología para incorporar la Inteligencia Artificial a las compras del Estado colombiano por medio de recomenda- ciones para la agregación de demanda a través de procesos de depuración y tratamiento de datos, así como de los algoritmos de Reglas de Asociación y Filtros Colaborativos. Los resultados alcanzados de desempeño lift ratio y F-Score

muestran una ámplia utilidad de los procesos para las entidades compradoras en la implementación de esquemas de agregación de adquisiciones por medio de Acuerdos Marco de Precios o esquemas de agregación propios. Adicionalmente, la metodología desarrollada pueden actuar como base para desarrollar nuevas herramientas que permitan mejorar la gestión de las compras públicas en las entidades estatales.

Referencias

1. Colombia Compra Eficiente. ¿Qué es agregar demanda? Accessed: 2024- 04-25. 2022. url: https://colombiacompra.gov.co/content/que-es- agregar-demanda.
2. Departamento Nacional de Planeación (DNP). Documento Conpes 4048. Tech. rep. Departamento Nacional de Planeación, 2021. url: https:// colaboracion.dnp.gov.co/CDT/Conpes/Econ%C3%B3micos/4048.pdf.
3. Google. Google Colaboratory. https://colab.research.google.com/ notebook. Accessed: 2024-04-19.
4. IBM. CRISP-DM Help Overview. https :// www . ibm . com / docs / es / spss-modeler/saas?topic=dm-crisp-help-overview. Accessed: 2024- 04-19.

Part II
Propuestas tesis doctorales

El propósito de este compendio es proporcionar una visión general de las tendencias de investigación emergentes en las áreas de ciencias de la computación e información de estudiantes de doctorado, ofreciendo a la comunidad académica y a los profesionales del sector una referencia valiosa sobre las direcciones de la investigación avanzada. Además, estas propuestas pueden inspirar a futuros investigadores a explorar temas nuevos y desafiantes en este campo dinámico.

Cada propuesta incluye una introducción al problema de investigación, preguntas clave, objetivos, una revisión de la literatura, la metodología propuesta y una discusión sobre las posibles contribuciones al conocimiento y sus aplicaciones en diferentes áreas de la computación.

The purpose of this compendium is to provide an overview of emerging research trends in computer and information sciences from doctoral students, offering the academic community and industry professionals a valuable reference on the directions of advanced research. Additionally, these proposals can inspire future researchers to explore novel and challenging topics in this dynamic field.

Each proposal includes an introduction to the research problem, key questions, objectives, a literature review, the proposed methodology, and a discussion on potential contributions to knowledge and their applications in various areas of computing.

O propósito deste compêndio é fornecer uma visão geral das tendências de pesquisa emergentes nas áreas de ciências da computação e informação de estudantes de doutorado, oferecendo à comunidade acadêmica e aos profissionais do setor uma referência valiosa sobre as direções da pesquisa avançada. Além disso, estas propostas podem inspirar futuros pesquisadores a explorar temas novos e desafiadores neste campo dinâmico.

Cada proposta inclui uma introdução ao problema de pesquisa, perguntas-chave, objetivos, uma revisão da literatura, a metodologia proposta e uma discussão sobre as possíveis contribuições ao conhecimento e suas aplicações em diferentes áreas da computação.

17
El awareness y la interacción en un ambiente virtual de aprendizaje de enseñanza superior

Monroy, Juan [0000-0003-1757-2016]

Resumen La propuesta de investigación se centra en el modelamiento de un awareness de un ambiente virtual de aprendizaje que fomente la interacción de los participantes, tanto con los recursos dispuestos en el entorno, como con los otros integrantes. Se centra en la interacción del participante a partir de la influencia que tenga tanto la conciencia sobre el trabajo colaborativo (group awareness) como la conciencia de la actividad a desarrollar (activity awareness). Está enmarcada dentro de la lí-nea de investigación Interacción Persona-Ordenador (IPO) y el área temática es el awareness en entornos virtuales de aprendizaje.

Palabras clave Group awareness, Activity awareness, Ambiente virtual de aprendizaje AVA

17.1 Problema a resolver

La falta de un modelo de awareness para un ambiente virtual de aprendizaje de enseñanza superior, a partir del cual se fomente la interacción de los estudiantes, tanto con los recursos dispuestos para su formación, como con los demás participan-tes.

17.2 Justificación

Los ambientes virtuales de aprendizaje se han convertido en el espacio a través del cual se logran integrar un conjunto de elementos relevantes al proceso de formación de los estudiantes, tal y como lo menciona Boude [1]; se busca que sean espacios robustos para atender poblaciones diversas, tanto en sus necesidades, como en su características propias a su proceso de formación, de cara a una sociedad de compe-tencias globales, como se plantea en los documentos de Kommers et al, Auer et al.[2, 3]. Se empezaron a proyectar desde el surgimiento del e-learning Salvat [4] co-mo modalidad de formación a distancia y al uso de Internet en la educación supe-rior. La llegada de la pandemia generó un movimiento a nivel investigativo para abordar la migración a la formación virtual y revisar

Universidad de Lleida
Catalunya, España, C/ Jaume II, 69. E-25001 LLEIDA
e-mail: jm24@alumnes.udl.cat, e-mail: juan.monroy@unad.edu.co

a profundidad las interacciones entre los actores del proceso de formación, se han hecho estudios enfocados al im-pacto de las interacciones en espacios virtuales, desde el papel predominante del tutor como mediador Rossia et. al [5] y desde la introducción de una herramienta de visualización que permita evaluar dimensiones de la interacción Chavez [6]. También se destacan algunos artículos sobre la confianza en la interacción en un espacio colaborativo contrastado con la misma situación en espacios físicos, aspectos predic-tivos de participación con evidencias experimentales, Peña-Ayala [7] e investigacio-nes relacionadas con los recursos, herramientas y TIC aplicadas en el abordaje de los contenidos.

Ahora, si se parte del concepto de que el awareness es la percepción del grupo y de las actividades que realizan los usuarios en un espacio de trabajo colaborativo Bibbo et al. [8] y de que en las revisiones de literatura realizadas para este proyecto no se han encontrado investigaciones que involucren el awarenes con esos procesos de interacción del estudiante, se justifica iniciar una investigación para abordar el mode-lado de ese reconocimiento consciente que debe tener un estudiante, tanto de las actividades como del grupo de trabajo involucrados en su proceso de formación, y así, potenciar su interacción con los recursos dispuestos en el AVA y por ende mejo-rar su proceso de formación.

17.3 Pregunta de investigación

¿Cómo fomentar la interacción en un ambiente virtual de aprendizaje de enseñanza superior a partir del awareness de sus participantes?

17.4 Aportes esperados

Se espera lograr un modelo de awareness de un ambiente virtual de aprendizaje que favorezca la interacción en ambientes virtuales de aprendizaje de enseñanza supe-rior, en procesos de educación superior. Los resultados de dicha investigación se pro-yectan socializar y consolidar a través de: Revisión sistemática estado del arte, Cinco artículos publicados en revistas indexadas y ponencias

17.5 Metodología

Investigación con enfoque mixto, por una parte, para establecer la influencia del awareness en la interacción del estudiante en el AVA, y por la otra, para validar el reflejo de esa influencia en los procesos de interacción; de tipo aplicada y con alcan-ce exploratorio. Se plantea realizar a través de las siguientes fases: 1) Elaborar el estado del arte sobre awareness e interacción en AVA, 2) Caracterizar el awareness 3) Establecer el nivel de incidencia del awareness, 4) Modelar y validar el awareness. La investigación se basará en las interacciones de los estudiantes con los recursos y de-más participantes del grupo dentro de un AVA donde se implemente el modelo. La muestra se definirá una vez se establezca el entorno y la población participante y su análisis se realizará con software especializado de acuerdo a las características de la información recopilada

Referencias

1. Boude Figueredo, O. (2011). Pediatic: desarrollo de competencias en TIC a través del aprendizaje por proyectos. Educación Médica Superior, 25(2), 116-124. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-21412011000200009&lng=es&nrm=iso
2. Kommers, P., Smyrnova-Trybulska, E., & Morze, N. (2019). Universities in the Networked Society Cultural Diversity and Digital Competences in Learning Communities Introduction to This Book. Universities in the Networked Society, 10. https://doi.org/10.1007/978-3-030-05026-9
3. Auer, M. E., Hortsch, H., & Sethakul, P. (Eds.). (2020). The Impact of the 4th Industrial Revolution on Engineering Education: Proceedings of the 22nd International Conference on Interactive Collaborative Learning (ICL2019)–Volume 2 (Vol. 1135). Springer Nature. https://doi.org/10.1007/978-3-030-40274-7
4. Salvat, B. G. (2018). La evolución del e-learning: del aula virtual a la red. RIED. Revista Iberoamericana de Educación a Distancia, 21(2). https://www.redalyc.org/journal/3314/331455826005/331455826005.pdf
5. Rossia, P. G., Ranierib, M., Lic, Y., & Perifanoud, M. (2019). Interaction, feedback and active learning: where we are and where we want to go Interazione, feedback
6. Chávez, J., & Romero, M. (2014). The relationship between group awareness and participation in a computer-supported collaborative environment. In Learning Technology for Education in Cloud. MOOC and Big Data: Third International Workshop, LTEC 2014, Santiago, Chile, September 2-5, 2014. Proceedings 3 (pp. 82-94). Springer International Publishing.
7. Peña-Ayala, A. (Ed.). (2019). Educational networking: A novel discipline for improved learning based on social networks. Springer Nature. https://app.dimensions.ai/details/publication/pub.1122410761
8. Bibbo, L. M., Giandini, R. S., & Pons, C. (2016, December). Sistemas colaborativos con awareness: requisitos para su modelado. In Simposio Argentino de Ingeniería de Software (ASSE 2016)-JAIIO 45 (Tres de Febrero, 2016).

18
Modelo para evaluar la calidad del awareness en ambientes virtuales de aprendizaje en educación superior

Pilar Alexandra Moreno[0000-0002-6990-6105]

Resumen La línea de investigación de la propuesta es Human-Computer Interaction (HCI) y el área temática es la calidad del awareness en ambientes virtuales de aprendiza-je. Con el proyecto se busca desarrollar y aplicar un modelo para evaluar la calidad del awareness en ambientes virtuales de aprendizaje para educación superior partiendo de la premisa que la evaluación de la calidad del awareness ofrecerá elementos para que la calidad de los ambientes virtuales de aprendizaje se pueda garantizar. Corresponde a la propuesta de tesis doctoral que se encuentra en su etapa inicial y revisión de estado de arte.

Palabras clave Awareness, calidad awareness, ambiente virtual aprendizaje.

18.1 Problema

El problema se orienta hacia la necesidad de evaluar la calidad del awareness de un ambiente virtual de aprendizaje (AVA) para educación superior porque se espera que un buen awareness ayude en la experiencia del estudiante y, por tanto, ayude a identificar si el AVA está siendo útil para los estudiantes, si están alcanzando los objetivos.

18.2 Justificación

Según Bananthy [1] "Los ambientes de aprendizaje están compuestos de componen-tes interrelacionados e integrados que interactúan entre sí". Para Davidson [2] los elementos del sistema AVA son la comunidad virtual de aprendizaje, la infraestructura administrativa y la infraestructura técnica, allí interactúan instructores y estudian-tes, por tanto, es importante analizar el awareness como conciencia de grupo para que los estudiantes puedan ser conscientes de lo que está ocurriendo alrededor de su contexto de interacción, como un factor fundamental para su motivación y "engagement". Hasta el momento, hay estudios que miden la calidad del AVA, se han identificado criterios que permitan evaluar su pertinencia y efectividad, encontrando varios modelos algunos

Universidad de Lleida, C/ Jaume II, 69. E-25001 LLEIDA. Catalunya, España.
e-mail: pilar.moreno@unad.edu.co

orientados hacia la medición del diseño de experiencia de usuario UX en Zuo [3] “User Centered Design Model, Frameworks de evaluación de UX” y otros orientados a la medición de calidad del AVA en Kasote [4] “Modelo de calidad de aprendizaje en línea, Metodologías de evaluación de AVA, Framework de diseño instruccional”; pero no se han encontrado trabajos formales que definan la calidad de awareness ni modelos de calidad que permita medir el awareness en AVA.

Así, este proyecto de investigación resulta valioso ya que es necesario primero, definir “calidad del awareness de un ambiente virtual” para luego diseñar un modelo que permita su evaluación como ayuda para la validación de la experiencia del estu-diante en el AVA. Y en este proceso, se debe caracterizar el contexto, analizar crite-rios de evaluación del awareness dentro del sistema, identificando tecnologías, ele-mentos visuales, auditivos, recursos para los estudiantes, valorando aspectos relacio-nados con la sobrecarga cognitiva que pueda presentarse.

El término calidad se define de acuerdo con Crosby [5] como “conformance to re-quirements”, es decir, significa “conformidad con los requisitos o cumplimiento de los requisitos”. En cuanto a evaluación de la calidad en AVA está el trabajo citado por Moore y Kreasley [6] y realizado por el Institute for Higher Education and Policy [7] donde se resalta “es importante que a lo largo de la duración del curso, los estudian-tes tengan acceso a un ambiente virtual apropiado, con herramientas adecuadas, con asistencia e instrucciones sobre los medios electrónicos utilizados”. Ahora, tenien-do en cuenta el campo que ocupa esta investigación, HCI, es necesaria la definición de “awareness”, tomando a Endsley [8] “knowing what is going on”, interpretándose como el conocimiento que se ofrece al usuario de un sistema sobre lo que ocurre y hacen los otros usuarios, con quienes interactúa en el mismo sistema. Para validarla, se resalta lo citado por Garrido [9] quien orienta la visión de awareness con la defini-ción de Dourish [10] “awareness involucra conocer quién está alrededor, qué activi-dades ocurren, quién habla con quién; esto proporciona un punto de vista de otro usuario en el entorno diario de trabajo”.

18.3 Pregunta de investigación

¿Cómo evaluar la calidad del awareness de AVA? Así, la hipótesis inicial a comprobar con el trabajo es “La calidad del awareness de un AVA podría influir en el proceso de formación del estudiante para alcanzar los objetivos de aprendizaje”

18.4 Metodología

El tipo de investigación siguiendo a Sampieri [11] es aplicada, con diseño no experimental y alcance exploratorio bajo un enfoque mixto según Tashakkori [12] , cualitativo para definir criterios de calidad a medir en AVA; y cuantitativo, porque se procesan datos en la validación del modelo. Fase 1- Estado del arte que soporte el desarrollo del modelo, Fase 2- Definición de datos y criterios que el modelo de calidad deberá evaluar, proceso ETL. Fase 3-Creación del modelo y Fase 4-Validación del modelo en donde se prueba y evalúa para su respectivo afinamiento y diseño final.

18.5 Aportes esperados

Como resultado del proyecto se propone un modelo de medición de calidad del awareness en AVA, además, se espera contribuir con la definición de los criterios de evaluación y los datos que permiten medir la calidad del awareness. Propuesta, publicación y validación del modelo.

Referencias

1. Bananthy, B.B. (1987). Instructional Systems Design. In Davidson-Shivers, G., y Rasmessen, K. (2006). Web Based Learning: Design, Implementation and Evaluation (p. 19). Upper Saddle River, New Jersey: Pearson Education.
2. Davidson-Shivers, G., y Rasmessen, K. (2006). Web Based Learning: Design, Implememntation and Evaluation. (p.20). Upper Saddle River, New Jersey: Pearson Education.
3. Zuo, Y., Zhang, Y., & Wang, Y. (2023). User Crosby, P.B. (1979). Quality is Free: The Art of Making Quality Certain Mentor book. McGraw-Hill, New York 1979. ISBN, 0070145121, 9780070145122.
4. Kasote, S., & Vijayaraghavan, K. (2023). Human-Computer Interaction: Human Values and Quality of Life. [PDF].
5. Crosby, P.B. (1979). Quality is Free: The Art of Making Quality Certain Mentor book. McGraw-Hill, New York 1979. ISBN, 0070145121, 9780070145122.
6. Moore, M., y Kearsley, G. (2005). Distance Education: A Systems View (2nd ed.). Belmont, California. Thomson Learning. (p. 203).
7. Institute for el Institute for Higher Education and Policy. (2002). Benchmarks for Success in Internet-Based Distance Education. In Moore, M., y Kearsley, G. 2005. Distance Education: A Systems View (2nd ed.) (pp. 203). Belmont, California. Thomson Learning.
8. Endsley, M. (1995). Towards a Theory of Situation Awareness in Dinamic Systems. Human Factors 37(1), pp. 32-64.
9. Garrido, J. (2014). Solución Integral para Mejorar la Colaboración y el Awareness en Entornos Healthcare Ubicuos y Sensibles al Contexto. Tesis Doctoral. Universidad de Castilla - La Mancha. Departamento de Sistemas Informáticos. (p. 38).
10. Dourish, P y Bly, S. (1992). Portholes: supporting awareness in a distributed work group. P. Bauersfeld, J. Bennett, and G. Lynch (eds.): CHI'92 Conference Proceedings: ACM Conference on Human Factors in Computing Systems, 3–7, Monterey, California. New York: ACM Press, pp. 541-547.
11. Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, M. P. (2015). Metodología de la investigación (6.ª ed.). McGraw-Hill.
12. Tashakkory, A. y Teddlie, C. (1998). Mixed methodology; Combining qualitative and quantitative approaches. (pp 19). London, Sage Publications.

19

Identificación de Factores Determinantes en el Aprendizaje Aplicables en Contextos Virtuales y Medibles con Técnicas no Invasivas

Vargas García, Andrés[1][0009-0005-8742-5929]
Chavarro Porras, Julio Cesar[2][0000-0001-8876-8855]

La educación es esencial para el desarrollo de una sociedad y su progreso económico, social y cultural, mejorar su calidad proporciona a los estudiantes herramientas para enfrentar desafíos, desarrollar habilidades y ser productivos; además, fomenta ciudadanos responsables y críticos. Las TIC han transformado la educación al permitir formas de aprendizaje interactivas y colaborativas, facilitando el acceso a información y recursos en línea, y promoviendo el aprendizaje autónomo [2]. Ejemplos de esta transformación son la educación en línea, los recursos educativos abiertos y las plataformas digitales.

Aunque la educación en línea ofrece numerosas ventajas y oportunidades, también presenta algunos desafíos y problemas, entre ellos encontramos aspectos como la motivación, la atención y la memoria, aspectos que el docente requiere validar durante el proceso de aprendizaje y que son difíciles de medir en ambientes mediados por TIC, es por esto, que surge la pregunta ¿Es posible mediante la medición de la interacción del usuario frente a objetos virtuales de aprendizaje validar el proceso enseñanza aprendizaje usando métodos no invasivos? Nuestro objeto de estudio se centra en la identificación de factores que afectan el proceso de aprendizaje y que pueden ser medidos de manera no invasiva cuando este proceso se realiza en ambientes virtuales de aprendizaje.

La psicólogia define y concibe el aprendizaje como [1]: “El aprendizaje es un cambio relativamente permanente en las asociaciones o representaciones mentales como resultado de la experiencia”. Desde el punto de vista fisiológico el aprendizaje se da cuando se producen interconexiones entre las neuronas, en el fortalecimiento de la sinapsis ya existente y en la creación de nuevas. [2, 3, 4, 5]. Para esto se basa en los procesos cognitivos, que son las operaciones mentales que realiza el cerebro para procesar la información. Estos procesos permiten la adquisición, almacenamiento, manipulación y utilización de la información que recibimos del entorno.

La integración de las neurociencias con la educación ha permitido generar conocimientos fundamentales sobre las bases neurales del proceso de aprendizaje, investigaciones actuales posibilitaron mayores conocimientos sobre los procesos cognitivos involucrados en el aprendizaje a través del estudio de funciones cerebrales superiores y complejas, como lo son [6]: lenguaje, memoria, atención y motivación. A pesar que los OVA han buscado mejorar en aspectos estéticos, metodológicos y de contenido, el cumplimiento de estas características por parte de un entorno virtual de aprendizaje no garantiza que se dé de manera óptima el proceso enseñanza-aprendizaje, es por esta razón, que

Universidad Tecnológica de Pereira. Pereira, Risaralda, Colombia.
e-mail: avargas@utp.edu.co[1] e-mail: jchavar@utp.edu.co[2]

se han estudiado mecanismos para evaluar aspectos afectivos enfatizando en aspectos emocionales.

En [7] se proporciona un marco actualizado de factores humanos que influyen en la eficacia de los Entornos Virtuales de Aprendizaje, acá, los métodos para muestrear las emociones se dividen en tres categorías, la primera por medio de cuestionarios que se realizan durante el curso, esto toma instantáneas de las emociones, este método tiene limitantes debido a la intrusión en el proceso de aprendizaje, la segunda categoría consiste en un cuestionario una vez finalizado el proceso de aprendizaje, esta forma tiene como limitante que las emociones asociadas con el aprendizaje son de corta duración y como las emociones se encuestan después de finalizar el proceso de aprendizaje, no refleja las emociones reales que se vivieron durante el proceso.

Investigadores se han centrado en herramientas de tipo cuestionario para medir las emociones durante el proceso de enseñanza de estudiantes frente a OVA como: UEQ[8], AttrakDiff [9], SAM (Self-Assessment Manikin), [10] LEM-Tool [11], entre otras, Estos cuestionarios utilizan escalas de emociones que pueden resultar confusas para las personas que participan en las evaluaciones.

la tercera categoría detecta emociones durante el proceso de aprendizaje mediante métodos no intrusivos como la detección de expresiones faciales, detección de miradas, análisis de datos de textos generados, esta categoría contrarresta las limitaciones de las otras dos, pero hay que solucionar el problema de cómo capturar, analizar y correlacionar las expresiones faciales, medidas corporales, variables físicas y características de la voz con las emociones y estas a su vez con la efectividad del proceso enseñanza aprendizaje, esto sin usar métodos intrusivos que afecten o entorpezcan el proceso de aprendizaje.

El método propuesto en nuestra investigación se compone por un sistema de detección y procesamiento de imágenes fijas y en movimiento, identificando emociones durante el proceso de interacción del estudiante con el OVA en tiempo real. El sistema combina técnicas de procesamiento de imágenes y aprendizaje automático para analizar y clasificar el contenido visual, con los resultados del procesamiento de imágenes y los resultados del proceso del aprendizaje se establece la correlación entre estas variables y así determinar procesos cognitivos como la atención, la motivación y la memoria influyen en el proceso y la calidad del aprendizaje, de tal forma que se puedan realizar ajustes a las OVA.

Se planea la ejecución de un experimento con usuarios reales el cual consiste en tomar uno o varios OVA, los cuales se han validado y cumplen estándares establecidos de calidad evaluados siguiendo metodologías de evaluación de objetos de aprendizaje como: LORI [12], COdA [13] y ECOBA [14], allí se aplicará a los estudiantes seleccionados y se cruzaran los resultados obtenidos del sistema de detección, procesamiento de imágenes y predicción de resultados con los resultados obtenidos por los estudiante del proceso de aprendizaje, de tal forma que validemos el objetivo propuesto.

Este artículo aborda la selección de las variables con un enfoque cualitativo y se centra en identificar las variables que deben ser medidas en ambientes virtuales de aprendizaje, durante una sesión de trabajo de un estudiante que interactúa con el OVA a través de la plataforma.
Esto permitirá tanto a docentes como a estudiantes información relevante de su proceso, hacer ajustes al contenido, establecer mecanismos para intervenir y modificar situaciones que afectan el aprendizaje del alumno, todo esto sin la necesidad de estar docente y estudiante frente a frente.

Referencias

1. Ormrod, J.E., et al., Aprendizaje humano. Vol. 4. 2005: Pearson Educación Madrid, Spain.

2. Byrnes, J.P. and N.A. Fox, The educational relevance of research in cognitive neuroscience. Educational Psychology Review, 1998. 10: p. 297-342.
3. Greenough, W.T., J.E. Black, and C.S. Wallace, Experience and brain development. Child development, 1987: p. 539-559.
4. Merzenich, M.M., Cortical plasticity contributing to child development, in Mechanisms of cognitive development. 2001, Psychology Press. p. 79-108.
5. Rosenzweig, M.R., Multiple models of memory. The brain, cognition, and education, 1986: p. 347-371.
6. Izaguirre, M., Neuroproceso de la enseñanza y del aprendizaje. 2017: Alpha Editorial.
7. Syed, M., et al., Implicit and Explicit Emotions in MOOCs. International Educational Data Mining Society, 2019.
8. Laugwitz, B., T. Held, and M. Schrepp. Construction and evaluation of a user experience questionnaire. in HCI and Usability for Education and Work: 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society, USAB 2008, Graz, Austria, November 20-21, 2008. Proceedings 4. 2008. Springer.
9. Hassenzahl, M., M. Burmester, and F. Koller. AttrakDiff: A questionnaire to measure perceived hedonic and pragmatic quality. in Mensch & computer. 2003. Springer Heifelberg.
10. Bradley, M.M. and P.J. Lang, Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry, 1994. 25(1): p. 49-59.
11. Huisman, G., et al. LEMtool: measuring emotions in visual interfaces. in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013.
12. Vargo, J., et al., Learning object evaluation: computer-mediated collaboration and interrater reliability. International Journal of Computers and Applications, 2003. 25(3): p. 198-205.
13. Fernández, P., R. Domínguez, and R. Armas, Herramienta COdA de Evaluación de la Calidad de Objetos de Aprendizaje, desarrollada en el marco de los Proyectos de Innovación y Mejora de la Calidad de la Docencia. 2012, PIMCD 268/2010-2011 y PIMCD 236/2011-2012 financiados por el Vicerrectorado.
14. Ruiz González, R., J. Muñoz Arteaga, and F. Álvarez Rodríguez. Formato para la Determinación de la Calidad en los Objetos de Aprendizaje. in Primera Conferencia Latinoamericana de Objetos de Aprendizaje, LACLO. Coordinan: Repositorio de Conocimiento Europeo (ARIADNE), Corporación Latinoamericana de Redes Avanzadas (CLARA). Guayaquil, Ecuador. 2006.

20

Aplicación de Aprendizaje por Refuerzo y Series de Tiempo para el Análisis del Pronostico del Inventario Cooperativo de Medicamentos Asociados al Tratamiento de Enfermedades de Alto Costo

Wilmar Guillermo Rodriguez Lopez[1]
Marco Javier Suarez Barón[2]
Oscar Javier García Cabrejo[3]

Resumen La gestión de inventarios en cadenas de suministro medico permite afrontar retos como dificultad para adquirir, administrar y gestionar sus suministros, siendo po-sible gracias a esta incrementar el bienestar social, reducir la tasa de mortalidad y garantizar la disponibilidad de los medicamentos. Siendo la incertidumbre de la demanda de los medicamentos una de las mayores problemáticas en la gestión de las instituciones médicas, generando un impacto negativo en los costos, disponibilidad, gestión, logística e inventarios. Por lo cual, la presente investigación pro-pone el desarrollo de un modelo predictivo basado en modelos de programación profunda de aprendizaje por refuerzo y un novedoso modelo de inventarios de-nominado Inventario Cooperativo.

Palabras clave Forecasting, Reinforce Learning, Inventories, Medical Supply Chain.

20.1 Problema de Investigación

En la actualidad la tendencia global indica que la confiabilidad es un requerimiento innegociable en cualquier cadena de suministro [1, 2]. No obstante, la cadena de suministro de medicamentos presenta dificultades en la previsión de la demanda, generando un impacto negativo en los costos, la disponibilidad, la gestión, la logística y los inventarios [3, 4]. De forma similar, algunos autores han identificado brechas de acceso en la adquisición de medicamentos debido a su poca disponibilidad, proble-mas de suministro, impacto económico, limitaciones en la prestación del servicio de salud y condiciones geográficas [5], [6]. En particular en Colombia la Cuenta de Alto Costo (CAC) se encarga de los registros de gestión de patologías como: artritis reuma-toide, hemofilia, enfermedad renal crónica, VIH, cáncer y hepatitis C [7]. Siendo la gestión de estas enfermedades un tema de interés nacional debido a su forma de gestión [8].

Por otra parte, al indagar en la literatura se encuentra que la aplicación de modelos de pronósticos a medicamentos en el sector medico es mínima, y que los modelos y técnicas de modelado mejoran su eficiencia al emplear herramientas basadas en aplicaciones de Machine Learninig [9, 10, 11]. Esto a causa de considerar que los modelos de series de tiempo por si solos producen predicciones

[1,2,3]Universidad Pedagógica y Tecnológica de Colombia – UPTC
e-mail: wilmar.rodriguez@uptc.edu.co[1]
e-mail: marco.suarez@uptc.edu.co[2]
e-mail: oscar.garcia04@uptc.edu.co[3]

erradas debido a con-diciones de ruido y falta de linealidad de los datos, volviendo con ello más compleja su determinación [12].

20.2 Hipotesis y Pregunta de Investigación

Tomando en cuenta lo especificado previamente, se propone como pregunta de investigación la siguiente: ¿Cómo gestionar la predicción demanda de inventario cooperativo de los medicamentos asociados al tratamiento de las enfermedades de alto costo mediante la aplicación de aprendizaje por refuerzo y series de tiempo?

A partir de la cual se plantean hipótesis para su desarrollo teniendo en cuenta co-mo foco principal las métricas encontradas en la literatura asociadas a la evaluación de modelos similares en cuanto al tipo de variables y problemáticas, siendo conside-rada para la hipótesis el MAPE (Mean Absolute Percentage Error - MAPE).

Ho: El MAPE generado por el modelo de pronóstico propuesto es menor o igual al 5Ha: El MAPE generado por el modelo de pronóstico propuesto es mayor al 5

20.3 Contribuciones o aportes esperados de la investigación

A continuación, se relacionan las contribuciones u aportes esperados como fruto del desarrollo del trabajo doctoral:

- Estado del arte de los modelos de pronóstico (forecasting)
- Estado del arte de los modelos de inventarios existentes
- Arquitectura lógica de los modelos propuestos
- Modelo de pronóstico para la predicción del inventario cooperativo del medicamento relacionados con las enfermedades de alto costo
- Modelo de visualización
- Validación del modelo de pronóstico

20.4 Plan de Evaluación de los Resultados

De acuerdo con los objetivos planteados en la propuesta doctoral, se han considera-do la generación de un artículo de revisión bibliográfica. Así mismo, para los resultados relacionados con el modelo computacional, arquitectura lógica, algoritmo de indexación y validación del modelo, se propondrá la publicación de artículos científicos que evidencien los resultados obtenidos. Finalmente, en cuanto al modelo de visualización se deberá diseñar y desarrollar un modelo de visualización del modelo de pronóstico propuesto con una interfaz llamativa y de fácil operación. Así mismo los resultados serán compartidos con comunidades de aprendizaje orientadas al ma-nejo de los pronósticos como Forecast Social Media y en espacios pre-print de revis-tas indexadas isi-Scopus, a fin de recibir retroalimentación.

Referencias

1. D. Salite, “Traditional prediction of drought under weather and cli-mate uncertainty: analyzing the challenges and opportunities for small-scale farmers in Gaza province, southern region of Mozam-bique,” Natural Hazards, vol. 96, no. 3, pp. 1289–1309, Apr. 2019, doi: 10.1007/s11069-019-03613-4.
2. V. G. Kossobokov and A. A. Soloviev, “Testing Earthquake Predic-tion Algorithms,” Journal of the Geological Society of India, vol. 97, no. 12, pp. 1514–1519, Dec. 2021, doi: 10.1007/s12594-021-1907-8.
3. A. Dixit, S. Routroy, and S. K. Dubey, “A systematic literature review of healthcare supply chain and implications of future research,” Int J Pharm Healthc Mark, vol. 13, no. 4, pp. 405–435, 2019, doi: 10.1108/IJPHM-05-2018-0028.
4. G. Stecca, I. Baffo, and T. Kaihara, “Design and operation of strate-gic inventory control system for drug delivery in healthcare industry”, IFAC-PapersOnLine, vol. 49, no. 12, pp. 904–909, 2016, doi: 10.1016/j.ifacol.2016.07.890.
5. S. V. Rajkumar, “The high cost of prescription drugs: causes and so-lutions,” Blood Cancer J, vol. 10, no. 6, 2020, doi: 10.1038/s41408-020-0338-x.
6. P. A. H. O. PAHO, Access to High-Cost Medicines in the Americas: Sit-uation, Challenges and Perspectives, no. 1. 2010. [Online]. Availa-ble: https://www.paho.org/hq/dmdocuments/2010/High-cost-Med-Tech-Series-No-1-Sep-15-10.pdf
7. C.-C. de A. Costo, “Cuenta de Alto Costo - CAC.” Accessed: Jan. 05, 2023. [Online]. Available: https://cuentadealtocosto.org/site/quienes-somos/
8. O. Moreno Mattar, “Financiamiento de medicamentos de alto costo en Colombia.” Accessed: Jan. 04, 2023. [Online]. Available: https://www.neuroeconomix.com/es/financiamiento-de-medicamentos-de-alto-costo-en-colombia/
9. Z. Kang, C. Catal, and B. Tekinerdogan, “Machine learning applica-tions in production lines: A systematic literature review,” Comput Ind Eng, vol. 149, no. August, p. 106773, 2020, doi: 10.1016/j.cie.2020.106773.
10. S. Pouyanfar et al., “A Survey on Deep Learning: Algorithms , Tech-niques ,” ACM Comput Surv, vol. 51, no. 5, pp. 1–36, 2018.
11. J. Bedi and D. Toshniwal, “Deep learning framework to forecast elec-tricity demand,” Appl Energy, vol. 238, no. July 2018, pp. 1312–1326, 2019, doi: 10.1016/j.apenergy.2019.01.113.
12. V. Ramos, W. Yamaka, B. Alorda, and S. Sriboonchitta, “High-frequency forecasting from mobile devices’ bigdata: an application to tourism destinations’ crowdedness,” International Journal of Con-temporary Hospitality Management, vol. 33, no. 6, pp. 1977–2000, 2021, doi: 10.1108/IJCHM-10-2020-1170.

21
Modelo computacional para determinar el nivel adecuado de madurez en cosecha del aguacate Hass

Carlos Meneses [1] [0000-0002-8192-6889]
Julio Chavarro[2] [0000-0001-8876-8855]

Resumen The quality of Hass avocado exports depends largely on achieving optimal ripeness at harvest, a task that is difficult to achieve visually. Non-invasive technological tools using image analysis have shown promise in addressing this problem, although their effectiveness and computational characteristics remain uncertain. This study presents a computational model aimed at accurately determining the appropriate level of Hass av-ocado harvest maturity through digital image processing. Image processing contributes significantly by providing an accurate, objective and con-sistent method for assessing maturity, reducing human error and supporting the overall quality control process.

Palabras clave computational model, level of maturity, Hass avocado

21.1 Problema

21.1.1 Justificación

Según informe en la revista Semana [1] , el cultivo de aguacate Hass es reciente en Colombia, apenas hace aproximadamente 2 décadas iniciativas particulares iniciaron con proyectos productivos de este fruto y recientemente se ha venido transformando en un pilar de desarrollo econó-mico y social del país [2], [3] donde uno de los propósitos principales es la participación en mercados internacionales. Las exportaciones de aguacate Hass superan los 200 millones de dó-lares en 2023 y Colombia se ha convertido en el tercer exportador mundial según cifras de la Asociación Nacional de Comercio Exterior (Analdex) [4]. El panorama es promisorio para la industria aguacatera, sin embargo, la investigación y la adopción de tecnologías para el adecuado acompañamiento técnico del cultivo no crece en la misma proporción [5]. El aguacate es un fruto climatérico donde el nivel de madurez para consumo humano no se pro-duce en el árbol sino después de la cosecha [6], por esta razón es importante conocer el mo-mento óptimo para cosechar la fruta. Determinar las características físicas del fruto de aguacate Hass para saber que está en su punto óptimo en el que se debe cosechar, se ha convertido en un reto debido a que no es fácil a simple vista determinarlas [7], [8]. Los mercados internacionales requieren frutos de calidad y durabilidad del producto en el

[1,2] Universidad Tecnológica de Pereira, Pereira, Colombia
e-mail: cmeneses@utp.edu.co

transporte. De acuerdo con Cerdas Araya et Al [9], si un aguacate se cosecha antes de su punto óptimo para exportar, va tener en el mercado un proceso de maduración prematura, lo cual según Gamble et Al [10] altera su sabor y textura pues al no alcanzar el porcentaje de materia seca adecuada, se presentan defec-tos internos que inciden en la percepción del consumidor. De otro lado, si permanece dema-siado en el árbol, aumentan los problemas fitosanitarios y su vida útil se verá reducida. Para medir el grado de madurez y calidad del fruto se emplean técnicas invasivas como el con-tenido de materia seca (CMS), el contenido de aceite (CA) y el contenido de clorofila en la cáscara [11]. Las técnicas analíticas basadas en análisis de imágenes ofrecen ventajas en la de-tección (no invasiva) de los parámetros de madurez y de calidad del fruto del aguacate [12]. Un estudio realizado por Herrera et Al. [13]. permite conocer indicadores preliminaries de madurez fisiológica y comportamiento de los frutos durante el proceso de maduración pos-cosecha, basados en indicadores visuales de madurez como la luminosidad, color y matiz (tonal-idad) de la cascara y el tamaño, entre otros. La clasificación de frutos para exportación se lleva a cabo mediante el aprendizaje automático (Machine Learning) y su objetivo es estudiar y construir algoritmos que pueden aprender y hacer predicciones sobre los datos recolectados. Algunos estudios como el indicado en [14], señalan que los algoritmos más comunes son redes neuronales, regresión no lineal, agrupación jerárquica, árbol de decisión, lógica difusa, análisis discriminante lineal y procesamiento digital de imágenes, entre otros. En la mayoría de soluciones propuestas a partir de imágenes digitales, no específican las técnicas utilizadas y en otras, no se indica el modelo computacional subyacente. La determinación del estado óptimo de madurez del aguacate Hass en el momento de la cosecha se traduce en una mayor calidad del fruto y en ganancias económicas para el productor. Por ello, la aplicación de tecnologías automatizadas no invasivas, como el procesamiento digital de imá-genes, permitiría al productor ofrecer un producto de alta calidad que garantice al consumidor un alimento con los nutrientes ideales.

21.1.2 Definición del Problema

Se identifica la carencia de modelos computacionales con procesamiento digital de imágenes para determinar el grado de madurez en cosecha de un aguacate Hass.

21.2 Hipótesis

Es posible desarrollar un modelo computacional basado en técnicas de procesamiento digital de imágenes que permita determinar con precisión el nivel adecuado de madu-rez de cosecha del aguacate Hass.

21.3 Contribuciones o aportes esperados

Los aportes esperados con este trabajo de investigación son: - Modelo computacional desarrollado. - Socialización con ponencias en eventos nacionales e internacionales: o XVIII Congreso Colombiano de Computación. o JCC 2024 Jornadas Chilenas de Computación. - Artículos científicos - Proyecto de Maestría.

21.4 Evaluación de Resultados

Para determinar si la imagen contiene el objeto de interés para su análisis y la identificación en ella de las variables de interés del modelo, se comparan las diferentes técnicas de aprendizaje de máquina midiendo su eficiencia mediante cálculos de complejidad computacional. La eficacia del modelo se mide comparando la detección del grado de madurez del fruto cal-culado con distintas técnicas utilizadas en el modelo propuesto, frente a la línea base obtenida a partir de pruebas fisico-químicas como porcentaje de materia seca, dureza y nivel de clorofila que sirven para medir el grado de madurez de un fruto.

Referencias

1. R. Uribe, "Cartama, casi dos décadas dedicadas al aguacate Hass", Semana, 2 May 2018.
2. J. C. Diaz Vasquez, C. Ardila López y M. A. Guerra Aranguren , "Estudio de caso sobre la admisibilidad del aguacate Hass colombiano en el mercado estadounidense: oportunidades en el Este de Asia", Revista Mundo Asia Pacífico, vol. 8, nº 14, 2019.
3. Minagricultura, "Memorias al congreso de la República 2018-2019", 2017.
4. Analdex - Asociación Nacional de Comercio Exterior, "Analdex - Retos de producción, logísticos y climáticos, los principales dolores de cabeza para las exportaciones del aguacate Hass", [En línea]. Available: https://www.analdex.org/2024/05/29/retos-de-produccion-logisticos-y-climaticos-los-principales-dolores-de-cabeza-para-las-exportaciones-del-aguacate-hass/. [Último acceso: 15 jul 2024].
5. A. Hurtado Salazar y J. S. Arias Garcia, "Biología reproductiva y alternancia del aguacate Hass en el trópico andino de Caldas", Manizales, 2023.
6. V. Hershkovitz, H. Friedman, E. E. Goldschmidt, O. Feygenberg y E. Pesis, "Induction of ethylene in avocado fruit in response to chilling stress on tree", Journal Plant Phisiology, pp. 1855-1862, nov 2009.
7. J. L. Rocha Arroyo, S. Salazar García y A. E. Bárcenas Ortega, "Determinación irreversible a la floración del aguacate 'Hass' en Michoacán", Revista Mexicana de Ciencias Agrícolas, vol. 1, nº 4, 2010.
8. E. López Chaveli, "Efecto de la reducción de la floración con ácido giberélico sobre la producción del aguacate", Valencia, 2020.
9. M. d. M. Cerdas Araya, M. Montero Calderon y O. Somarribas Jones, "Verificación del contenido de materia seca como indicador de cosecha para aguacate (Persa americana) cultivar Hass en zona intermedia de producción de de los Santos Costa Rica", Agronomia Costarricense [online]., vol. 38, nº 1, pp. 207-214, jan-jun 2014.
10. J. Gamble, R. Harker, S. Jaeger, A. White, C. Bava y M. Beresford, "The impact of dry matter, ripeness and internal defects on consumer perceptions of avocado quality and intentions to purchase", Postharvest Biology and Technology, vol. 57, nº 1, pp. 35-43, Jul 2010.
11. P. J. Hofman, J. Bower y A. Woolf, "Harvesting, packing, postharvest technology, transport and processing. The avocado: botany, production and uses", CABI Digital Library, pp. 489-540, 2013.
12. L. S. Magwaza y S. Tesfay, "A Review of Destructive and Non-destructive Methods for Determining Avocado Fruit Maturity", Food and Bioprocess Technology, vol. 8, nº 10, pp. 1995-2011, 01 Oct 201J. A. Herrera Gonzalez, S. Salazar Garcia, H. E. Martinez Florez y J. E. Ruiz Garcia, "Indicadores preliminares de madurez fisiológica y comportamiento poscosecha del fruto de aguacate Mendez", Fitotecnia Mexicana, vol. 40, nº 1, pp. 55-63, 2017.
14. J. F. I. N. a. U. L. Opara, "Machine learning applications to non-destructive defect detection in horticultural products", Biosystems Engineering, pp. 60-83, 2020.

22
Desarrollo de un sistema inteligente para el soporte en la toma de decisiones relacionadas con confirmación metrológica en procesos industriales

Carolina Márquez Narváez[1]
David Cuesta Frau[2]
Jorge Iván Padilla Buriticá[3]

Resumen Este trabajo propone un sistema de inteligencia artificial para mejorar la confirmación metrológica de instrumentos de medición de flujo, optimizando precisión, confiabilidad y eficiencia. Utilizará aprendizaje automático y redes neuronales para automatizar la validación y calibración, reduciendo costos operativos mediante sensores de bajo costo. El sistema será adaptable a diversos entornos industriales, minimizando tiempos de inactividad y evitando métodos intrusivos. Será desarrollado en Python con una interfaz gráfica, facilita la gestión y análisis de datos. Esta propuesta busca transformar la metrología en la Industria 4.0, garantizando precisión y trazabilidad.

Palabras clave Sistema Inteligente, Confirmación Metrológica, Toma de Decisiones, Instrumentación.

22.1 Problema

22.1.1 Problema de Investigación

La medición precisa del flujo es crucial en industrias como la hidráulica, alimentaria y energética, ya que los errores pueden generar desperdicios, aumentar costos y comprometer la seguridad [9, 7]. La confirmación metrológica es esencial para cumplir con estándares internacionales, pero los métodos tradicionales enfrentan problemas como errores humanos y altos costos de mantenimiento (ISO, 2003; Vázquez Ramírez, 2019). La Industria 4.0 introduce tecnologías como IA, aprendizaje automático e IoT para mejorar la confirmación metrológica, pero su aplicación enfrenta desafíos de integración y adaptabilidad [3]. Los métodos actuales son intrusivos y costosos, mientras que sensores no intrusivos y análisis con IA permiten mediciones más precisas, reduciendo costos y tiempos de inactividad [6]. Se propone un sistema basado en IA para evaluar, calibrar y validar mediciones de flujo de manera continua, alineado con estándares internacionales, adaptable y escalable, optimizando la metrología en la Industria 4.0.La pregunta de investigación es: ¿Cómo

Universidad Politécnica de Valencia, España[1,3]
Universidad Autónoma de Manizales, Colombia[1]
Instituto Tecnológico Metropolitano de Medellín, Colombia[2]
e-mail: carolina.marquez@upv.es,david.cuesta@upv.es
e-mail: jorge.padilla@itm.edu.co,

puede un sistema de IA mejorar la precisión y eficiencia en la confirmación metrológica de instrumentos de medición de flujo cumpliendo con estándares internacionales?

22.2 Justificación del Problema

La metrología es fundamental para garantizar la precisión y exactitud en las mediciones, lo que impacta directamente en la calidad de productos, la eficiencia de los procesos y la seguridad en diversas industrias [1, 2]. La confirmación metrológica asegura que los instrumentos cumplen con estándares internacionales como la ISO 10012 y la ISO/IEC 17025, manteniendo la trazabilidad y confiabilidad en las mediciones [4]. Sin embargo, los métodos tradicionales dependen de procesos manuales y son susceptibles a errores humanos y altos costos de mantenimiento, lo que afecta la productividad [8].

Para optimizar la confirmación metrológica, los sistemas basados en inteligencia artificial (IA) automatizan la validación y calibración, mejorando la precisión y reduciendo la incertidumbre mediante análisis continuo en tiempo real [9]. Aunque su implementación enfrenta desafíos de integración y escalabilidad en distintos entornos industriales [5], estos sistemas pueden transformar la metrología, alineándose con los principios de la Metrología 4.0, anticipando desviaciones y asegurando la conformidad con los estándares internacionales [5].

22.3 Hipótesis

Un sistema basado en inteligencia artificial puede mejorar notablemente la precisión, confiabilidad y eficiencia en la confirmación metrológica de instrumentos de medición de flujo, superando a los métodos tradicionales. Esto se logra mediante la evaluación, calibración y validación continua de las mediciones, garantizando el cumplimiento de los estándares internacionales de metrología.

22.4 Contribuciones esperadas

Contribuciones Esperadas Esta investigación busca realizar avances significativos en el campo de la metrología y la inteligencia artificial mediante las siguientes contribuciones:

- Desarrollo de un Sistema de IA para Confirmación Metrológica
- Optimización de Procesos Metrológicos
- Mejora de la Eficiencia y Reducción de Costos
- Escalabilidad y Adaptabilidad en Entornos Industriales Diversos
- Publicaciones y Difusión de Conocimientos

22.5 Resultados Parciales

Durante la investigación, se realizaron dos experimentos clave. En el primer experimento, se verificaron dos sensores de nivel, uno de bajo costo y otro de la marca Siemens, en un tanque de pruebas del Instituto Tecnológico Metropolitano (ITM). Las pruebas, realizadas en condiciones controladas y reales, mostraron que el sensor de bajo costo, calibrado con modelos de regresión

lineal, puede ofrecer una precisión comparable a los sensores comerciales más caros. En el segundo experimento, se llevó a cabo la confirmación metrológica de caudal utilizando métodos indirectos basados en normativas como ISO 17025 e ISO 10012. Se utilizó un banco de pruebas con control PID (Proporcional-Integral-Derivativo), capturando datos críticos para validar el proceso de confirmación metrológica. También se ha completado la recolección y caracterización de información técnica y administrativa, así como la identificación de normativas nacionales e internacionales relevantes. Además, se ha diseñado un esquema de reglas metrológicas basado en el conocimiento de expertos, e implementado un prototipo inicial del sistema con un motor de inferencia.

Referencias

1. BIPM. (2022, December 21). “BIPM”. Retrieved from https://www.bipm.org/en/worldwide-metrology/#metrology.
2. Brown, R. J. (2021). Measuring measurement – What is metrology and why does it matter? “Measurement”, 168(108408).
3. Dudnikov, A., Ivankova, O., Gorbenko, O., & Kelemesh, A. (2021). Effect of Vibration Treatment on Increasing the Durability of Tillage Equipment Working Bodies. “Eastern-European Journal of Enterprise Technologies”, 2(1), 110.
4. ISO International Organization for Standardization. (2003). “Sistemas de gestión de las mediciones — Requisitos para los procesos de medición y los equipos de medición”.
5. . Leonov, O. A., & Shkaruba, N. Z. (2020). The quality of products of machinery manufacturing plants largely depends on metrological assurance of measurements. “Journal of Physics: Conference Series”, 1515(032010), 1.
6. Pilsung, K., Dongil, K., & Sungzoon. (2016). Semi-supervised support vector regression based on self-training with label uncertainty: An application to virtual metrology in semiconductor manufacturing. “Expert Systems with Applications”, 51, 85-106.
7. Tsimikas, J., & Zervas, P. (2023). Flow Measurement in Energy Systems: Reducing Error and Waste. “Energy Efficiency in Industry”, 45(2), 78-92.
8. Vázquez Ramírez, V. M. (2019). “Metrología y Normalización”.
9. Zhang, S.-L., & Zhang, Y.-L. (2014). Novel robust stability criteria of neutral-type bidirectional associative memory neural networks. “International Journal of Computer Science Issues (IJCSI)”, 11(1), 10.

Part III
Propuestas tesis de maestría

Esta sección, desarrollada por estudiantes de maestría, destaca por su enfoque innovador y relevancia práctica. Cada trabajo aborda aspectos clave de la computación, como el desarrollo de metodologías y algoritmos aplicados a contextos reales. Estas investigaciones no solo ofrecen soluciones computacionales a problemas complejos, sino que integran áreas de la computación de manera efectiva, desarrollando herramientas basadas en sistemas de información. Sus contribuciones son esenciales para el avance del conocimiento, brindando una base sólida para futuras investigaciones y tecnologías emergentes.

This section, developed by master's students, stands out for its innovative approach and practical relevance. Each project tackles key aspects of computing, such as the development of methodologies and algorithms applied to real-world contexts. These studies not only provide computational solutions to complex problems but also integrate different areas of computing effectively, developing tools based on information systems. Their contributions are essential for advancing knowledge, providing a solid foundation for future research and emerging technologies.

Esta seção, desenvolvida por estudantes de mestrado, destaca-se por sua abordagem inovadora e relevância prática. Cada trabalho aborda aspectos essenciais da computação, como o desenvolvimento de metodologias e algoritmos aplicados a contextos reais. Estas pesquisas não só oferecem soluções computacionais para problemas complexos, mas também integram áreas da computação de forma eficaz, desenvolvendo ferramentas baseadas em sistemas de informação. Suas contribuições são fundamentais para o avanço do conhecimento, oferecendo uma base sólida para futuras investigações e tecnologias emergentes.

23
Code is Art, Art is Code

Nicolás Londoño[1] [0009-0002-9677-2963]
Nicolás Cardozo[2] [0000-0002-1094-9952]

Abstract Interact with code should be available to everyone, no matter their background. *SCuLPT* is a programming language and artistic ecosystem designed to challenge the *status quo* of programming languages, and foster computational thinking. *SCuLPT* disrupts standard programming interfaces proposing an intuitive 3D interface. This is done by creating a specification of programming constructs that become tangible artistic objects such that, when composed, generate computable programs and art pieces. The language provides a set of tools aimed to make art a tool for coding and programming an art form. The constructed language is a fully capable, Turing complete language that serves as a creative outlet. We perform an empirical validation of the experience with novice and expert programmers on computing specific programs. The results show programming proficiency in solving small computational programs across the two populations, and some really neat art!

Keywords Programming languages, Human-Computer Interface, Art

23.1 Introduction

The typewriter interface is the reigning interface to materialize programs into code since the early 50s. Interface design is lacking to support all users. Users are not created equally. Work in the user experience and human-interface integration show reports in which technology and its designs can discriminate against users due to several factors, such as different physical or cognitive capacities, race, and gender [2]. Moreover, programming has a steep learning curve and a high entry level [1]. Different approaches to computing are needed to lower the entry threshold and make computing more approachable to everyone. To lower the entry threshold of programming and computing, more approachable interfaces and methods are required. New fields in computing are emerging with different intents, while keeping the same interfaces [5]. To explore new methods of computing, we must look beyond current tools, into how to approach programming.

The Simple Cubic Language for Programming Tasks (*artlang*) is an artistic sandbox and programming language that allows the creation of physical sculptures that compile to valid programs. Conventional languages and their creations are intended to be executed by a computer, while the

Systems and Computing Engineering Department, Universidad de los Andes
Santafé de Bogotá, D.C, Colombia
e-mail: @unad.edu.co

aesthetic aspects of programs is left to programmers' creativity. By using physical objects to represent programming constructs, the creations in
emphartlang are focused on aesthetics not just on creating computable programs.

23.2 Objectives

The objective of *SCuLPT* is to build a fully capable programming language addressing the concerns described previously about the state of programming, computation, and programming languages. *SCuLPT* posits a paradigm that breaks existing preconceptions about programming languages and moves beyond the typewriter interface by leveraging artistic creativity as an alternative.

The success of the language is measured from two perspectives. First, the language should be simple to foster computational thinking, without the intrinsic difficulties of current programming languages. Second, the language should be flexible enough to express programs in different ways, enabling program construction as an artistic expression, with the computational behavior as a side-effect of the built 3D model.

23.3 Methodology

There are two main aspects to *SCuLPT*: (1) the physical specification of its constructs, and (2) the programming language constructs to enable computation. *SCuLPT*'s focus on usability and accessibility is displayed by just using simple distinctive 3D physical objects and shapes for each component on the language. There are no keywords or similar looking symbols. This allows users of any background to code foregoing knowledge, language or capacity barriers (e.g., lack of coding literacy, cognitive disability, cultural language bias) that arise from conventional programming languages.

SCuLPT is designed to be a Turing complete language. Specifically, *SCuLPT* is a stack-based language that uses (physical) blocks as instructions, executing operations on the stack. Code blocks are composed together to create complete programs, where each block represents a different function or operation on the stack. Physically, each block has a set of features that must be connected to other blocks in order to function correctly.

23.4 Evaluation of *SCuLPT*

We will evaluate *SCuLPT* from three perspectives. The language's accessibility and usability, its ability to foster computational thinking, and the inherent value the sculptures produced as art pieces. For said objectives we will conduct an experiment where participants will be asked to solve a set of problems using *SCuLPT*, measuring the ease of use and time taken to complete a task. The (artistic and programming) background of the participants will be taken into account to evaluate the language's accessibility and usability. After the experiment, we will evaluate the participants' performance and their feedback on the language using psychoanalytical tests following the Technology Acceptance Model [3] and Art Affinity Index [4]. The evidence collected, will enable us to perform a comparative, quantitatively and qualitatively analysis of the language's usability, and its ability to foster computational thinking via artistic expression in contrast to traditional methods to introduce programming.

References

1. Y. Bosse and M. A. Gerosa. Why is programming so difficult to learn? patterns of dificulties related to programming learning mid-stage. SIG-SOFT Softw. Eng. Notes, 41(6):1-6, 2017. ISSN: 0163-5948. https://doi.org/10.1145/3011286.3011301.
2. A. Ko. How my broken elbow made the ableism of computer programming personal. Nature, 2023. ISSN: 0028-0836. https://doi.org/10.1038/d41586-023-02885-y.
3. N. Maranguniéc and A. Graniéc. Technology acceptance model: a literature review from 1986 to 2013. Universal Access in the Information Society, 2014. https://doi.org/10.1007/s10209-014-0348-1.
4. W. Tschacher, C. Bergomi, and M. Trondle. The art affinity index (aai): an instrument to assess art relation and art knowledge. Empirical Studies of the Arts, 33(2):161-174, 2015. https://doi.org/10.1177/0276237415594709.
5. H. Yang and L. Zhang. Promoting creative computing: origin, scope, research and applications. Digital Communications and Networks, 2(2):84-91, 2016. ISSN: 2352-8648. https://www.sciencedirect.com/science/article/pii/S2352864816000043.

24
Integração de uma solução Learning Record Warehouse em um Serviço de Predição de Risco Acadêmico

Juary Costa-Rocha[1][0000-0001-7028-6524]
Cristian Cechinel[2][0000-0001-6384-409X]

Abstract A interoperabilidade semântica é crucial para a eficácia dos sistemas de Análise de Aprendizagem, atuando como o meio pelo qual dados de diversas fontes podem ser integrados e interpretados por um aplicativo central sem perder seu significado original. Entre os diversos mecanismos que promovem a interoperabilidade no Learning Analytics, destacam-se duas especificações: Caliper Analytics e Experience API (xAPI). Esses padrões fornecem as estruturas necessárias para harmonizar os fluxos de dados entre plataformas e aplicativos educacionais, criando um ecossistema coeso para análise de dados e geração de insights.

Keywords Ambientes Virtuais de Aprendizagem, Computação em Nuvem, Data Warehouse, Learning Analytics, Learning Record Store

24.1 Problema e hipótese da pesquisa

As motivações para execução deste projeto são devido à tendência da adoção do ensino digital, que pode acontecer de forma totalmente digital (Educação a Distância) ou de forma híbrida, referidos como aprendizagem combinada [4], tornando, desta forma, o ambiente de ensino digital e sistemas de gestão de aprendizagem, como parte integrante das atividades diárias de ensino. Essa tendência foi intensificada pela pandemia do COVID-19 em 2020, movendo quase toda a educação para o formato online [1], levando a um rápido crescimento no interesse de ferramentas de apoio ao ensino e aprendizagem digital [1]. Com essa adoção, obtêm-se dados em abundância, que necessitam de um sistema de armazenamento pensado para receber dados provenientes de diferentes fontes [6].

Muitas soluções de LA (Learning Analytics) são entregues na forma de SaaS (Software as a Service) [1], permitindo que a mesma solução possa ser utilizada por diferentes instituições. Ou seja, quando se trata de diversas instituições, estas também possuem diferentes tipos de dados tanto com relação ao conteúdo quanto a natureza e utilizam formas de armazenamento e transmissão dos mais variados tipos possíveis. É neste contexto que esta dissertação está inserida e busca responder a seguinte pergunta de pesquisa: Como um Learning Record Warehouse pode ser integrado a uma solução SaaS de Learning Analytics para Predição de Risco Acadêmico de modo suportar o recebimento de dados provenientes de diferentes Ambientes Virtuais de Aprendizagem e diferentes outras fontes?

Universidade Federal de Pelotas, Pelotas, Brasil
e-mail: \{jcrocha,cristian.cechinel\}@inf.ufpel.edu.br

A integração de um Learning Record Warehouse com uma solução SaaS de Learning Analytics permitirá uma predição mais precisa do risco acadêmico ao suportar a coleta e análise de dados provenientes de múltiplas fontes de forma padronizada. Essa integração pode melhorar a interoperabilidade dos sistemas, oferecendo uma visão holística do desempenho acadêmico e permitindo intervenções mais eficazes.

Para abordar essa questão de pesquisa, o objetivo geral da pesquisa é desenhar e construir a integração de um Learning Data Warehouse para receber dados de plataformas e sistemas de informação educativos e ser utilizado em serviços de predição de acadêmicos em risco.

24.2 Trabalhos relacionados

Nos últimos anos, os sistemas educacionais testemunharam um rápido crescimento na integração de tecnologias on-line como parte da experiência de aprendizado do aluno. Essa integração de tecnologias educacionais, juntamente com a coleta de dados dos alunos, levou ao aumento da aplicação de LA para a tomada de decisões orientadas por dados para melhorar o aprendizado dos alunos [3].

Nesse sentido, com a crescente adoção de tecnologias, os sistemas educacionais não estão mais isolados e acabam tendo que ter a capacidade de pegar componentes e dados educacionais desenvolvidos em um local com um conjunto de ferramentas e plataformas e usá-los em outro local com outras ferramentas ou plataformas. Quando se tem sistemas colaborativos, lidar com fontes de dados heterogêneas é um grande desafio, mas a integração de dados de fontes heterogêneas usando LA permite uma análise abrangente, o que pode resultar em uma tomada de decisão mais assertiva no campo da educação.

A integração de dados está intimamente relacionada à interoperabilidade, que envolve os níveis semântico, técnico, jurídico e organizacional [7]. Com relação à interoperabilidade técnica e semântica, há duas especificações de dados amplamente conhecidas (padrões do setor) que visam ao domínio educacional e à LA, a saber, a Experience API (xAPI) e o Caliper Analytics (Caliper) [6].

Ao adotar essas especificações, um vocabulário comum é estabelecido, o que é necessário para garantir que todas as atividades de aprendizagem entre todos os diferentes sistemas descrevam a mesma experiência. Ao formalizar o vocabulário usado, um conjunto de atributos e regras sobre os dados é estabelecido, o que determina como eles serão armazenados, recuperados e acessados por outros componentes, sistemas ou atividades [2]. Esses benefícios contribuem para um intercâmbio mais eficaz e contínuo de dados para fins analíticos. Podemos pensar na interoperabilidade como uma forma de reduzir a energia gasta na transformação e integração de dados manualmente. Nesse sentido, podemos usar especificações já estabelecidas no mercado, como xAPI ou Caliper Analytics.

No estudo realizado por [8], a especificação Caliper Analytics é discutida para oferecer um modelo de informações extensível, vocabulários controlados e uma API para registrar eventos de aprendizagem. Entretanto, reconhece-se que muitas atividades de aprendizagem ainda não foram modeladas pelo grupo de trabalho do Caliper. Para resolver essa lacuna, o documento propõe o envolvimento da comunidade SoLAR para contribuir com o desenvolvimento de novos perfis de atividades de aprendizagem dentro da especificação Caliper.

O [6] examina o estado atual da análise de aprendizagem (LA) com foco na integração de dados em contextos de ensino superior. O documento descreve a escalabilidade na LA como um desafio a ser alcançado. A revisão resultou em 20 estudos analisados. As descobertas revelam que algumas fontes de dados, principalmente os Sistemas de Gerenciamento de Aprendizagem (LMS), são utilizadas com frequência em estudos de LA. No entanto, o número de fontes de dados integradas tende a ser limitado, e formatos de dados semelhantes são frequentemente combinados, o que representa um desafio técnico relativamente simples. Além disso, a literatura geralmente carece de informações detalhadas sobre os processos de integração de dados nos sistemas implementados,

e as especificações de dados educacionais, como a xAPI, são subutilizadas, apesar de seus possíveis benefícios.

24.3 Metodologia

A metodologia que vai ser adotada para o desenvolvimento deste trabalho inclui as seguintes etapas: revisão de literatura e benchmarking de tecnologias, avaliação de fontes de informação educativas, criação do modelo de armazenamento, implementação do Learning Data Warehouse e validação da proposta. Cada uma dessas etapas está relacionada com um dos objetivos específicos como se pode olhar na Tabela 24.1.

Table 24.1: Metedologia

Etapa	**Objetivo**
1 Revisão de literatura e benchmarking de tecnologias	Estudar as tecnologias existentes através da revisão de literatura e processo de benchmarking para identificar os elementos relevantes e opções de melhora no suporte da tomada de decisão.
2 Avaliação de fontes de informação educativas	Caracterização das plataformas e sistemas de informação educativos que podem ser modelados no Learning Data Warehouse para definir o alcance que terá este trabalho nesse aspecto.
3 Criação do modelo de armazenamento	Propor um modelo de Learning Data Warehouse a partir da caracterização das tecnologias e fontes de informação educativas existentes que permita a padronização e centralização de dados e posteriormente a aplicação de Learning Analytics.
4 Integração do Learning Data Warehouse ao serviço de predição	Integração do Learning Data Warehouse como o ao serviço de predição para aproveitar as vantagens da computação na nuvem.
5 Validação da proposta	Definir parâmetros validação do Learning Data Warehouse proposto a partir das métricas definidas para mensurar o impacto da proposta.
6 Difusão de resultados	Paralelo a todos os objetivos

Neste momento, a dissertação está em um estado significativo de progresso, finalizando os objetivos específicos associados à Integração do Learning Data Warehouse como o ao serviço de predição para aproveitar as vantagens da computação na nuvem e à definição de parâmetros validação do Learning Data Warehouse proposto a partir das métricas para mensurar o impacto da proposta.

A validação da implementação será realizada por meio de testes de performance, essenciais para garantir a qualidade dos sistemas. Esses testes têm como objetivo avaliar aspectos não funcionais, como velocidade, estabilidade, confiabilidade e escalabilidade. Durante esse processo, diferentes tipos de testes poderão ser aplicados para examinar o comportamento do sistema e identificar áreas com desempenho insatisfatório, tanto em condições favoráveis quanto adversas.

Os principais indicadores de desempenho abrangem o tempo de resposta, a quantidade de usuários simultâneos e a taxa de transferência, os quais podem oferecer uma visão completa do desempenho do sistema.

De acordo com o exposto, a data estimada para a defesa é o segundo semestre de 2024. Isso também se baseia na presença de dois produtos científicos, um publicado e o outro apresentado em evento científico.

24.4 Resultados parciais e esperados

Como resultado deste trabalho irá se buscar o fortalecimento do conhecimento técnico e investigativo, revisando o estado da arte atual respeito de tecnologias e técnicas de tratamento e armazenamento de dados educacionais de modo a adquirir o conhecimento que permita integrar um learning Data Warehouse a um serviço de Predição de Risco Acadêmico.

Como parte do progresso obtido no desenvolvimento do cronograma proposto, o processo de revisão da literatura e comparação das especificações de interoperabilidade para dados educacionais foi concluído, conforme refletido em um artigo enviado e aprovado para um evento acadêmico.

Da mesma forma, a caracterização das fontes de dados e os experimentos associados à comparação das especificações de interoperabilidade existentes foram realizados e são apresentados em [5].

Acknowledgements Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por meio do Programa de Pós-Graduação em Computação da Universidade Federal de Pelotas (UFPel - Chamada Interna PPGC nº 01/2022)

References

1. Dondorf, T.: Learning analytics for moodle: facilitating the adoption of data privacy friendly learning analytics in higher education. Ph.D. thesis, Dissertation, Rheinisch- Westfälische Technische Hochschule Aachen, 2022
2. Lockwood Hills Federal, L.: Total learning architecture: 2018 reference implementation specifications and standards (2019)
3. Paik, J.H., Himelfarb, I., Yoo, S.H., Yoo, K., Ha, H.: The relationships among school engagement, students' emotions, and academic performance in an elementary online learning. In: Proceedings of the 14th Learning Analytics and Knowledge Conference. pp. 219–230 (2024)
4. Rasheed, R.A., Kamsin, A., Abdullah, N.A.: Challenges in the online component of blended learning: A systematic review. Computers & Education 144, 103701 (2020)
5. Rocha, J.C., Hernández-Leal, E., Ramos, V., Munoz, R., Queiroga, E., Primo, T., Cechinel, C.: Implementação de um learning record warehouse para diferentes especificações de interoperabilidade em um learning management system. pp. 25–27 (2024). https://doi.org/10.5281/ZENODO.12084382, https://zenodo.org/records/12084382
6. Samuelsen, J., Chen, W., Wasson, B.: Integrating multiple data sources for learning analytics—review of literature. Research and Practice in Technology Enhanced Learning 14(1), 11 (2019)
7. Union, E.: New european interoperability framework promoting seamless services and data flows for european public administration (2017)
8. Whyte, A., Nayak, P., Johnston, J.: Lak16 workshop: Extending ims caliper analytics ™ with learning activity profiles. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. pp. 490–491 (2016)

25
Modelo híbrido para el balanceo de carga en agentes de compilación

Escobar Naranjo, Juan Camilo [1[0009-0008-3262-1668]]
Duque-Méndez, Nestor Dario [2[0000-0002-4608-281X]]

Resumen En el contexto de la computación en la nube, la necesidad de estrategias de autoescalado efectivas es cada vez más crítica, especialmente con el auge del escalado predictivo basado en inteligencia artificial. Sin embargo, a pesar de los avances en el escalado automático en la nube, hay una notable carencia de estudios que se centren en la aplicación de estos métodos al escalado dinámico de agentes de ejecución para pipelines de integración y despliegue continuo. Este artículo propone la adaptación de modelos de escalado híbridos para gestionar runners de manera autoalojada, aprovechando el modelo de pago por uso de la nube, con el fin de abordar esta laguna en la investigación y desarrollar estrategias de escalado avanzadas que respondan eficazmente a las cambiantes demandas de carga de trabajo.

Palabras clave contenedores, balanceo de cargas, agentes, aprendizaje máquina, devops.

25.1 Problema

25.1.1 Contexto

La virtualización y el uso de contenedores en la nube se han convertido en herramientas de gran utilidad para el despliegue de aplicaciones. Rossi et al., (2019)[1] afirman que esto se debe a que los contenedores son una forma eficiente de empaquetar y distribuir aplicaciones, lo que facilita su implementación y gestión en entornos de nube. Los contenedores, especialmente en la nube, ofrecen la posibilidad de manejar las variaciones de carga de trabajo al explotar la elasticidad tanto vertical como horizontal Rossi et al., (2019)[1]. Sin embargo, esta asignación no toma en cuenta las fluctuaciones futuras de la carga de trabajo, lo que puede afectar el rendimiento y el costo de funcionamiento de las aplicaciones en caso de cambios significativos en la demanda.

Universidad Nacional de Colombia Sede Manizales[1,2] e-mail: jcescobarn,ndduquemen@unal.edu.co

25.1.2 Brecha de conocimiento

La revisión de la literatura destaca la necesidad de estrategias de autoescalado efectivas, particularmente en el escalado predictivo basado en inteligencia artificial, pero no se encontraron estudios que aborden específicamente el escalado dinámico de agentes de ejecución para pipelines. Zampetti et al., (2021)[2] describen un pipeline de integración y despliegue continuo como un proceso automatizado en servidores dedicados para detectar errores tempranamente y mejorar la calidad del software, crucial para la entrega continua de software de alta calidad. Estos pipelines requieren unidades de procesamiento llamadas runners, las cuales, según Chandrasekara y Herath, (2020)[3] son máquinas físicas o virtuales que se encargan de ejecutar procesos sin intervención humana. En el caso de Github Actions estos pueden ser gestionados por Github o en infraestructura propia. La computación en la nube y su modelo de pago por uso, junto con la gestión autoalojada de agentes, presentan una oportunidad para adaptar los modelos de escalado híbridos a este dominio específico, abordando una laguna en la investigación y probando estrategias de escalado avanzadas que se adapten a las cambiantes demandas de carga de trabajo.

25.2 Contribuciones

Este proyecto de tesis contribuirá a la mejora de costos y la reducción de tiempos en el despliegue de aplicaciones centrados en la nube publica de AWS. Al desarrollar y validar algoritmos de escalado predictivo basados en inteligencia artificial, adaptados específicamente para la gestión dinámica de runners en pipelines de integración y despliegue continuo, se ofrecerán soluciones prácticas que permitirán a las empresas definir estrategias para el balanceo de carga. Esto no solo mejorará la utilización de recursos y la calidad del servicio, sino que también reducirá los costos operativos asociados con la sobreprovisión y la subprovisión de infraestructura en entornos de nube. La comunidad académica se beneficiará de esta investigación mediante la introducción de nuevos algoritmos transformados para este dominio específico, contribuyendo al conocimiento teórico y metodológico, proporcionando un marco replicable para futuras investigaciones. De esta manera, se espera que este trabajo inspire y facilite el desarrollo de nuevas estrategias de escalado avanzadas que puedan adaptarse eficazmente a las cambiantes demandas de carga de trabajo en diversos contextos.

25.3 Evaluación de resultados

La identificación de métricas clave de rendimiento, como la métrica de tiempo, permitirá medir la efectividad del modelo híbrido entre predictivo y reactivo para el balanceo de carga en agentes de compilación. Se implementarán técnicas y algoritmos de aprendizaje maquina seleccionados, que serán evaluados mediante experimentos controlados y simulaciones, y comparados con métodos tradicionales de escalado dinámico. Además, se realizará una comparación entre los datos generados por el modelo y datos obtenidos previamente que reflejen el comportamiento del uso de agentes de compilación. El estudiante integrará estas técnicas en un entorno de prueba realista, presentando casos de estudio que demuestren la aplicación práctica del modelo y permitiendo una valuación integral de su rendimiento y resultados, incluyendo tanto los costos computacionales como los tiempos de respuesta. La documentación y presentación de estos resultados garantizará que las evidencias sean claras, coherentes y alineadas con los objetivos específicos del proyecto.

Referencias

1. Rossi, F., Nardelli, M., & Cardellini, V. (2019). Horizontal and vertical scaling of container-based applications using reinforcement learning. In: IEEE International Conference on Cloud Computing, CLOUD, 2019-July, pp. 329–338. https://doi.org/10.1109/CLOUD.2019.00061
2. Zampetti, F., Geremia, S., Bavota, G., & Di Penta, M. (2021). CI/CD Pipelines Evolution and Restructuring: A Qualitative and Quantitative Study. In: Proceedings - 2021 IEEE International Conference on Software Maintenance and Evolution, ICSME 2021, pp. 471–482. https://doi.org/10.1109/ICSME52107.2021.00048
3. Chandrasekara C. and Herath P., Hands-on Azure Pipelines, Apress, 2020. Available: https://doi.org/10.1007/978-1-4842-5902-3.

26
Predicción e interpretación del rendimiento estudiantil en educación superior utilizando técnicas basadas en IA

Karol Yulieth Pavas Leiva[1]
Eduardo José Villegas Jaramillo[2]
Nestor Dario Duque Mendez[3]

Resumen Esta investigación busca contribuir a abordar problemáticas críticas en el sistema educativo colombiano como lo es predecir e interpretar el rendimiento estudiantil mediante la aplicación de técnicas de inteligencia artificial, al explorar el potencial de las redes neuronales y el aprendizaje por transferencia en el análisis de datos en entornos educativos. Se espera entregar una herramienta que permita com-prender mejor el contexto y los factores determinantes en el rendimiento estudian-til y así ampliar la compresión de este fenómeno y otros relacionados, tales como la deserción estudiantil. Este estudio no solo proporcionará una interpretación in-tegral del rendimiento académico, sino que también sentará una base sólida para futuras investigaciones en el campo de la minería de datos educativos y avanzará en la comprensión del uso de las redes neuronales con transferencia de aprendiza-je en problemáticas diferentes a las imágenes.

Palabras clave rendimiento estudiantil, minería de datos educativos, redes neuronales, transferencia de aprendizaje.

26.1 Problema

La educación superior en Colombia enfrenta desafíos interrelacionados en términos de cobertura, deserción y rendimiento. Uno de los factores clave en la vida estudiantil y laboral es el rendimiento académico. En este contexto, el Instituto Colombiano para la Evaluación de la Educación (ICFES) [6], mide el rendimiento de los estudiantes de la educación superior a través de la prueba estandarizada Saber Pro. Según el informe de resultados de los exámenes Saber TyT y Saber Pro de 2023, el puntaje promedio global en Saber Pro fue de 145 en una escala de 0 a 300, un valor que se ha mantenido constante desde 2021. Es crucial analizar y predecir el rendimiento académico de los estudiantes, para comprender su impacto en problemáticas como la deserción. Los entornos educativos generan un gran volumen de datos que en los últimos años ha despertado un considerable interés en la comunidad académica con el objetivo de aprovechar su gran potencial de análisis y generación de conocimiento [2]. La minería de datos educativos (EDM) ha mostrado avances satisfactorios en el análisis de datos generados en entornos educativos, pero aún enfrenta desafíos significativos tales como la falta de técnicas propias del dominio específico y la interpretabilidad de

[1,2,3]Universidad Nacional de Colombia
Km 7 vía al Magdalena.
e-mail: kpavasl,ndduqueme@unal.edu.co

los resultados [1]. A pesar de la creciente disposición de datos generados en entornos educativos, en Colombia, estos no se han aprovechado adecuadamente, limitando el análisis a la estadística descrip-tiva y a los modelos matemáticos, denotando una ausencia de la aplicación de técnicas avanzadas de minería de datos en entornos educativos.

Entre los avances más recientes aplicados a la EDM se encuentra el uso de técnicas de inteligencia artificial (IA), sin embargo, se observa que las investigaciones se enfo-can en el resultado específico de la predicción, olvidando la identificación y caracte-rización de los factores que influyen significativamente en su rendimiento académico [3]. Dentro de las técnicas de IA más frecuentemente utilizadas, se encuentran el desarrollo de modelos usando redes neuronales artificiales (RNA) para la predicción de la calificación final de los estudiantes [4] y el uso de árboles de decisión (AD) para este mismo propósito [5]. En este contexto, el uso de redes neuronales (RN) y el aprendizaje por transferencia (AT) que comúnmente son utilizadas en el procesamiento de imágenes, ofrecen una opor-tunidad para ampliar la comprensión del rendimiento estudiantil, abordando problemas como la falta de información etiquetada en las bases de datos y el desbalance entre las categorías, conduciendo a un análisis más robusto del desempeño de los estudiantes, abriendo la posibilidad de realizar una interpretación amplia de esta problemática, proporcionando información valiosa para la toma de decisiones estratégicas en el ámbito educativo.

26.2 Pregunta de investigación

¿Cómo la utilización de técnicas basadas en IA puede mejorar y en qué medida la predicción e interpretación del rendimiento estudiantil, en los entornos de educación superior en Colombia?

26.3 Contribuciones esperadas

En este contexto se espera contar con un modelo predictivo de RN con AT capaz de estimar el rendimiento estudiantil comparándolo con otras técnicas de RN utilizadas para este mismo propósito, bajo diferentes métricas y medidas de desempeño. Ade-más, se busca su aplicación práctica para ofrecer a las instituciones educativas y demás interesados una herramienta que proporcione una interpretación del rendi-miento estudiantil. Este enfoque no solo entregará resultados cuantitativos, sino que también permitirá comprender de manera integral el comportamiento y contexto de este fenómeno.

26.4 Evaluación de resultados

Finalmente, después del desarrollo y entrenamiento del modelo, se llevará a cabo un proceso de EDM siguiendo una metodología adecuada; iniciando con la recolección y preprocesamiento de los datos obtenidos de las instituciones de educación superior de Colombia, luego, la aplicación del modelo previamente entrenado para concluir con la evaluación mediante técnicas de validación cruzada y comparación del ren-dimiento a través de métricas previamente establecidas, con técnicas comúnmente usadas en el campo de EDM tales como las RNA y los AD.

Referencias

1. Hernandéz Leal, E. (2023). MODELO DE DOMINIO ESPECÍFICO PARA ANÁLISIS Y MINERÍA DE DATOS EDUCATIVOS [Tesis doctoral inédita]. Universidad Nacional de Colombia.
2. Baker, R. S., & Inventado, P. S. (2014). Educational Data Mining and Learning Analytics. In Learning Analytics (pp. 61–75). Springer New York. https://doi.org/10.1007/978-1-4614-3305-7_4.
3. Reyes, N. S., Morales, J. B., Moya, J. G., Teran, C. E., Rodriguez, D. N., & Altamirano, G. C. (2019). Modelo para predecir el rendimiento académico basado en redes neuronales y analítica de aprendizaje. Revista Ibérica de Sistemas e Tecnologias de Informação, (E17), 258-266.
4. Usman, O. L., & Adenubi, A. O. (2013). Artificial neural network (ANN) model for predicting students' academic performance. Journal of Science and Information Technology, 1(2), 23-37.
5. Hu, Y. H., Lo, C. L., & Shih, S. P. (2014). Developing early warning systems to predict students' online learning performance. Computers in Human Behavior, 36, 469-478.
6. Icfes. (2024). Resultados de los exámenes Saber TyT y Saber Pro Publicación preliminar de resultados individuales 2023.

27
Detección Automatizada de Problemas de Erosión Arquitectónica para aplicaciones Android

Juan Camilo, Acosta Rojas[1][0009-0001-8282-3400]
Camilo Andrés, Escobar Velasquez[2][0000-0001-8414-9301]]

Resumen La ingeniería de software involucra diferentes esfuerzos con el objetivo de desarrollar software de calidad. Dentro de estos esfuerzos se encuentra la definición de la arquitectura del sistema, priorizando los diferentes atributos de calidad de acuerdo a las necesidades de este. Sin embargo, debido a diferentes razones, el desarrollo de la solución puede desviarse de esta arquitectura, generando problemas en el rendimiento del sistema, y por ende, afectando la buena experiencia del usuario. Este impacto en la calidad del software puede verse representado en una mayor proporción en soluciones de software para dispositivos móviles, dado que estos cuentan con recursos más limitados. Los trabajos previos enfocados en esta área estudian y detectan problemas en diferentes áreas como seguridad, conectividad, entre otras. Sin embargo, no se ha estudiado a profundidad el impacto de la ”erosión arquitectónica” en soluciones móviles. El objetivo de este proyecto de investigación es detectar y localizar problemas de erosión arquitectónica en aplicaciones móviles por medio de mecanismos automatizados de análisis estático y dinámico basados en modelos de deep learning.

Palabras clave Erosión Arquitectónica · Software · Arquitectura · Android.

27.1 Detección Automatizada de Problemas de Erosión Arquitectónica para aplicaciones Android

La correcta realización del ciclo de vida del software es fundamental para la sostenibilidad de un proyecto de software a lo largo del tiempo. Para ello cada una de sus fases es esencial, desde su diseño, en donde se plantea los estándares y prioridades de acuerdo al negocio. Sin embargo, pese a los grandes esfuerzos por elaborar de manera estricta las fases de diseño e implementación, por diferentes factores se debe invertir una gran porción de recursos en procesos de mejora y refactoring, para poder mantener la disponibilidad, y por ende, el rendimiento de este proyecto. Es por ello, que se decide por encontrar metodologías que permitan detectar y corregir de manera temprana las violaciones al diseño inicial, que pueda comprometer el rendimiento del proyecto de software, y por ende, la buena experiencia de usuario. Dentro de ello, se encuentran las violaciones a la arquitectura del sistema, dichas violaciones son consideradas dentro del fenómeno de ”erosión arquitectónica”[1].

Universidad de los Andes
e-mail: jcescobarn@unal.edu.co

El impacto de la erosión arquitectónica se puede ver representado en la degradación de la calidad del software y su desempeño [1, 2]. Esto podría representar un problema de mayor magnitud para los ecosistemas móviles, dado que se cuenta con menos recursos, representando un gasto más significativo para dispositivos móviles[3, 4]. De igual forma, estos problemas podrían representar una degradación de la experiencia de usuario, siendo esta una variable del éxito de aplicaciones móviles[5]. Estos problemas podrían llevar a que los usuarios desinstalen la aplicación. Es por esto, que la definición de metodologías y mecanismos de detección y localización de problemas de erosión arquitectónica es necesaria en aplicaciones Android. Pese a los grandes avances en investigaciones para solución de issues de seguridad, conectividad y performance para el ecosistema mobile, existe una gran deuda técnica respecto a implementaciones de detección de violaciones arquitectónicas[6, 7, 8, 9, 10, 11, 12]. Pese a esto, se han presentando grandes avances en el ecosistema del lado del servidor, logrando el desarrollo de herramientas desarrolladas bajo el paradigma basado en modelos para poder buscar y detectar diferentes reglas ingresadas a dichas herramientas basadas en el árbol de invocaciones o de sintaxis abstracta [13, 14, 15]. En conclusión, las investigaciones realizadas son un buen fundamento, pero no tienen en cuenta las características únicas del ambiente de desarrollo Android.

27.1.1 Objetivos y Preguntas de Investigación

El objetivo principal de esta investigación es habilitar la detección automatizada de problemas de erosión arquitectónica para cualquie aplicación móvil durante la etapa de desarrollo. Para esto se explorarán las herramientas actuales de análisis de código y las oportunidades de extensión que existen por medio del uso de técnicas de deep learning. Se plantean resolver las siguientes preguntas de investigación durante el desarrollo de esta tesis:

- ¿Qué usos potenciales de las herramientas actuales se puede implementar para detectar y localizar erosión arquitectónica para desarrollo móvil?
- ¿Qué posible uso e impacto podrían tener las técnicas de deep learning en la detección y localización de erosiones arquitectónicas?
- ¿Qué herramienta se podría proponer para detectar y corregir erosión arquitectónica en aplicaciones Android?

27.2 Contribución esperada

Dado estos planteamientos, se establece, como productos finales, los siguientes:

- Dataset de aplicaciones móviles con problemas de erosión arquitectónica.
- Desarrollo de una solución automatizada para la detección y localización de problemas de erosión arquitectónica, que se ajuste a los procesos de desarrollo de aplicaciones móviles.
- Validación de la solución propuesta con desarrolladores de aplicaciones móviles y estudios empíricos del impacto de la solución de los problemas de erosión en diferentes atributos de calidad.

Referencias

1. De Silva, L., Balasubramaniam, D. (2012). Controlling software architecture erosion: A survey. Journal of Systems and Software, 85(1), 132–151. doi:10.1016/j.jss.2011.07.036

2. Mendoza, C. Bocanegra, J. Garcés, k. Casallas, R. ”Architecture violations and visualization in continuous integration pipeline”. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.3004.
3. Neves, C. Lin, C. Nigam, s. Patapas, D. Eguiluz, A. Islam, T. Malavolta, I. ”A Study on the Energy Consumption and Performance of Single-Activity Android Apps”. 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10190841
4. Kayande, D. Shrawankar, U. ”Performance analysis for improved RAM utilization for Android applications”, 2012. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/6349500
5. Noor, A. Mehmood, M. Daas, T. ”End Users’ Perspective of Performance Issues in Google Play Store Reviews”. 2022. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-031-21388-545
6. Talukder, A. Shariar, H. Qian, K. Rahman, M. Ahamed, S. Wu, F. Agu, E. “DroidPatroo: A Static Analysis Plugin For Secure Mobile Software Development”. 2019 IEEE 43rd Annual Computer Software And Applications Conference (COMPSAC). [Online]. Available: https://par.nsf.gov/servlets/purl/10104316.
7. “Improve Your Code With Lint Checks”. [Online]. Avalable: https://developer.android.com/studio/write/lint.
8. “Guide App To Architecture”. [Online]. Available: https://developer.android.com/topic/architecture.
9. Barstad, V. Goodwin, M. Gjøsæter, T. “Predicting Souce Code Quality with Static Analysis and Machine Learning”. [Online]. Available: Predicting source code quality with static analysis and machine learning.
10. Rozo, A. Velazques, C. Acero, J. Linares, M. Bavota, G. ”Detecting Connectivity Issues in Android Apps”. 2022. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9825832
11. Kukkonen, H. kurkela, V. Infotech Oulu, Oasis Group and Department of Information Processing Science. ”DEVELOPING SUCCESSFUL MOBILE APPLICATIONS”. 2003. [Online]. Available: https://www.researchgate.net/profile/Harri-OinasKukkonen/publication/228785470DevelopingSuccessfulMobileApplications/links/02e7e52838f746f1f1000000/Developing-Successful-Mobile-Applications.pdf
12. Wang, Y. liu, X. Mao, W. Wang, W. ”DCDroid: automated detection of SSL/TLS certificate verification vulnerabilities in Android apps”. 2019. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3321408.3326665
13. Zhong, C. Zhang, H. Huang, H. Chen, Z. Li, C. Liu, X. Li, S. ”DOMICO: Checking conformance between domain models and implementations”. 2023. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/spe.3272
14. Wang, T. Li, B. ”EsArCost: Estimating repair costs of software architecture erosion using slice technology”. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/abs/pii/S0164121223002704
15. Ruichen, H. ”An Automated Approach to Check Software Architecture Erosion”. 2023. [Online]. Available: https://pure.tue.nl/ws/portalfiles/portal/319375239/HuR.pdf

28
Planificación automática para el modelado del proceso de innovación abierta

Tatiana Pachón Ávalos[1]
Jhon Wilder Sánchez-Obando[2][0000-0001-8525-5835]
Néstor Darío Duque-Méndez[3][0000-0002-4608-28IX]

Resumen La innovación abierta ofrece diferentes beneficios a las organizaciones, mejorando así los resultados en el desarrollo de proyectos, facilitando la expansión empresarial. Sin embargo, gestionar la innovación abierta implica enfrentar desafíos significativos, como la gestión de grandes volúmenes de datos y la coordinación entre las partes interesadas, lo cual puede llegar a ser complejo y costoso sin la automatización adecuada. Esta tesis se orienta aplicar un modelo de planificación automática para la innovación abierta. Actualmente, no existe un modelo que aborde completamente y de manera efectiva estos desafíos, lo que representa un vacío de conocimiento en el campo. La planificación automática tiene el potencial de mejorar la precisión y eficiencia de los procesos de innovación abierta al permitir la generación automatizada de planes adaptativos.

Palabras clave Innovación abierta, planificación automática. Modelado de procesos

28.1 PROBLEMA

28.1.1 Contexto

La gestión de la innovación abierta puede implicar la generación, recolección y análisis de grandes cantidades de datos y conocimientos provenientes de múltiples fuentes, lo que puede resultar una tarea costosa y compleja si se realiza de manera manual. El factor de tiempo puede requerir la participación de múltiples partes interesadas lo que hace que la coordinación de los esfuerzos y la sincronización del tiempo sea más difícil cuando el proceso se lleva a cabo manualmente. También puede ser difícil de identificar y acceder a la información para el éxito del proyecto de innovación abierta (Chesbrough 2019).

[1,3]Universidad Nacional de Colombia, Facultad de Administración, Maestria en Administración.
e-mail: tpachon,dduqueme@unal.edu.co
[2]Universidad de Manizales, Facultad de Ciencias Contables, económicas y administrativas
e-mail: jwsanchez@umanizales.edu.co

28.1.2 Brecha del conocimiento

La revisión de literatura destaca la falta de un modelo automatizado adecuado para la planificación en la innovación abierta, ya que limita la capacidad de las organizaciones para poder adaptarse ágilmente a las diferentes demandas del mercado y aprovechar al máximo las oportunidades que brinda la innovación, donde la implementación de herramientas de planificación automática pueden mejorar la eficiencia operativa, y la precisión en la toma de decisiones en el entorno empresarial[3] y Saetti, 2019; Marrella y Chakraborti, 2019).

28.2 Contribuciones

Este proyecto de tesis contribuirá a la mejora de procesos de innovación abierta aportando diferentes alternativas que permitan su aplicación de una manera que facilite las tareas a desarrollar y permita disminuir el tiempo y costo para la ejecución de proyectos en las organizaciones. Aportando también métodos y herramientas que faciliten la gestión de proyectos innovadores. La comunidad académica se beneficiará de esta investigación, ya que se exploran nuevas aplicaciones de inteligencia artificial en la gestión de la innovación, proporcionando modelos prácticos para entender los modelos complejos de innovación abierta.

28.3 Evaluación de resultados

Se llevara a cabo un estudio de caso detallado en un entorno real para medir la precisión y adaptabilidad del modelo con la generación de comparaciones con la gestión manual, analizando el impacto estratégico en la capacidad del proyecto para lograr capturar oportunidades , se evaluara como el modelo facilita la colaboración entre diferentes actores y mejora la coordinación de actividades en proyectos de innovación , la aceptación y usabilidad del modelo se evaluaran mediante métricas que permitan evidenciar la efectividad del modelo. Se realizará un análisis comparativo de los datos para evidenciar la hipótesis sobre la mejora del desempeño y proporcionar recomendaciones practicas para la implementación futura de la planificación automática en la gestión de proyectos de innovación abierta.

Referencias

1. Heinrich, B., Bolsinger, M., & Bewernik, M. (2009). Automated planning of process models: The constructions of exclusive choices.
2. de Leoni, M., & Marrella, A. (2017). Aligning real process executions and prescriptive process models through automated planning. Expert Systems with Applications, 82, 162–183.
3. Heinrich, B., Krause, F., & Schiller, A. (2019). Automated planning of process models: The construction of parallel splits and synchronizations. Decision Support Systems, 125, 113096.
4. Wynn, D. C., & Clarkson, P. J. (2018). Process models in design and development. Research in Engineering Design, 29, 161–202.
5. Heinrich, B., Bolsinger, M., & Bewernik, M. (2009). Automated planning of process models: The constructions of exclusive choices.
6. Santos, F. P., Teixeira, A. P., & Guedes Soares, C. (2018). Maintenance planning of an offshore wind turbine using stochastic Petri nets with predicates. Journal of Offshore

7. Mannhardt, F., de Leoni, M., Reijers, H. A., & van der Aalst, W. M. P. (2016). Balanced multi-perspective checking of process conformance. Computing, 98, 407–437.

29

Análisis de protocolos para transmisión de datos utilizando sensores remotos en cuerpos de aguas continentales
Analysis of protocols for data transmission using remote sensors in continental water bodies

Alex Osorio-Rasero[1][0009-0005-8517-8207]
Javier Montoya[2][0000-0003-0507-452X]
Plinio Puello[3][0000-0002-3908-3005]

29.1 Descripción del problema

En el Departamento de Bolívar, es considerado un departamento de relevancia dada su biodiversidad en cual se destacan su flora y fauna, además de tener muchas cuencas de agua donde existen ríos, lagos y lagunas; tales como: canal del dique, canal matute, ciénaga de la virgen etc. Por tal motivo existen entidades que se encargan de promover el uso adecuado del ambiente, y recursos naturales, otro proceso es monitorear los cuerpos de agua. Según entrevistas realizadas estas entidades cuentan con diferentes puntos de monitoreo de calidad de agua in-situ y la recolección de datos periódicas [1, 2, 3].

Efecto 1: Se tiene limitaciones en los equipos tecnológicos para transmitir los datos.	**Efecto 2**: Se generan pérdidas de datos porque la información que queda en el dispositivo se satura y se pierde la información.	**Efecto 3**: Se capturan datos esporádicos in-situ cuando se visitan los diferentes puntos.
Problema: No se encuentra diseñado un sistema que transmita los datos en tiempo real de las variables ambientales en cuerpos de agua a una base de datos usando protocolos de transmisión y comunicación		
Causa 1: No se cuenta con un inventario de equipos que permitan la transmisión de datos en tiempo real en el departamento de Bolívar.	**Causa 2**: No se cuenta con desarrollos propios de sistemas tecnológicos que permitan la transmisión de datos de monitoreo en tiempo real en el departamento de Bolívar.	**Causa 3**: No se cuenta con desarrollos propios de bases de datos para la captura de información en el departamento de Bolívar.

Universidad Nacional de Colombia, Km 7 vía al Magdalena. e-mail: jcescobarn@unal.edu.co · Duque-Méndez Nestor-Dario
Universidad Nacional de Colombia, Km 7 vía al Magdalena. e-mail: name@email.address

29.2 Justificación

Las soluciones inteligentes para el monitoreo de la calidad del agua están ganando importancia con el avance e innovación de sensores novedosos. Los trabajos de Gámez [4], Geetha y Gouthami [5], presentan una descripción detallada de los avances realizados en el campo del monitoreo inteligente de la calidad del agua. Las anteriores experiencias aportan razones por las cuales es relevante contar con desarrollos propios de tecnología y abaratamiento de costos, siendo estos claros ejemplos en un contexto histórico sobre lo que es tradicional y también sobre lo que nos deparará el futuro para facilitar el análisis de la calidad del agua mediante el desarrollo de la ciencia y la tecnología. [6] Por tal razón, esta propuesta de investigación propone realizar el desarrollo de un sistema tecnológico que permita el envío en tiempo real de las variables ambientales en cuerpos de agua y una base de datos mediante protocolos de transmisión y comunicación, esto con el fin de dar apoyo a las entidades que se encargan del monitoreo de la calidad del agua para que logren reemplazar sus procesos tradicionales manuales por sistemas automatizados que les permitan tomar decisiones de manera oportuna, en el caso que exista una novedad en los procesos de revisión. [7, 8] Este desarrollo será innovador para las entidades de monitoreo del departamento de bolívar porque apoyará en el proceso acceso inmediato, la detección temprana de los diferentes resultados arrojados, mayor eficiencia en el monitoreo y fortalecerá la toma de decisiones basada en datos [9, 10].

29.3 Pregunta de Investigación

¿Cómo transmitir variables ambientales en cuerpos de agua en tiempo real entre sensores y una base de datos usando protocolos de transmisión y comunicación? 4 Aportes esperados Desarrollo de un Sistema de Transmisión de Datos en Tiempo Real: Creación de un sistema innovador que permita la transmisión en tiempo real de las variables ambientales de cuerpos de agua a una base de datos centralizada. Este sistema mejorará la precisión y la disponibilidad de los datos, permitiendo a las entidades de monitoreo actuar de manera proactiva y oportuna [11, 12]. Mejoras en la Eficiencia de Monitoreo Ambiental: Al automatizar la recolección y transmisión de datos, se reducirán los errores humanos y se optimizará el tiempo de respuesta frente a eventos críticos, como la detección de contaminación o cambios abruptos en la calidad del agua [11, 13]. Reducción de Costos y Dependencia Tecnológica: La implementación de tecnologías y desarrollos propios disminuirá la dependencia de soluciones externas y costosas, lo que a su vez podría reducir significativamente los costos operativos de las entidades responsables del monitoreo ambiental en el departamento de Bolívar. Fortalecimiento de la Capacidad de Toma de Decisiones Basada en Datos: Con un sistema que proporciona datos en tiempo real, las entidades podrán tomar decisiones informadas y basadas en evidencia, lo que mejorará la gestión de los recursos hídricos y la preservación del medio ambiente en la región [14, 15].

29.4 Evaluación de resultados

La evaluación de los resultados del proyecto se llevará a cabo en varias etapas, asegurando que los objetivos planteados se cumplan de manera efectiva. A continuación, se detallan los criterios y métodos de evaluación:

1. **Pruebas de Funcionamiento del Sistema:** Se realizarán pruebas exhaustivas del sistema de transmisión en tiempo real para verificar su fiabilidad y precisión en la captura y envío de datos.

Estas pruebas se llevarán a cabo en diferentes puntos de monitoreo de los cuerpos de agua en Bolívar, evaluando la consistencia y calidad de los datos transmitidos [16, 17].
2. **Comparación con Métodos Tradicionales:** Se compararán los resultados obtenidos mediante el nuevo sistema con los métodos tradicionales de monitoreo manual. Se evaluará la eficiencia, precisión y tiempo de respuesta, buscando mejoras significativas en estos aspectos [18, 19].
3. **Canálisis de Impacto en la Toma de Decisiones:** Se medirá el impacto del sistema en la toma de decisiones de las entidades encargadas del monitoreo ambiental. Esto incluirá la evaluación de la rapidez y efectividad con la que se responden a las alertas o novedades detectadas por el sistema [20, 21, 22].
4. **Satisfacción de las Entidades de Monitoreo:**Se realizarán encuestas y entrevistas con los usuarios del sistema para evaluar su satisfacción con la herramienta, la facilidad de uso y los beneficios percibidos en sus operaciones diarias [23, 24, 25].
5. **Monitoreo Continuo y Ajustes:** Tras la implementación inicial, se llevará a cabo un monitoreo continuo del sistema para identificar posibles mejoras o ajustes. Este proceso incluirá la recolección de retroalimentación y la realización de actualizaciones necesarias para optimizar el funcionamiento del sistema [26, 27, 28].

Referencias

1. Zuluaga, J., González, A., & Pérez, L. (2021). Evaluación del impacto de actividades agrícolas en la calidad del agua en la cuenca del río Bogotá, Colombia.
2. Geetha, S. Internet of things enabled real time water quality monitoring system (2017)
3. bina, N. A., & Cheng, S. C. (2022). A Review of Unmanned System Technologies with Its Application to Aquaculture Farm Monitoring and Management. In Drones(Vol. 6, Issue 1). MDPI.
4. Budiarti, R. P. N., Tjahjono, A., Hariadi, M., & Purnomo, M. H. (2019). Development of IoT for Automated Water Quality Monitoring System. 2019 International Conference on Computer Science, Information Technology, and Electrical Engineering (ICOMITEE), 1, 211–216. https://doi.org/10.1109/ICOMITEE.2019.8920900.
5. International Standard, & 29148, I. I. (2018). Systems and software engineering — Life cycle processes — Requirements engineering Ingénierie. Iso/Iec/Ieee 29148:2018, 2012(40).
6. LeidyMarcelaBuitragoMarquez2019. (2019). Redes LoRaWAN. Revisión de componentes funcionales en aplicaciones IoT.
7. A concise guide for IOT based Water Quality Monitoring. An End-To-End IoT Solutions Provider. (2021). Retrieved November 3, 2021.
8. B. Kitchenham, S. Charters. “Guidelines for Performing Systematic Literature Reviews in Software Engineering (version 2.3)”. Technical Report, Keele University and University of Durham, 2007.
9. Ubina, N. A., & Cheng, S. C. (2022). A Review of Unmanned System Technologies with Its Application to Aquaculture Farm Monitoring and Management. In Drones (Vol. 6, Issue 1). MDPI. https://doi.org/10.3390/drones6010012.
10. Jamroen, C., Yonsiri, N., Odthon, T., Wisitthiwong, N., & Janreung, S. (2023). A standalone photovoltaic/battery energy-powered water quality monitoring system based on narrowband internet of things for aquaculture: Design and implementation. Smart Agricultural Technology, 3, 100072. https://doi.org/10.1016/j.atech.2022.100072
11. Ullo, S. L., & Sinha, G. R. (2020). Advances in smart environment monitoring systems using iot and sensors. In Sensors (Switzerland) (Vol. 20, Issue 11). MDPI AG. https://doi.org/10.3390/s20113113.
12. Kumar, P. M., & Hong, C. S. (2022). Internet of things for secure surveillance for sewage wastewater treatment systems. Environmental Research, 203. https://doi.org/10.1016/j.envres.2021.111899.

13. Ahmad, M. I., Mohd, M. H., Nordin, R., Mohamed, F., Abu-Samah, A., & Abdullah, N. F. (2021). Ionizing radiation monitoring technology at the verge of internet of things. In Sensors (Vol. 21, Issue 22). MDPI. https://doi.org/10.3390/s21227629.
14. Goyal, P., Sahoo, A. K., Sharma, T. K., & Singh, P. K. (2019). Internet of Things: Applications, security and privacy: A survey. Materials Today: Proceedings, 34, 752–759. https://doi.org/10.1016/j.matpr.2020.04.737.
15. Kimothi, S., Singh, R., Gehlot, A., Akram, S. V., Malik, P. K., Gupta, A., & Bilandi, N. (2022). Intelligent energy and ecosystem for real-time monitoring of glaciers. Computers and Electrical Engineering, 102. https://doi.org/10.1016/j.compeleceng.2022.108163.
16. Jan, F., Min-Allah, N., & Düştegör, D. (2021). Iot based smart water quality monitoring: Recent techniques, trends and challenges for domestic applications. In Water (Switzerland) (Vol. 13, Issue 13). MDPI AG. https://doi.org/10.3390/w13131729.
17. Adhiparasakthi Engineering College. Department of Electronics and Communication Engineering, Institute of Electrical and Electronics Engineers. Madras Section, & Institute of Electrical and Electronics Engineers. (n.d.). Proceedings of the 2018 IEEE International Conference on Communication and Signal Processing (ICCSP) : 3rd - 5th April 2018, Melmaruvathur, India.
18. Institute of Electrical and Electronics Engineers, & PPG Institute of Technology. (n.d.). Proceedings of the 5th International Conference on Communication and Electronics Systems (ICCES 2020) : 10-12, June 2020.
19. N, S. M., Natarajan, K., & Shobha, K. R. (n.d.). An IoT based 6LoWPAN enabled Experiment for Water Management.
20. Nemade, B., & Shah, D. (2022). An efficient IoT based prediction system for classification of water using novel adaptive incremental learning framework. Journal of King Saud University - Computer and Information Sciences, 34(8), 5121–5131. https://doi.org/10.1016/j.jksuci.2022.01.009.
21. Bhimanpallewar, R. N., Rama Narasingarao, M., & Vaddeswaram, K. L. E. F. (n.d.). A HYBRID MACHINE LEARNING BASED IOT APPLICATION TO MONITOR, ASSESS AND IMPROVE PRECISION IN AGRICULTURE SYSTEM.
22. Simeon Okechukwu Ajakwe, Ihunanya Udodiri Ajakwe, Taesoo Jun, Dong-Seong Kim, Jae-Min Lee, CIS-WQMS: Connected intelligence smart water quality monitoring scheme, Internet of Things, Volume 23, 2023, 100800,ISSN 2542-6605.
23. T. Lakshmi Narayana, C. Venkatesh, Ajmeera Kiran, Chinna Babu J, Adarsh Kumar, Surbhi Bhatia Khan, Ahlam Almusharraf, Mohammad Tabrez Quasim, Advances in real time smart monitoring of environmental parameters using IoT and sensors,Heliyon,Volume 10, Issue 7,2024, e28195,ISSN 2405-8440.
24. Babangida Isyaku, Kamalrulnizam bin Abu Bakar, Nura Muhammed Yusuf, Mohammed Abaker, Abdelzahir Abdelmaboud, Wamda Nagmeldin, Software defined wireless sensor load balancing routing for internet of things applications: Review of approaches, Heliyon, Volume 10, Issue 9, 2024, e29965, ISSN 2405-8440.
25. Waheb A. Jabbar, Tan Mei Ting, M. Fikri I. Hamidun, Ajwad H. Che Kamarudin, Wenyan Wu, Jamil Sultan, AbdulRahman A. Alsewari, Mohammed A.H. Ali, Development of LoRaWAN-based IoT system for water quality monitoring in rural areas, Expert Systems with Applications, Volume 242,2024,122862, ISSN 0957-4174.
26. Jan, Farmanullah, Nasro Min-Allah, and Dilek Düştegör. 2021. "IoT Based Smart Water Quality Monitoring: Recent Techniques, Trends and Challenges for Domestic Applications" Water 13, no. 13: 1729.
27. M. H. Jewel and A. Al Mamun, "Internet of Things (IoT) for Water Quality Monitoring and Consumption Management," 2022 4th International Conference on Sustainable Technologies for Industry 4.0 (STI), Dhaka, Bangladesh, 2022, pp. 1-5.
28. Park, Jungsu, Keug Tae Kim, and Woo Hyoung Lee. 2020. "Recent Advances in Information and Communications Technology (ICT) and Sensor Technology for Monitoring Water Quality" Water 12.

Part IV
Póster

Este congreso ofrece una plataforma para la presentación de pósteres que abordan diversos temas en las áreas de ciencias de la computación. Los trabajos presentados abarcan desde innovaciones técnicas hasta aplicaciones prácticas, destacando avances en inteligencia artificial, procesamiento de datos, seguridad informática, desarrollo de software y más. El espacio está diseñado para fomentar el intercambio de ideas, impulsar la colaboración interdisciplinaria y proporcionar una visión integral de las tendencias emergentes en el campo.

This conference provides a platform for the presentation of posters covering various topics in the fields of computer science. The presented works range from technical innovations to practical applications, highlighting advances in artificial intelligence, data processing, software development, and more. The event is designed to foster the exchange of ideas, promote interdisciplinary collaboration, and offer a comprehensive view of emerging trends in the field.

Este congresso oferece uma plataforma para a apresentação de pôsteres que abordam diversos temas nas áreas de ciências da computação. Os trabalhos apresentados vão desde inovações técnicas até aplicações práticas, destacando avanços em inteligência artificial, processamento de dados, segurança cibernética, desenvolvimento de software e mais. O evento é projetado para promover o intercâmbio de ideias, incentivar a colaboração interdisciplinar e fornecer uma visão abrangente das tendências emergentes no campo.

30
Herramienta automatizada de segmentación para la evaluación de isquemias cerebrales en tomografías axiales computarizadas (TAC)

E. Ortiz[1]
J. Rivera[2]
M. Granja[3]
A. Salazar[4]
M. Hernández Hoyos[5]

30.1 Introducción

Las isquemias cerebrales son una de las principales causas de discapacidad en el mundo. El puntaje ASPECTS (Alberta Stroke Program Early CT Score) es una medida estándar mundial para evaluar el daño cerebral en pacientes con accidente cerebrovascular isquémico, cuantificando diez regiones específicas del cerebro. La cuantificación manual del puntaje ASPECTS es costosa en tiempo y recursos y está sujeta a una alta variabilidad interobservador.

30.2 Objetivos

Evaluar el impacto de la isquemia cerebral en el territorio de la arteria cerebral media, integrando técnicas de procesamiento de imágenes y aprendizaje automático en un algoritmo para segmentar y cuantificar las regiones ASPECTS.

30.3 Materiales y Métodos

El método está dividido en tres etapas principales: detección automática de los cortes axiales a nivel ganglionar y cortical, estandarización de las imágenes y segmentación y cuantificación de las regiones ASPECTS (Fig. 30.1).

[1,2,5]Universidad de los Andes
[1,2]Departamento de Ingeniería de Sistemas y Computación, Bogotá – Colombia
[5]Laboratorio de Electrofisiología y Telemedicina, Bogotá - Colombia
[2]Hospital Universitario Fundación Santa Fe de Bogotá, Departamento de Imágenes Diagnósticas, Bogotá – Colombia
[2]Hospital Universitario Fundación Santa Fe de Bogotá, Departamento de Neurología, Bogotá – Colombia

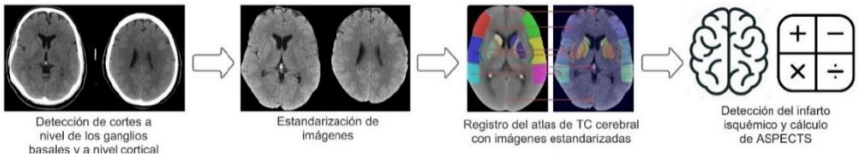


Fig. 30.1: Flujo operativo del algoritmo propuesto.

La detección automática de los cortes axiales a nivel ganglionar y cortical se hace a través de una red neuronal convolucional que devuelve los dos cortes de la imagen con mayor probabilidad de pertenecer a dichas regiones. La estandarización de dichos cortes consiste en redimensionarlos a un tamaño estándar, normalizar sus intensidades, reducir el ruido y eliminar el cráneo. Se utiliza un atlas de cerebro para segmentar las regiones ASPECTS, refinado mediante transformaciones rígidas y no rígidas. Se calcula la diferencia de intensidad media contralateral para regiones con intensidades de 15-50 HU. En la validación del algoritmo se utilizan 85 volúmenes de TAC (34 pacientes con isquemia cerebral, 51 controles) con píxeles de 0.5x0.5 mm y un grosor de corte de 0.5-10 mm. El patrón de oro del puntaje ASPECTS se determina mediante un acuerdo de dos tercios entre tres neurorradiólogos.

30.4 Resultados

Después de la segmentación automatizada, el 93% de las regiones (n=850) se consideraron ”Buen ajuste”, y un 2% ”requiere ajuste alto”. La puntuación ASPECTS se dicotomizó (ASPECTS ¡ 6, ¿ 6). El algoritmo logró una sensibilidad de 0.83, especificidad de 0.96 y un AUC de 0.9, lo que es comparable con el rendimiento de los neurorradiólogos. Además, se obtuvo un acuerdo sustancial con el estándar de oro (Kappa=0.77).

30.5 Conclusiones

Nuestra herramienta automatizada de ASPECTS se desempeñó de manera similar a los neurorradiólogos, demostrando su potencial para ahorrar tiempo y recursos en la evaluación de isquemias cerebrales. Esta tecnología puede servir como una herramienta de ayuda diagnóstica, contribuyendo a reducir la variabilidad interobservador en entornos clínicos.

Acknowledgements Agradecimientos. Universidad de los Andes, Hospital Universitario Fundación Santa Fe de Bogotá, Ministerio de Ciencia, Tecnología e Innovación (Grant 926-93362)

Referencias

1. Yu, Z. Chen, Y. Yu, H. Zhu, D. Tong, y Y. Chen, “An automated ASPECTS method with atlas-based segmentation”, Comput Methods Programs Biomed, vol. 210, p. 106376, oct. 2021, doi: 10.1016/j.cmpb.2021.106376.

2. C. Maegerlein et al., “Automated Calculation of the Alberta Stroke Program Early CT Score: Feasibility and Reliability”, Radiology, vol. 291, n.o 1, pp. 141-148, abr. 2019, doi: 10.1148/radiol.2019181228.

31
Análisis de factores que influyen en la selección de carreras de Ingeniería: Ingeniería de Sistemas

Emilcy Hernández-Leal[1]
Gloria Piedad Gasca-Hurtado[2]
Daniela Higuita Agudelo[3]

31.1 Introducción

El perfil vocacional es fundamental para la selección del programa académico de un joven porque orienta al aspirante a partir de sus intereses y aptitudes. Sin embargo, tanto los intereses como las aptitudes pueden estar influenciadas por factores relativos a su género [1], [2]. Dicha influencia sesga las percepciones de los individuos sobre su competencia en diversas tareas relevantes para la selección de su profesión, controlando la capacidad real [3]. Con el propósito de identificar factores que influyen en la selección del programa académico de un aspirante, este trabajo realiza un análisis exploratorio con estudiantes del programa Ingeniería de Sistemas de la Universidad de Medellín. Se utiliza un instrumento tipo encuesta, para recoger información relacionada con aspectos culturales y percepciones de los individuos y conducir un análisis cuantitativo de aspectos socioeconómicos, familiares y de trayectoria escolar.

31.2 Objetivos

El principal objetivo es la identificación de factores que influyen en la selección del programa de formación en Ingeniería de Sistemas, teniendo en cuenta aspectos culturales que puedan sesgar la percepción del individuo por su género. Se evaluó la influencia de estereotipos de género y expectativas sociales en la selección de un programa académico como el mencionado. La identificación de dichos factores facilita determinar la influencia de las creencias culturales sobre el género. Los resultados de este estudio, aunque preliminares, están encaminados a mejorar los procesos de selección de los aspirantes universitarios, tales como las pruebas vocacionales, a partir del diseño de estrategias que aporten en la disminución de la brecha de género que culturalmente se presenta.

[1,2,3]Universidad de Medellín
Carrera 87 #30-65, Medellín, Colombia
Semillero ARKADIUS - Minería de Datos
e-mail: ejhernandez@udemedellin.edu.co

31.3 Materiales y Métodos

Para realizar este estudio se construyó un instrumento de recolección de datos, tipo encuesta, organizado en tres dimensiones, 1) demográfica, definida a partir de seis preguntas de caracterización individual; 2) creencias culturales, constituida por ocho preguntas de factores familiares y 3) académica, con ocho preguntas de factores escolares. Este instrumento se construyó a partir de una revisión de literatura sobre factores asociados a la selección de programas STEM y del contexto de la Universidad en estudio. La muestra fue de 210 individuos, que corresponde al 45,7% del total de estudiantes matriculados en el programa Ingeniería de Sistemas en el primer semestre de 2024, por lo que se considera una muestra representativa. Todos los individuos aceptaron el consentimiento informado y las respuestas fueron tratadas de forma anonimizada. El análisis de resultados incluyó visualizaciones de orden diagnóstico y descriptivo de los datos.

31.4 Resultados

Del total de individuos encuestados, 202 aceptaron el consentimiento informado y 8 abandonaron el instrumento. La distribución en términos de género arrojó que la población está comprendida en 154 hombres, 45 mujeres y 3 individuos que no informan su género. Las edades de los encuestados están entre los 17 y los 32 años. Dentro de los datos demográficos se encuentra el estrato socioeconómico de la población, dado el contexto donde se ejecutó la encuesta. El 76% de los encuestados pertenece a los estratos socioeconómicos 1, 2 y 3, el 20% a estratos 4, 5 y 6 y el porcentaje restante no informa. Para el tipo de juegos de infancia, las categorías que sobresalen son videojuegos con un 79% para hombres, frente al 35% para mujeres. Otra categoría de juegos fueron los de aprendizaje (desarrollo cognitivo), donde el comportamiento para los dos géneros es igual, el 42%.

31.5 Conclusiones

Se evidencia que existen diferencias importantes frente a la brecha de género relacionadas con la cultura y los estereotipos que sesgan la selección de programas académicos. Estas diferencias se marcan desde la educación básica y desde el entorno familiar. Adicionalmente, se continúa presentando en la generación de padres de familia de los encuestados una predominancia de madres dedicadas a labores de casa, administrativas o de cuidado, frente a padres hombres con profesiones en ingeniería, comercio, servicios de seguridad o transporte. Como trabajo futuro propone validar el instrumento y profundizar en el análisis de resultados con minería de datos y de texto.

Acknowledgements A la Convocatoria 52 Fondo Semilla para nuevos investigadores Universidad de Medellín, que apoya el desarrollo del proyecto de investigación titulado “Diseño de una estrategia basada en gamificación para promover la participación de mujeres en ciencia y tecnología”.

Referencias

1. Avolio, B., Chávez, J., Vílchez-Román, C.: Factors that contribute to the underrepresentation of women in science careers worldwide: a literature review. Social Psychology of Education. 23, 773–794 (2020). https://doi.org/10.1007/S11218-020-09558-Y/METRICS.

2. Vieira De Mello, A., Petró, V., Meincke Melo, A., Fonseca Finger, Alice Bustamante Sá, M.: Egressas de Cursos de Computação: o quê as influenciou a escolherem um curso na área? In: Women in Information Technology (WIT). pp. 113–123. SBC (2023). https://doi.org/10.5753/WIT.2023.229508.
3. Correll, S.J.: Gender and the career choice process: The role of biased self-assessments. American Journal of Sociology. 10, 1691–1730 (2001).https://doi.org/10.1086/321299/0.

32
Modelo de autómata celular neuronal para aplicaciones en criptografía

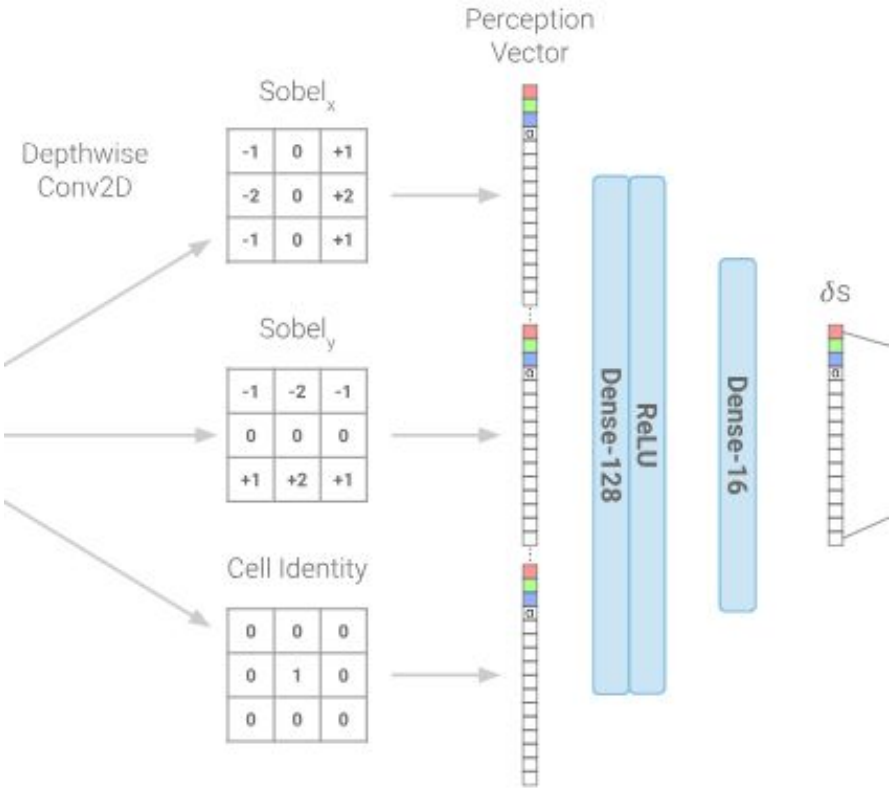
Alejandro Patiño Bedoya [1]
Elkin Duván Eraso Riascos[2]
José Antonio Valencia Aricapa[3]
Juan Carlos Riaño Rojas[4]

Introducción El autómata celular neuronal (ACN) se emplea en criptografía debido a su bajo costo computacional, lógica sencilla, escalabilidad y relativa facilidad de implementación. Además, sin la desventaja de un bajo nivel de seguridad, ya que las reglas del autómata celular neuronal no pueden ser determinadas fácilmente. El autómata celular neuronal emplea una red neuronal para inferir las reglas de un autómata celular ordinario y alcanzar un determinado objetivo, como encriptar información. Esta técnica resulta útil ya que las reglas generadas por la red neuronal son aleatorias y cambiantes, lo que dificulta su descifrado. En este trabajo, se presenta una descripción de un autómata celular neuronal diseñado para aprender de cada imagen de emoji y foto de persona y animales, en específico, con el objetivo de generar un patrón de crecimiento y evaluar su viabilidad como método de encriptación. El autómata celular neuronal se configura como una matriz tridimensional, donde cada célula busca expandirse hacia la imagen objetivo provista. Una vez que el autómata ha sido entrenado adecuadamente, se ejecuta el programa durante un número determinado de épocas, y así, encriptar la información contenida en otra imagen. Los resultados obtenidos indican que en un entorno de ejecución con GPU T4, se logra un proceso de encriptación de 4 minutos en promedio y eficiente, con un alto MSE, lo que sugiere un alto nivel de seguridad. Y en el caso de la decriptación se obtiene casi instantáneamente y un bajo MSE, lo que sugiere que se reduce la cantidad de información perdida. Objetivos

- Implementar un autómata celular neuronal para encriptar una imagen de una nanopartícula.
- Implementar una forma para que el autómata celular neuronal decripte la imagen de una nanopartícula.

Materiales y Métodos Se utilizará una arquitectura de red neuronal de Convolutional neural network (CNN) al que se le aplicaran unos kernels de sobel en x y y, que implica una relación de derivada en el espacio, que explica como interactúa cada pixel con su entorno. Además, se tiene un kernel de identidad de cada pixel, se aplana la matriz y después se ingresa en una CNN de 3 capas de profundidad, para después obtener un vector de 16 dimensiones que serán las reglas de cada pixel.

[1,2,3,4]Universidad Nacional de Colombia Sede - Manizales
Laboratorio de investigación PCM Computational Applications
Facultad de Ciencias Exactas
e-mail: \{alpatinob,eeraso,javalenciaa,jcrianoro\}@unal.edu.co

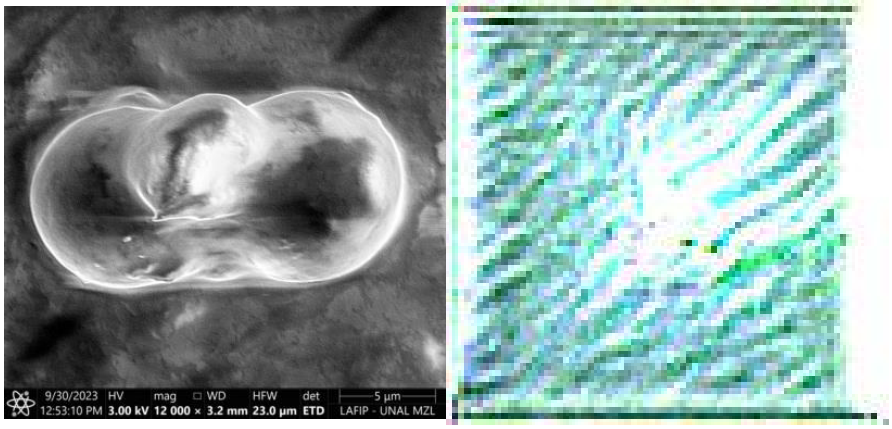


32.1 Resultados

Se implementó con satisfacción un autómata celular neuronal que fuera capaz de encriptar de una nanopartícula, lo hace en el entorno de ejecución GPU-T4 de Google colab en 1 minuto en promedio. Teniendo en cuenta una medida de encriptación con respecto al MSE de 4524.205. lo que indica que es un alto nivel de encriptación, sin embargo, puede variar por la misma naturaleza del autómata celular neuronal. A su vez el autómata neuronal en si mismo es capaz de reconstruir una imagen con un MSE de 0.007 con respecto a la imagen objetivo, lo que implica que se reproduce el 99.7% de la variabilidad de la imagen y se hace en solo 1 segundo.

32.2 Conclusiones

En conclusión, es factible crear un autómata celular neuronal para encriptar y decriptar una imagen con una mínima perdida de información. Se intentó implementar el ACN haciendo un preprocesamiento de la imagen con PCA, sin embargo, no funcionó, se tiene la hipótesis de que puede ser porque los vectores que conforman la matriz de PCA son independientes entre sí.



Acknowledgements Agradecimientos. Agradecimientos especiales a RevloG, profesores, padres y amigos.

Referencias

1. Alexander Mordvintsev, Ettore Randazzo, Eyvind Niklasson, Michael Levin.: Growing Neural Cellular Automata. 10.23915/distill.00023 (2020)
2. Ettore Randazzo, Alexander Mordvintsev, Eyvind Niklasson, Michael Levin, Sam Greydanus.: Self-classifying MNIST Digits. 10.23915/distill.00027.002 (2020)
3. L. Hernández Encinas, A. Hernández Encinas, S. Hoya White, A. Martín del Rey, G. Rodríguez Sánchez: Cifrado de imágenes usando autómatas celulares con memoria. https://digital.csic.es/handle/10261/21257 (2004)
4. Cortés Martinez, L.M., Alvarado Nieto, L.D. y Amaya Barrera, E.I..: Composite cellular automata based encryption method applied to surveillance videos YNA. 87, 213 (abr. 2020), 212–221. DOI:https://doi.org/10.15446/dyna.v87n213.81859.

33
ADOPET: Transformando la adopción de mascotas con innovación digital

Oscar Ivan Torres Chicue [1]
Johan Sebastián Lorza Pulido[2]
Heriberto Fernando Vargas Losada[3]

33.1 Introducción

El desarrollo de ”AdoPet: Transformando la Adopción de Mascotas” se centró en crear una aplicación web moderna para apoyar el proceso de adopción de mascotas que se ejecuta en el departamento del Caquetá, aunque es un sistema web nuevo, se da la posibilidad de dar cobertura nacional o internacional de acuerdo con su evolución. Utilizando una arquitectura hexagonal, la aplicación fue construida con .NET Core para el backend, Microsoft SQL Server para la gestión de datos y Angular para el frontend. Esta estructura permitió desarrollar una plataforma escalable, eficiente y fácil de mantener, asegurando una experiencia de usuario óptima y un manejo efectivo de los datos. La implementación de ASP.NET Core permitió la creación de una API RESTful eficiente y escalable, mientras que Microsoft SQL Server ofreció una solución robusta para el manejo de grandes volúmenes de datos con consultas optimizadas. Angular se utilizó para desarrollar una interfaz de usuario intuitiva y responsiva, mejorando significativamente la experiencia del usuario final. Por medio de esta sinergia de tecnologías se plantea una solución escalable, usable y dinámica para su utilidad.

33.2 Objetivo

Desarrollar una plataforma web eficiente y escalable que optimice el proceso de adopción de mascotas para el departamento del Caquetá.

33.3 Materiales y Métodos

El desarrollo de AdoPet, se centró en la creación de una aplicación web moderna y eficiente para contribuir en el proceso de adopción de mascotas. El proyecto utilizó una combinación de tecnologías y enfoques arquitectónicos para asegurar un sistema escalable y mantenible. De esta

[1,2,3]Universidad de la Amazonia
e-mail: \{os.torres,j.lorza,heri.vargas\}@udla.edu.co

manera, se planteó una estructura robusta y mantenible, para ello, se implementó una arquitectura hexagonal, conocida como arquitectura de puertos y adaptadores, se facilita la integración de la lógica de negocio central con diferentes interfaces externas, permitiendo la separación de responsabilidades y la integración de nuevas tecnologías (Salguero, 2020). El Backend se desarrolló con .NET Core, permitiendo la creación de una API RESTful de alto rendimiento y escalabilidad (Dykstra et al., 2023). Esta arquitectura soporta modularidad, facilitando el mantenimiento y la adición de nuevas funcionalidades. En cuanto a la gestión de datos, se optó por Microsoft SQL Server, una base de datos que proporciona robustez y eficiencia (West & Ray, 2024). Se implementaron consultas optimizadas y procedimientos almacenados para mejorar los tiempos de respuesta a los clientes en sus peticiones. Además, se configuraron datos relacionales a través de índices y claves foráneas para acelerar el acceso a los datos críticos. El Frontend se desarrolló bajo Angular, un framework que permite la creación de aplicaciones de una sola página (SPA) interactivas y responsivas (Angular, 2024). Angular, aseguró una interfaz de usuario consistente y accesible en diversos dispositivos. Los formularios reactivos con validaciones integradas y los servicios HTTP para la comunicación con la API contribuyeron a una experiencia de usuario fluida y segura.

33.4 Resultados

El desarrollo de la Aplicación web, se fundamentó bajo la implementación de la API RESTful con ASP.NET Core y es así, que permite el manejo de grandes solicitudes sin degradación significativa del rendimiento. El proceso de comunicación entre los componentes del sistema esenciales para

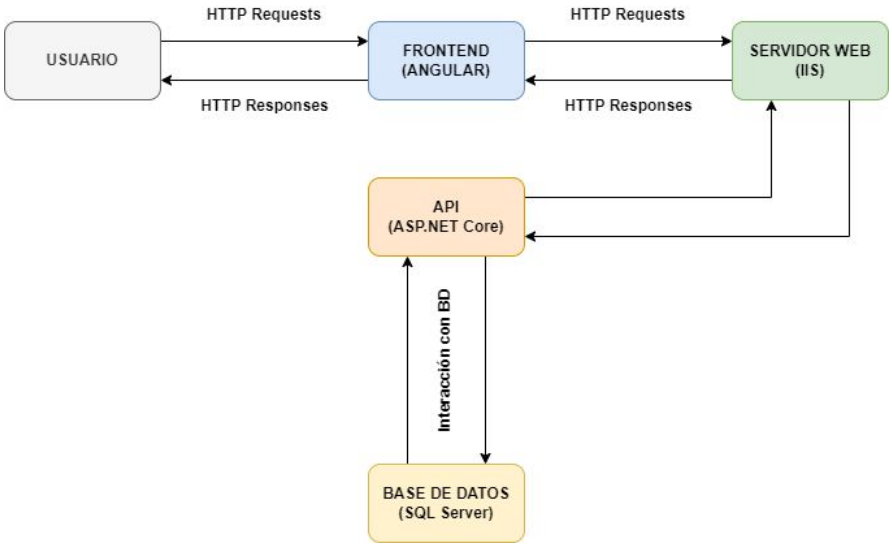


Fig. 33.1: Proceso de comunicación del sistema. Elaboración propia.

la aplicación se construyó con base en los usuarios que interactúan con la interfaz desarrollada en Angular, enviando solicitudes HTTP a la API RESTful a través del servidor web (Internet Information Server). La API procesa las solicitudes, interactúa con la base de datos para obtener o

modificar datos, y envía respuestas de la petición al Frontend. Este, a su vez, actualiza la interfaz de usuario según las respuestas recibidas, proporcionando una usabilidad ideal del sistema. La API

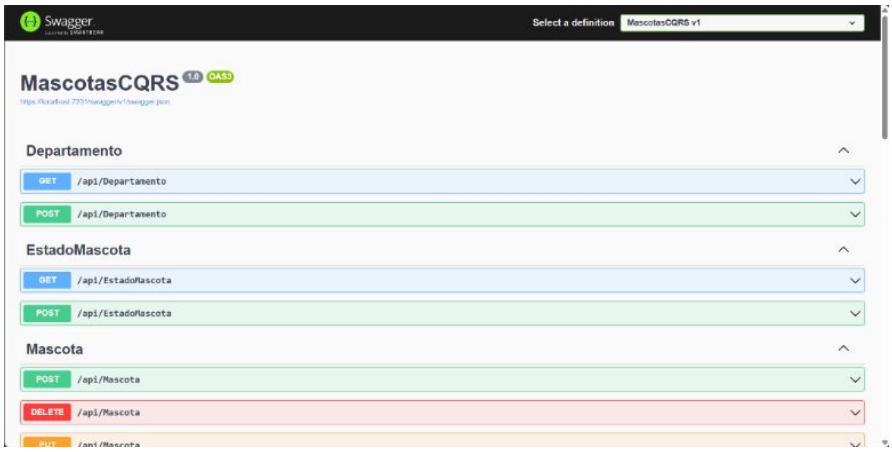


Fig. 33.2: Documentación de la API usando Swagger, Elaboración propia.

se documenta y fue validada utilizando Swagger, una herramienta que permite visualizar y probar las diferentes endpoints de la API de una manera sencilla y práctica. La imagen muestra la vista de Swagger de la API, donde se pueden observar los diferentes endpoints disponibles para gestionar departamentos, estados de mascotas y las propias mascotas. El desarrollo del frontend con Angular proporciona una interfaz de usuario atractiva y fácil de usar, los formularios reactivos y los servicios HTTP permiten una interacción fluida y segura con la API y de esta forma la aplicación web es amigable para el usuario final.

33.5 Conclusiones

La implementación del backend con ASP.NET Core y una arquitectura hexagonal ha demostrado ser altamente eficiente y escalable. La API RESTful desarrollada permite manejar altos volúmenes de solicitudes simultáneas sin degradación del rendimiento, asegurando una respuesta rápida y fiable. Esta estructura modular facilita el mantenimiento y la expansión futura del sistema, permitiendo una adaptación flexible a nuevas necesidades y funcionalidades. El desarrollo del frontend con Angular ha resultado en una interfaz de usuario atractiva, intuitiva y responsiva. Los formularios reactivos y los servicios HTTP integrados aseguran una comunicación fluida y segura con la API backend, mejorando la experiencia del usuario final. Esta combinación de tecnologías ha permitido crear una interfaz que es fácil de usar y accesible desde diversos dispositivos, aumentando la satisfacción y el compromiso de los usuarios. AdoPet ha demostrado ser una solución efectiva para abordar el problema de las mascotas desamparadas y en proceso de adopción. Esto no solo mejora el bienestar de los animales, sino que también optimiza la eficiencia operativa de las personas que dan en adopción a las mascotas. Al proporcionar una mayor visibilidad de las mascotas y facilitar el proceso de adopción, la aplicación contribuye significativamente a encontrar hogares adecuados para más animales, cumpliendo así su misión de transformar el proceso de adopción y mejorar la vida de las mascotas y sus adoptantes.

Referencias

1. Angular. (2024). Angular—Introduction to the Angular docs. https://v17.angular.io/docs
2. Dykstra, T., Anderson, R., Buck, A., Latham, L., Larkin, K., Alankuş, E., Cavuoti, S., Addie, S., Montemagno, J., Myers, A., & Dresko, A. (2023, noviembre 30). Creación de API web con ASP.NET Core. https://learn.microsoft.com/es-es/aspnet/core/web-api/?view=aspnetcore-8.0
3. Salguero, E. (2020, mayo 22). Arquitectura Hexagonal. Medium. https://medium.com/@edusalguero/arquitectura-hexagonal-59834bb44b7f
4. West, R., & Ray, M. (2024, febrero 14). What is SQL Server? - SQL Server. https://learn.microsoft.com/en-us/sql/sql-server/what-is-sql-server?view=sql-server-ver16

34
Hacia un sistema computacional en web para prototipos con Inteligencia Artificial para la gestión de recursos hídricos

Joseph Santiago Portilla-Martínez[1]
Raúl Rojas-Herrera[2]
Viviana Vargas-Franco[3]

34.1 Introducción

Los sistemas tradicionales de gestión de recursos hídricos a menudo no abordan de manera eficaz la incertidumbre del mundo moderno. Los sistemas basados en lógica difusa ofrecen una solución innovadora al proporcionar decisiones más realistas en modelos complejos. Este trabajo propone desarrollar un sistema computacional que utiliza lógica difusa para la gestión eficiente de recursos hídricos. Una herramienta accesible, dinámica y adaptable, permitiendo a los expertos introducir cambios en los modelos sin necesidad de programar, lo que facilita su aplicación en diversas regiones y condiciones específicas.

34.2 Objetivo

Desarrollar y desplegar un sistema computacional en web que permita gestionar prototipos basados en modelos de lógica difusa para la gestión de recursos hídricos.

34.3 Materiales y Métodos

Con el fin de aprovechar de manera óptima los recursos de dominio e infraestructura, el sistema se basa en una arquitectura multitenant que permite gestionar Sistema para prototipos con Inteligencia Artificial para la gestión de recursos hídricos eficientemente modelos de conocimiento independientes, cada uno con su conjunto de datos y reglas de conocimiento. Se implementa una API con arquitectura de microservicios en Spring Boot, la cual divide el sistema computacional web en tres servicios principales, cada uno con una función particular El núcleo del sistema de gestión de

[1,3]Universidad Nacional de Colombia - Sede Palmira
Grupo de Investigación: Monitoreo, modelación y gestión de cuencas hidrográficas - UNAL Palmira.
e-mail: portillajs@gmail.com, e-mail: vvargasf@unal.edu.co
[2]Universidad del Valle
e-mail: raulrojash@gmail.com

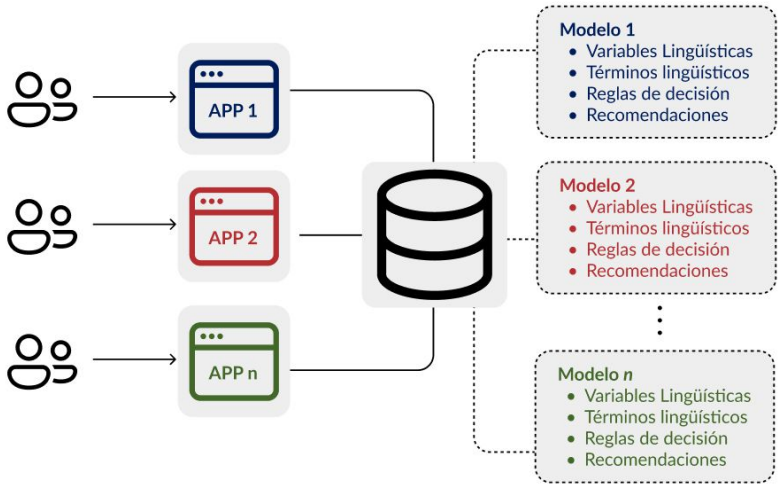


Fig. 34.1: Arquitectura Multitenant del modelo de datos

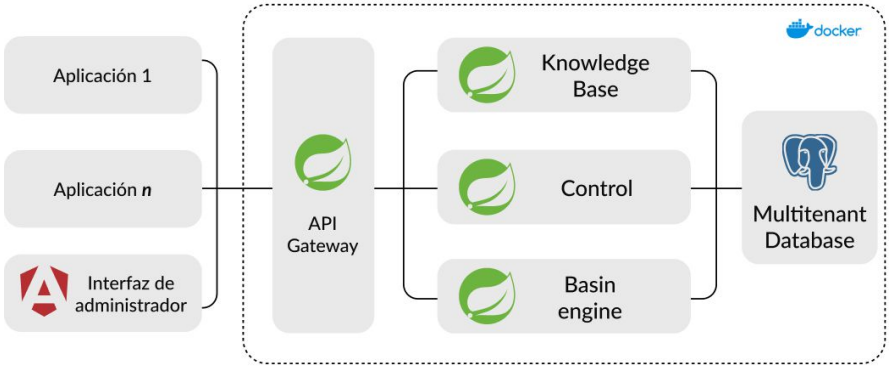


Fig. 34.2: Arquitectura de microservicios

recursos hídricos es el motor de lógica difusa desarrollado con la biblioteca JFuzzyLogic de Java.

34.4 Resultados

- **Escalabilidad**: Las arquitecturas multitenant y de microservicios aseguran que el sistema pueda escalar en función de las necesidades.
- **Adaptabilidad**: La contenerización brinda la capacidad de implementar el sistema en infraestructuras variables y zonas rurales sin conectividad.

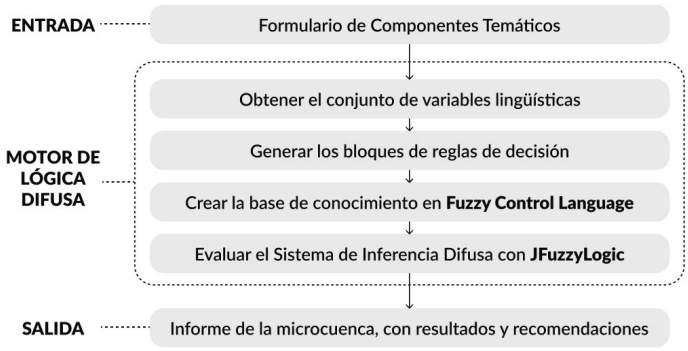


Fig. 34.3: Motor de lógica difusa desarrollado

- **Flexibilidad**: El sistema permite crear y modificar modelos de lógica difusa de manera dinámica, ajustando los parámetros para reflejar condiciones locales de manera precisa.

34.5 Conclusiones

El sistema desarrollado ofrece una solución innovadora para la gestión de recursos hídricos, permitiendo adaptar modelos de lógica difusa sin necesidad de contar con conocimientos en programación. Su arquitectura modular y la contenerización facilita su accesibilidad e implementación en diversos contextos.

Referencias

1. Vargas Franco, V. (2014). Modelo basado en conocimiento para la planeación de cuencas hidrográficas con el uso de inteligencia artificial. Universidad del Valle.
2. Cingolani, P., & Alcalá-Fdez, J. (2013). jFuzzyLogic: a java library to design fuzzy logic controllers according to the standard for fuzzy control programming. International Journal of Computational Intelligence Systems, 6(sup1), 61-75.

35
Code as Art, Art as Code

Nicolás Londoño [1]
Nicolás Cardozo[2]

35.1 Introduction

The typewriter interface, through the keyboard, has been the reigning interface to materialize programs into code since the beginning of programming. Interface designs lack what is needed to support every user. Naturally, users are not created equally. Work in the user experience and human-interface integration show reports in which technology and its designs can discriminate against users due to several factors, such as different physical or cognitive capacities, race and gender [3]. Moreover, programming has a steep learning curve and a high entry level [2]. Different approaches to computing are needed to lower the entry threshold and make computing more approachable to everyone. Current programming languages are based on the same principles that we have used since the 50s and centered around the typewriter. To lower the entry threshold of programming and computing, more approachable interfaces and methods are required. New fields in computing are emerging with different intents but still the same interfaces [4]. To explore new methods of computing, we must look further ahead than the tools we use, into how we approach code in general. To do so, we must shift attention to other mediums that can be used for computation, yet serve a complete different purpose. Such path may lie in the realm of art. Using coding tools to create art is not a new concept. Many artists today embrace code in their practice. For example, Sonic Pi is a coding suite that allows users to create music using code tailored for live coding performances [1]. The Simple Cubic Language for Programming Tasks (SCuLPT) is an artistic sandbox that ensures correct computation, shifting the focus away from computation, yet relying heavily on it to create physical sculptures that, by design, compile to valid programs.

35.2 Objectives

In the creation of SCuLPT we posit two objectives. First, we require to build a fully capable programming language addressing the concerns described previously about the state of programming, computation and programming languages. Second, we aim to evaluate the language in two main aspects. First, we intend to evaluate the ease of the language to foster computational thinking without the intrinsic difficulties of current programming languages. Second, we evaluate the properties

[1,2]Universidad de los Andes
e-mail: \{n.londonoc,n.cardozo\}@uniandes.edu.co

of SCuLPT as a paradigm that breaks existing preconceptions about programming languages and moves beyond the typewriter interface by leveraging artistic creativity as an alternative.

35.3 Methodology

SCuLPT is in fact two languages, one being the physical specification that defines the shapes and the other being the programming language that is used to define the behavior of each component. SCuLPT's focus on usability and accessibility is displayed by using simple but distinctive shapes and forms for each component on the language. This design choice allows for users to easily identify and differentiate between different components of the language, allowing people with any background to recognize each component foregoing knowledge, language or capacity barriers that arise from conventional programming languages. Code blocks are assembled together to create a program, with each block representing a different function or operation. Each block has a set of features that must be connected to other blocks in order to function correctly. SCuLPT is designed to be a Turing complete language. The language uses stacks that allow for Pop, Push, Move and Duplicate operations to store and manipulate data represented by any shape other than the reserved shapes used by the instructions. Rational numbers are the only type available and are defined by their literals. Arithmetic operations are performed in-place. SCuLPT allows for flow control through the use of loops of blocks, the Question block and Jump blocks. Regarding the language's aesthetics, SCuLPT is designed to be an open artistic sandbox hence the rules for how shapes should look are loose. This allows users to tinker with the components' properties. This not only provides a sense of freedom and creativity, but also implies that the language can be fitted to almost anyone's needs, allowing accessibility and customization for users with different physical or cognitive capacities.

35.4 Results

SCuLPT as a programming language is a functioning language capable of writing general programs. As an example, we have an implementation of Fibonacci in SCuLPT. SCuLPT 's visual aspect makes it such that a program can take any shape the user intends, as long as the connections are respected. This allows for a high degree of creativity and personalization in the code, making it a unique and personal piece of artistic expression. Notably, our Fibonacci program runs forever. However, it can be bound by adding a Question block and another stack to count the number of iterations desired. As showcased, the range of operations, while limited, allows for total control of a stack data structure, making it Turing complete with the assumption that stacks are infinite in size.

35.5 Conclusion

SCuLPT is an artistic sandbox and programming language that aims to lower the entry threshold to programming by leveraging artistic creativity as an alternative to the traditional typewriter interface. We have presented the design and implementation of Artlang, a language that allows users to create physical sculptures that compile to valid programs. Artlang's design choices make it possible for anyone to create programs without the need for prior programming knowledge, making it accessible and modifiable to a wide range of users. While the validations done so far are only preliminary, they show promising results in terms of usability and accessibility, while resulting in interesting pieces of art.

References

1. S. Aaron et al. 2016. The development of sonic pi and its use in educational partnerships: co-creating pedagogies for learning computer programming. Jour. of Music, Technology and Education, 9, 1.
2. Y. Bosse et al. 2017. Why is programming so difficult to learn? patterns of difficulties related to programming learning mid-stage. SIGSOFT Softw. Eng. Notes, 41, 6.
3. A. Ko. 2023. How my broken elbow made the ableism of computer programming personal. Nature.
4. H. Yang et al. 2016. Promoting creative computing: origin, scope, research and applications. Digital Communications and Networks, 2, 2.

36
Overview of article: Quantum Vision transformers for Quark-Gluon classification

Andrea Sánchez Jiménez Largo [1]
Laura Daniela Becerra [2]

36.1 Introduction

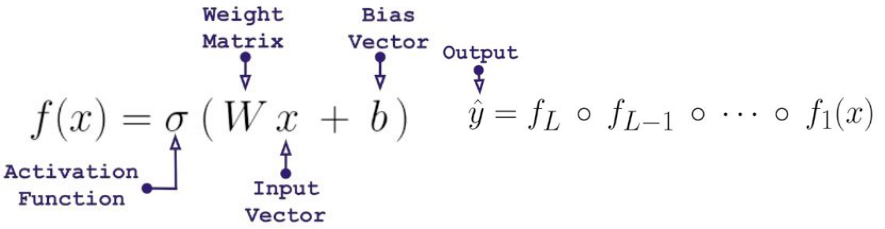
The imminent operation of the LH-LHC signals new challenges regarding the great amount of data it will generate. A promising approach to deal with this could be the application of Quantum Machine Learning (QML). In particular, a hybrid architecture of quantum-classical vision transformer is proposed here.This architecture integrates Variational Quantum Circuits (VQCs) into the attention mechanisms and Multi-Layer Perceptrons (MLPs) of the classical Vision Transformer (ViT) architecture.The proposed QViT is trained on multi-detector jet images from the CMS Open Data Portal, with the goal of discriminating between quark-initiated and gluon-initiated jets. The motivation to apply QML to this particular task stems from the limitations of current classical deep learning models, which are increasingly constrained by escalating computational demands inherent in processing large datasets.

36.2 Background

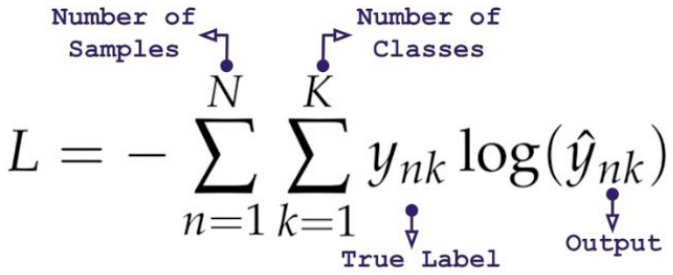
Deep Learning, the Transformer and the Vision Transformer: One of the most straightforward realizations of a neural network is a multi-layer perceptron (MLP), built in a feedforward configuration. Mathematically, MLPs are described as a composition of elementwise non-linearities with affine transformations (linear transformations followed by translation) of the data. The output

[1,2]Universidad Nacional de Colombia Sede - Medellin
Semillero de Investigación Quantum Computing UNAL
Medellín, Antioquia, 050034, Colombia.
e-mail: \{asanchezji,lbecerral\}@unal.edu.co

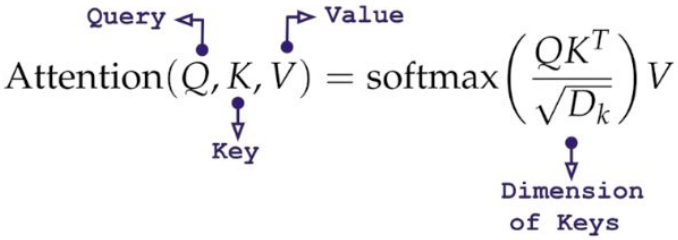
of a layer serves as the input of the next layer, until a final output is produced by the last layer.

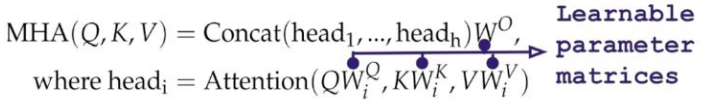


The learning process of the MLP involves adjusting the weights and biases in order to minimize a cost function. For classification tasks, a cost function often used is the cross-entropy loss function. It is defined as: Among minimization methods, Stochastic Gradient Descent and Backpropagation are



common alternatives. Multi-head self-attention mechanisms and MLPs are the basic components of a transformer building block. MHA allows the model to jointly attend to information from different representation subspaces at different positions.





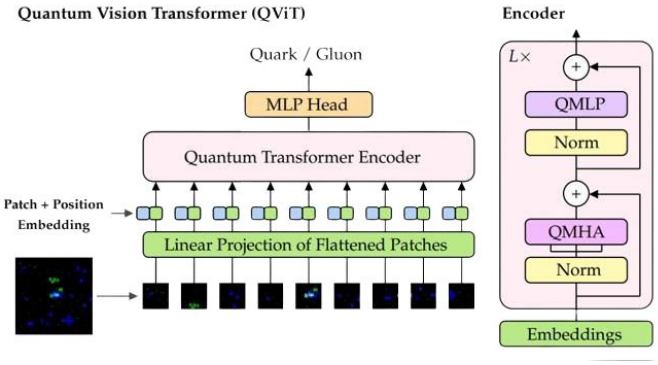
In ViTs, an image is split into a sequence of patches, which are linearly embedded and treated as input tokens for a transformer encoder.

Quantum Computing and Quantum Machine Learning: The Rx gate performs a single qubit rotation about the X axis in the Bloch sphere. The CNOT gate operates over two qubits, by flipping the target qubit only if the control qubit is 1 and leaving the control qubit unchanged. The main

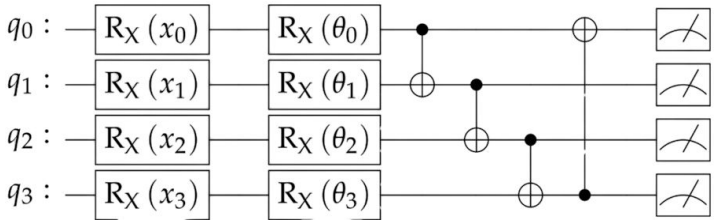
$$R_X(\theta) = \begin{bmatrix} \cos(\theta/2) & -i\sin(\theta/2) \\ -i\sin(\theta/2) & \cos(\theta/2) \end{bmatrix} \quad CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

idea behind QML is to use models executed on quantum computers, replacing subroutines of the model with quantum circuits. Such parameterized quantum circuits are called variational quantum circuits (VQCs). **High-Energy Physics and Jets:** In collider experiments, jets arise as a result of the hadronization of the fundamental elementary particles, which carry color charge, in particular quarks and gluons. The question of the precise origin of a given jet is highly non-trivial and remains a subject of active investigations in the literature.

36.3 Methods



The VQC configuration is as follows: the features of the input vector $x = (x_0, \ldots, x_{n-1})$ are embedded into the qubits by encoding them into their rotation angles. Then a layer of one-parameter single-qubit rotations acts on each wire. The parameters $\theta = (\theta_0, \ldots, \theta_{n-1})$ are learnable. A ring of $CNOT$ gates follows to entangle the qubit states, obtaining a behavior similar to a matrix multiplication. Finally, each qubit is measured, and the output is fed to the next component of the encoder. The proposed QViT and a classical ViT are trained with the same hyperparameters to have a meaningful baseline for comparison. Four transformer blocks with four attention heads



each and a hidden MLP size of four are used. These dimensions are small so that the circuits, which are simulated using classical computers, do not require many qubits. Of the data, 714510 images are allocated for training, 79390 for validation and 139306 for the test set. Receiver Operating Characteristic (ROC) curves and AUC scores are used to assess the classifiers performance. JAX and Flax are used to implement the classical parts of the hybrid model, as well as the classical baseline model, and to train both. For the VQCs, TensorCircuit is used, running numerical simulations.

36.4 Results

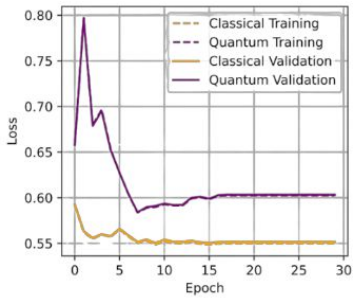


Fig. 36.1: Binary cross-entropy loss evolution.

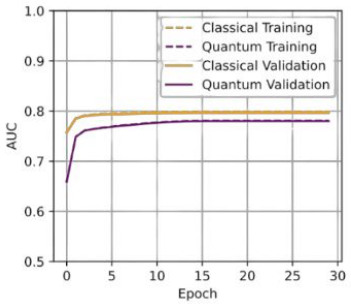


Fig. 36.2: AUC score evolution.

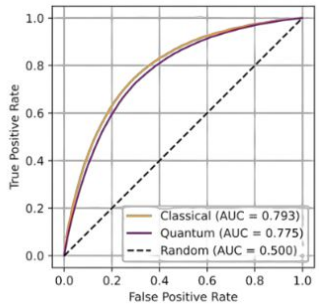


Fig. 36.3: ROC curves.

36.5 Conclusion

The trained model was benchmarked against a classical vision transformer with the same hyperparameters and a similar number of trainable parameters and was found to have comparable performance. The results achieved so far are encouraging and warrant future investigations. Acknowledgements.

Acknowledgements Acknowledgemet to project 52893 - Center of Excellence in Quantum Computing and Artificial Intelligence.

References

1. Comajoan Cara, M.; Dahale, G.R.; Dong, Z.; Forestano, R.T.; Gleyzer, S.; Justice, D.; Kong, K.; Magorsch, T.; Matchev, K.T.; Matcheva, K.; et al. Quantum Vision Transformers for Quark–Gluon Classification. Axioms, 13, 323 (2024).

37
Deep Learning Methods for Image Steganography: New Approaches to Concealing Videos Within Videos

Omar Esteban Vargas Salamanca [1]

37.1 Introducción

La esteganografía, el arte de ocultar mensajes dentro de un medio para asegurar que permanezcan inadvertidos, tradicionalmente implica la inserción de texto en imágenes. Los avances recientes han ampliado estas técnicas para incluir medios digitales como imágenes, audio y video. Inspirados en el trabajo de Baluja sobre esteganografía de imagen en imagen, nuestra investigación introduce la esteganografía de video en video utilizando redes neuronales convolucionales (CNNs). Este enfoque mejora la seguridad y la calidad en la transmisión de datos visuales, avanzando más allá de las técnicas convencionales.

37.2 Objetivos

37.2.1 Objetivo General:

Desarrollar y evaluar un sistema de esteganografía de video en video con redes neuronales convolucionales (CNNs) que permita ocultar y recuperar videos completos con alta calidad.

37.2.2 Objetivos Específicos:

- Diseñar una arquitectura de autoencoder para ocultar información visual en videos de manera indetectable. Optimizar el modelo y evaluar su rendimiento con la métrica SSIM.
- Comparar la eficacia del sistema con otros métodos de esteganografía existentes.

[1] Universidad de los Andes, FLAGLAB
e-mail: o.vargas@uniandes.edu.co

37.3 Materiales y Métodos

Datos: Conjunto de datos ImageNet para el entrenamiento de modelos.
Arquitectura del Modelo: Una arquitectura de autoencoder con Red de Preparación, Red de Ocultación (Encoder) y Red de Revelación. El modelo procesa fotogramas de video, incrustando información secreta dentro de videos de cubierta.
Entrenamiento: Más de 6000 imágenes redimensionadas a 124x124 píxeles, empleando técnicas de aumento de datos.
Métrica de Evaluación: SSIM para evaluar la integridad y calidad de los videos con esteganografía.

37.4 Resultados

Valores de SSIM: El SSIM del video de cubierta se mantuvo consistentemente por encima de 0.98, mientras que el SSIM del video secreto estuvo entre 0.86 y 0.93. Rendimiento: Altas tasas de inserción de datos con mínima distorsión perceptual. Aplicación: Se demostró la esteganografía de video en video con videos de 243 fotogramas, mostrando excelentes capacidades de ocultación y recuperación.

37.5 Conclusiones

Eficacia del Modelo: El sistema desarrollado basado en CNNs ha demostrado ser efectivo para la esteganografía de video en video, logrando ocultar y recuperar videos con alta calidad perceptual.
Dificultad en la Detección: Construir un identificador de esteganografía usando deep learning que detecte efectivamente el uso de este modelo es un desafío considerable, debido a la sutileza de las modificaciones introducidas en los videos ocultos.
La técnica propuesta ofrece mejoras significativas en comparación con métodos tradicionales, especialmente en la capacidad de ocultación y la calidad de recuperación de datos.

Referencias

1. Baluja, S. ”Hiding Images in Plain Sight: Deep Steganography.” In Advances in Neural Information Processing Systems, 2017.
2. Hashemi, S. H. O., Majidi, M. H., & Khorashadizadeh, S. ”Color Image Steganography Using Deep Convolutional Autoencoders Based on ResNet Architecture.

38
LS-SEG: Semi-Automatic Segmentation Methods for Lumbar Spine MRI in 3D Slicer

William A. Romero R. [1]
Susana Uribe Velásquez[2]
Daniel Restrepo Quiñones[3]
Duván A. Gómez[4]

38.1 Introduction

LS-SEG (Fig. 1) is a module extension for 3D Slicer [1]. This extension includes methods for vertebrae segmentation and Cobb angle calculation. The main LS-SEG algorithm requires only two points on the image for Cobb angle calculation; the traditional method requires four: two for the tangent line to the upper end-plate of the upper vertebra and two for the tangent line to the lower end-plate of the lower vertebra. Such efficiency is vital when managing images from a substantial patient cohort in a clinical study.

38.2 Objetive

Develop and deploy a module extension that enables rapid segmentation of vertebrae in lumbar MR images, with the capability to calculate the Cobb angle between the segmented vertebrae.

38.3 Materials and Methods

- Sagittal fast spin-echo images of the lumbar (Esaote G-scan Brio 0.25T).
- Imaging parameters: $\frac{TE}{TR} = \frac{125}{3000} ms$, in-plane resolution/slice $thickness = \frac{0.6}{4} mm$, echo train $length = 6$.
- Cobb angle:

$$\theta = tan_{-1} \left(\frac{|m_1 - m2|}{1 - m_1 m_2} \right) \tag{38.1}$$

- Vertebrae segmentation performed by a U-Net neural network [2] trained to predict a vertebra within a sub-image (Fig. 38.2).
- LS-SEG was developed in Python using SimpleITK [3] and TensorFlow [4].

[1]University of Montpellier, CARTIGEN, CHU de Montpellier, Montpellier, France
[2,3,4]Universidad EIA, Envigado, Antioquia, Colombia
e-mail: \{susana.uribe7,daniel.restrepo53,duvan.gomez36\}@eia.edu.co

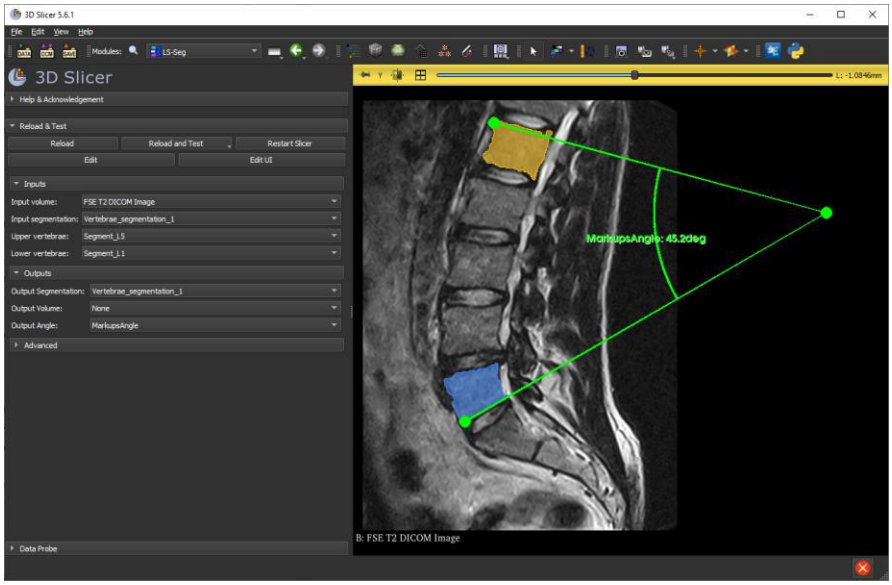


Fig. 38.1: LS-SEG Workspace in 3D Slicer.

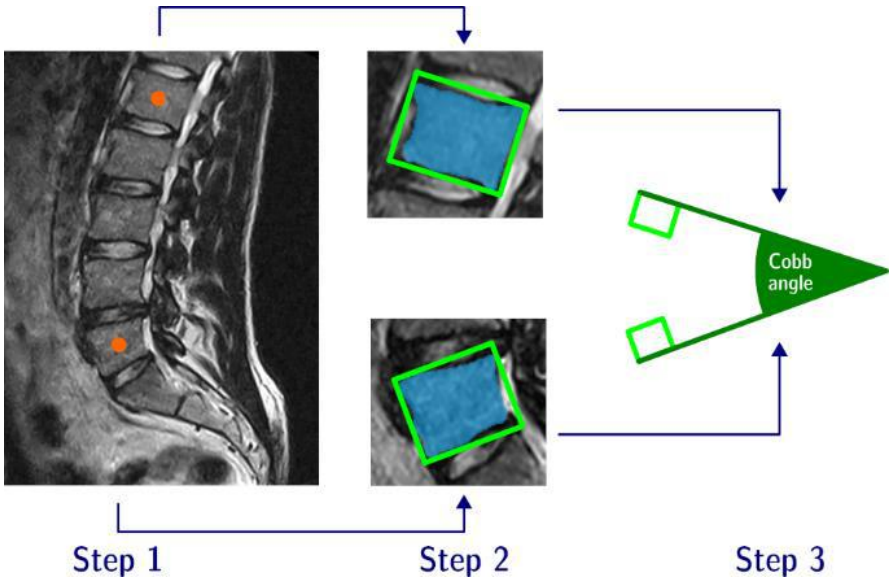


Fig. 38.2: Workflow for Cobb angle calculation. Step 1: Vertebrae selection (2 mouse click), Step 2: Sub-image extraction, vertebrae segmentation and bounding-box calculation, Step 3: tangent lines calculation and Cobb angle calculation.

38.4 Conclusion

LS-SEG is a module extension for 3D Slicer, primarily designed to compute the Cobb angle between two user-selected vertebrae. The algorithm includes a semi-automatic vertebral segmentation step as part of the geometric definition required for calculating the Cobb angle. LS-SEG. The source code and production version are available at https://github.com/CARTIGEN/LS-SEG.

References

1. Fedorov, A., et al.: 3d slicer as an image computing platform for the quantitative imaging network. Magnetic resonance imaging (2012).
2. Ronneberger, O., et al.: U-net: Convolutional networks for biomedical image segmentation. MICCAI 2015, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. pp. 234-241 (2015).
3. Lowekamp, B.C., Chen, D.T., Ibáñez, L., Blezek, D.: The design of simpleitk. Frontiers in neuroinformatics 7, 45 (2013).
4. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., et al.: Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467 (2016).

39

Arquitectura de software de un sistema de recomendación de inmueble

Oñate Acuña Diego David [1][0009-0001-0279-7274]
Saballeth Lora Camilo Andres [2] [0009-0009- 4876-0440]
Henríquez Miranda Carlos Nelson[3] [0000-0002-0439-4954]

Resumen En el proyecto se desarrolla un sistema de recomendación de inmuebles que ayude a los usuarios a encontrar propiedades adecuadas mediante el uso de la técnica de filtrado colaborativo. Dada la gran cantidad de información y la complejidad de las preferencias individuales, los sistemas de recomendación utilizan algoritmos avanzados para personalizar las búsquedas, basándose en patrones de comportamiento y preferencias previas. Entre las técnicas comunes se encuentran el filtrado basado en contenido, el filtrado basado en conocimiento y el filtrado colaborativo, siendo el enfoque híbrido el que combina varias de estas técnicas para mejorar la precisión. El sistema emplea el filtrado colaborativo basado en similitud de coseno para generar recomendaciones personalizadas. La encuesta realizada a los usuarios reveló una alta satisfacción general, sin embargo, se sugirieron mejoras. El modelo desarrollado muestra capacidad para enfrentar el desafío del arranque en ausencia de datos de usuario al solicitar prefe- rencias desde el registro del usuario, y se recomienda explorar técnicas híbridas y aprendizaje profundo para mejorar aún más la precisión y adaptabilidad del sistema en el futuro.

Palabras clave Sistema de recomendación, filtrado colaborativo, inmueble.

39.1 Introducción

En la búsqueda diaria de propiedades, encontrar el lugar adecuado se convierte en un proceso complejo debido a los exigentes requisitos de las personas, la diversidad de aplicaciones y fundamentalmente, debido a la gran cantidad de información para revisar. En este contexto, los usuarios dedican cada vez más tiempo a interactuar con aplicaciones que les proporcionen información relevante.

Un componente fundamental utilizado para mitigar el tiempo de búsqueda en las plataformas son los sistemas de recomendación, estos emplean algoritmos avanzados para prever y ofrecer elementos que puedan interesar al usuario basados en unos requerimientos o perfilamientos previos. Estos sistemas analizan patrones de comportamiento, preferencias históricas, similitudes entre usuarios y elementos para personalizar la experiencia de búsqueda [1]. Dado el continuo crecimiento de la

[1,2,3]Universidad del Magdalena, Magdalena, Colombia
Semillero de Investigación del Grupo de investigación y Desarro- llo en Sistemas y Computación - GIDSYC.
e-mail: chenriquezm@unimagdalena.edu.co

información en línea, la capacidad de filtrar y seleccionar contenido relevante se vuelve crucial, ya que simplifican la toma de decisiones y aumentan la satisfacción del usuario [2]. Como resultado, en la presente investigación se realizó un sistema de recomendación de inmuebles basado en filtrado colaborativo con el objetivo de mejorar la búsqueda diaria de propiedades. En el proceso de construcción del sistema se tuvieron limitaciones debido a la gran cantidad de datos e interacciones a tener en cuenta, una de las limitaciones fue una base de datos con gran capacidad computacional, asimismo, el problema del cold-start o arranque en frio se le dio solución de forma que se le preguntaba al usuario al momento de registrarse sus preferencias.

39.2 Objetivos

Desarrollar una arquitectura de software de un sistema de recomendación de inmuebles para ayudar a los usuarios a tomar decisiones basadas en sus preferencias mediante la técnica de filtrado colaborativo.

39.2.1 Objetivos específicos

1. Revisar el estado de la literatura de proyectos referentes a sistemas de recomendación en el sector inmobiliario para la identificación de herramientas y tecnologías empleadas para estos.
2. Determinar las herramientas, los componentes clave y sus interacciones para un sistema de recomendación en el sector inmobiliario.
3. Implementar el modelo e interfaz web, asegurando la coherencia con la arqui- tectura definida, creando una experiencia de usuario intuitiva.

39.3 Materiales y métodos

El proyecto se dividió en las fases mostradas en la Fig. 1. Donde se inicia con la carac- terización del proyecto esta fase implica realizar una búsqueda y el análisis de herra- mientas, luego se diseña la arquitectura, en esta fase se decidieron los distintos módulos y se escogieron las herramientas tecnológicas, por consiguiente, se implementa esta misma para la cual se recopilaron y limpiaron los datos de acuerdo con la investigación realizada y las herramientas seleccionadas. A continuación, se procedió a implementar la arquitectura del sistema y a construir la base de datos, estableciendo la conexión entre ambos para asegurar una integración efectiva, y por último se despliega el sistema de recomendación para esta fase se estudiaron las tecnologías más adecuadas que puedan soportar la aplicación. Por otro lado, para la página web se evaluaron distintos hostings con el objetivo de buscar una plataforma que ofrezca fiabilidad, escalabilidad y un rendimiento óptimo para el despliegue exitoso del sistema. En la fase de caracterización, se seleccionaron las herramientas

Fig. 39.1: Fases del proyecto

más adecuadas para el proyecto, optando por el filtrado colaborativo como la mejor opción. Este enfoque se basa en un modelo supervisado [3], el cual, según [4], funciona entrenando un algoritmo para que aprenda a mapear entradas a salidas deseadas. Durante este proceso, el algoritmo se entrena utilizando un conjunto de datos etiquetados, donde cada entrada está asociada con su correspondiente salida correcta o esperada. Para la fase de diseño, se eligieron los módulos (ver Figura 39.2) y herramientas del sistema, el anterior mencionado se compone de varios módulos que trabajan juntos para dar recomendaciones precisas. Primero, el módulo de webscraping recopila los datos relevantes de fuentes en línea, la cual fue selecciona la página de Metrocuadrado por su extensa información detallada sobre inmuebles. Luego, el módulo de procesamiento de lenguaje natural (PLN) limpia y estructura estos datos, como resultado se obtuvieron 49,191 inmuebles. Después, los datos procesados se almacenan en una base de datos a través de esta arquitectura, estos se usarán para entrenar el modelo y se consumen por el aplicativo web. Finalmente, el front-end se integra con la arquitectura para presentar las recomendaciones al usuario en una interfaz amigable. Para complementar el análisis del sistema de recomendación de inmuebles, se realizó una encuesta dirigida a usuarios que han utilizado la plataforma para buscar propieda- des. El objetivo de la encuesta fue evaluar la percepción de los usuarios sobre la efec- tividad y utilidad del sistema de recomendación implementado, la encuesta fue completada por 19 personas residentes en la ciudad de Santa Marta. A continuación, se presentan los resultados más destacados. Al observar la Fig. 39.3, los usuarios fueron encuestados acerca de "¿Qué tan conforme se siente con la experiencia de manejo del sistema? ¿Qué tan fácil fue el manejo?" los cuales respondieron que muy bueno el 58%, bueno el 34% y neutral el 8%.

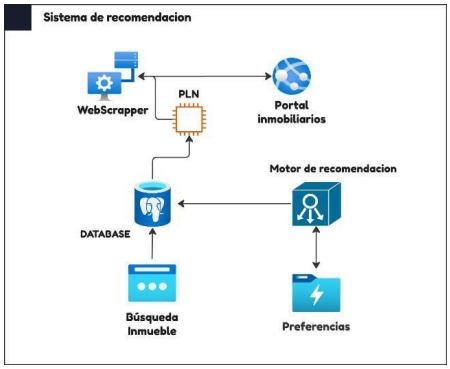


Fig. 39.2: Módulos de la aplicación. Fuente: Elaboración de los autores

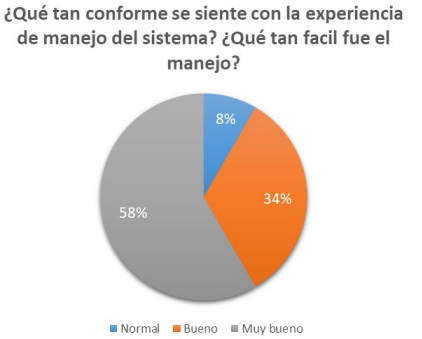


Fig. 39.3: Resultados encuesta de satisfacción. Fuente: Elaboración de los autores

39.4 Conclusiones

Para concluir, el modelo de recomendación desarrollado se destaca por su capacidad para abordar el problema del cold-start, solicitando las preferencias de los usuarios justo después de registrarse. Este enfoque ha permitido ofrecer recomendaciones precisas desde el primer momento, diferenciando este sistema de otros. Como trabajo futuro, se propone explorar técnicas híbridas y algoritmos de aprendizaje profundo para seguir

mejorando la precisión y adaptabilidad del sistema en función de las preferencias cam- biantes de los usuarios. Con estos avances, esperamos que el sistema no solo mantenga

su ventaja competitiva, sino que también ofrezca recomendaciones aún más precisas y contextualizadas en un entorno de datos en constante crecimiento.

Agradecimientos A los miembros del Grupo de Investigación y Desarrollo en Sistemas y Computación, de la Universidad del Magdalena por el invaluable apoyo y orientación durante el desarrollo del presente proyecto. Su experiencia y dedicación no solo han enriquecido nuestras investigaciones, sino que también han sido fundamentales para superar los desafíos técnicos y conceptuales que enfrentamos.

Referencias

1. K. T. Méndez Zapata y F. R. Quintuña Barbecho, “Desarrollo de un sistema de recomendación generalista basado en filtrado colaborativo para grupos de usuarios mediante un enfoque de agregación de factores probabilísticos”, bachelorThesis, 2021. Accedido: 13 de agosto de 2024. [En línea]. Disponible en: http://dspace.ups.edu.ec/handle/123456789/20460
2. D. Varela, J. Aguilar, J. Monsalve-Pulido, y E. Montoya, “Propuesta Arquitectónica de un Sistema de Recomendación Híbrido Adaptativo”, RISTI - Rev. Iber. Sist. E Tecnol. Inf., vol. E38, pp. 171-184, dic. 2020.
3. I. Portugal, P. Alencar, y D. Cowan, “The use of machine learning algorithms in recommender systems: A systematic review”, Expert Syst. Appl., vol. 97, pp. 205-227, may 2018, doi: 10.1016/j.eswa.2017.12.020.
4. Computational model of supervised machine learning classification, for the analysis of cardiovascular data and medical prognosis — Ecuadorian Science Journal”. Accedido: 13 de agosto de 2024. [En línea]. Disponible en: https://journals.gdeon.org/index.php/esj/article/view/83

40
Identificación de recursos hídricos en la superficie de Marte mediante percepción remota

Santiago Chaves Córdoba[1]
Edwin Francisco Valdes Arias [2]

40.1 Introducción

La percepción remota es una herramienta fundamental en la investigación espacial, permitiendo a científicos y exploradores recopilar datos sobre planetas, estrellas y otros cuerpos celestes sin necesidad de estar presentes físicamente. Los satélites equipados con sensores remotos, como cámaras, telescopios y radares, proporcionan información vital sobre el clima, la atmósfera, la geología y los recursos naturales, lo cual es esencial para la investigación científica y la gestión de recursos.

En el presente trabajo se aplicaron técnicas de teledetección para identificar áreas prometedoras con presencia de agua en el Valle Marineris (Marte). Utilizando datos del instrumento HiRISE a bordo de la nave espacial Mars Reconnaissance Orbiter (2005), se calcularon índices radiométricos como el NDVI, NDWI e ICIRON. Los resultados mostraron cuatro clases distintas en las imágenes analizadas, sugiriendo la posible presencia de agua. Se implementaron diferentes algoritmos de aprendizaje automático como K-means, Decision Tree, Random Forest, Artificial Neural Networks y Support Vector Machine, con el objetivo de clasificar las distintas clases identificadas en la zona de estudio. La validación de los modelos demostró que las redes neuronales artificiales presentaron la mayor exactitud de clasificación (98%) en comparación a los demás modelos.

40.2 Objetivos

- Recopilar datos de Marte para detectar áreas prometedoras con presencia de agua.
- Aplicar técnicas de teledetección y algoritmos de clasificación en la zona de estudio.
- Evaluar y validar los resultados de los diferentes modelos desarrollados.

[1,2]Universidad del Cauca
Grupo de Investigación Óptica y Láser (GOL)
e-mail: \{santchaves,evaldes\}@unicauca.edu.co

40.3 Materiales y Métodos

La presente investigación se realizó haciendo uso de datos descargados desde NASA Planetary Data System (PDS), donde se utilizó el instrumento High Resolution Imaging Science Experiment (HiRISE) el cual está a bordo de la nave espacial Mars Reconnaissance Orbiter (MRO), misión que fue lanzada el 12 de agosto de 2005 para el avance del conocimiento humano de Marte a través de la observación detallada. Las imágenes descargadas corresponden al área del Valles Marineris, lugar de alto interés científico por sus características de sistema de cañones de gran extensión en el ecuador de la superficie de Marte. Las imágenes contenían bandas en RGB e infrarrojo cercano que permitieron calcular los índices radiómetros como el NDVI, NDWI e ICIRON, procesamiento y cálculos que se realizaron haciendo uso de Matlab y sus funciones específicas para el procesamiento de imágenes.

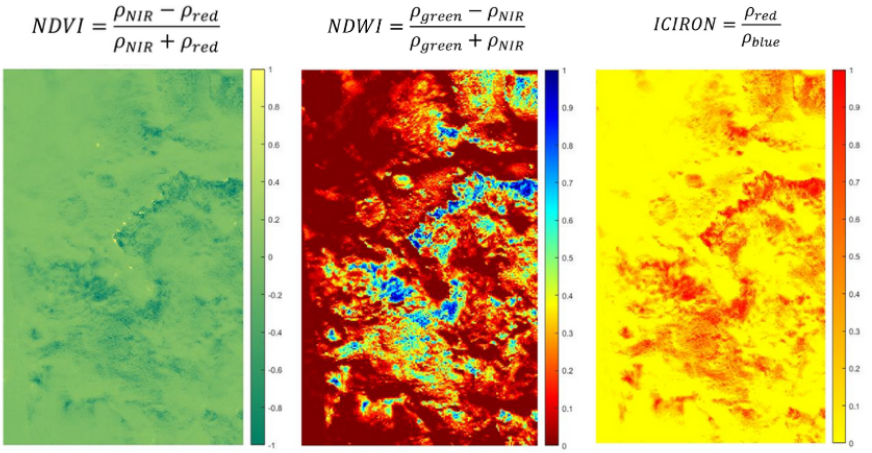


Fig. 40.1: Cálculo de los índices radiométricos. Derecha NDVI, Centro NDWI y Izquierda ICIRON.

La metodología implementada para el desarrollo de la investigación se basó en CRISP-DM, haciendo uso de las diferentes etapas acorde a los objetivos planteados.

Los algoritmos utilizados incluyeron modelos de aprendizaje supervisado y no supervisado como K-means, Decision Tree, Random Forest, Artificial Neural Networks y Support Vector Machine, con los cuales se realizó la clasificación de las 4 clases identificadas de acuerdo a los resultados de índice radiométrico NDVI (40.1 centro).

40.4 Resultados

Una vez se implementaron los modelos de aprendizaje automático en la clasificación de las clases identificadas, se procedió a realizar la validación por matriz de confusión de cada uno de los modelos, para medir la exactitud.

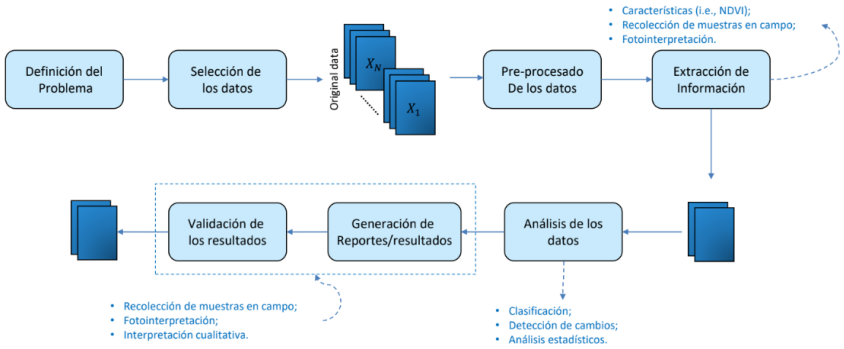


Fig. 40.2: Metodología implementada en la investigación.

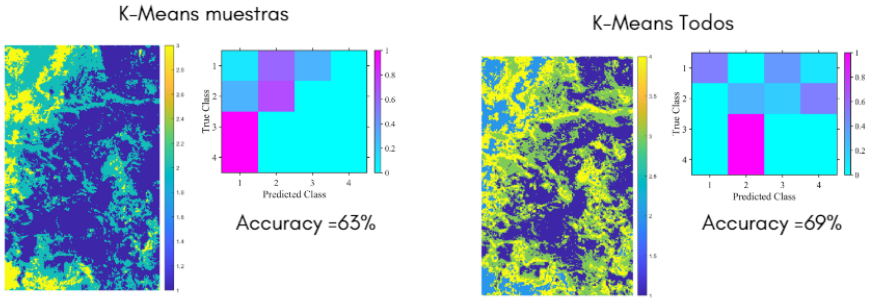


Fig. 40.3: Resultados de validación de modelos. A la derecha K-Means Muestras y a la Derecha K-Means Todos.

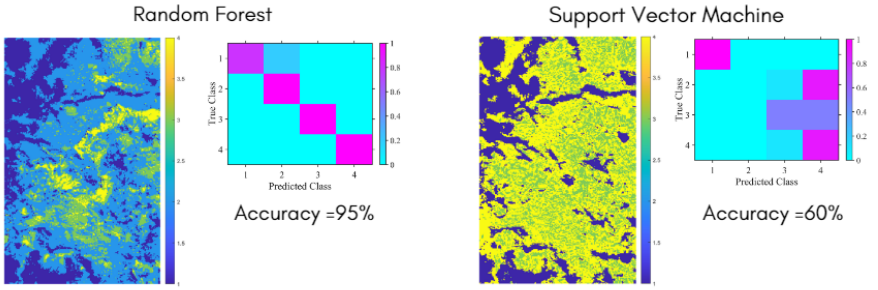


Fig. 40.4: Resultados de validación de modelos. A la derecha Random Forest y a la Derecha Support Vector Machine.

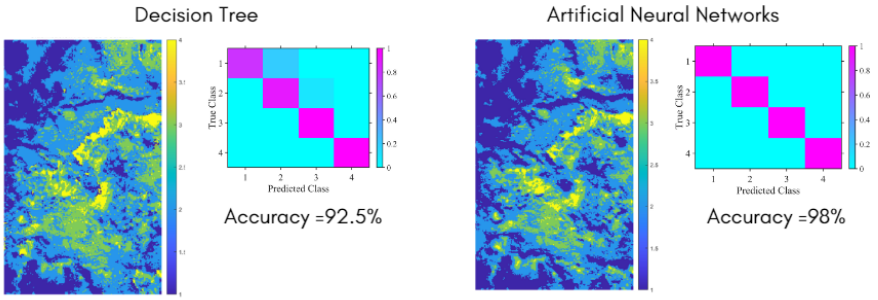


Fig. 40.5: Resultados de validación de modelos. A la derecha Decision Tree y a la Derecha Artificial Neural Networks.

40.5 Conclusiones

Las redes neuronales artificiales demostraron ser el algoritmo con mejores resultados en la evaluación realizada para clasificar áreas con posible presencia de agua en Marte, alcanzando una exactitud del 98%. Esta investigación resalta la efectividad de las técnicas de percepción remota y machine learning en la identificación de recursos hídricos fuera de nuestro planeta, proporcionando una base para futuras exploraciones y estudios científicos en el planeta rojo.

Referencias

1. Arnaud Valeille, Stephen W. Bougher, Valeriy Tenishev, Michael R. Combi, Andrew F. Nagy, Water loss and evolution of the upper atmosphere and exosphere over martian history, Icarus, Volume 206, Issue 1, 2010, Pages 28-39.
2. Michael J. Wolff, R. Todd Clancy, Melinda A. Kahre, Robert M. Haberle, François Forget, Bruce A. Cantor, Michael C. Malin, Mapping water ice clouds on Mars with MRO/MARCI, Icarus, Volume 332, 2019, Pages 24-49.
3. Rajaneesh, C.L. Vishnu, T. Oommen, V.J. Rajesh, K.S. Sajinkumar, Machine learning as a tool to classify extra-terrestrial landslides: A dossier from Valles Marineris, Mars, Icarus, Volume 376, 2022.

41
Búsqueda in silico de dianas fisiológicas en *Aedes aegypti* para el uso de compuestos biorracionales

Jimenez N Arancibia [1]
Mantilla A Javier G[2]

41.1 Introducción

El mosquito *Aedes aegypti* es un vector de gran importancia en la transmisión de enfermedades como el dengue, zika, chikungunya, fiebre amarilla y fiebre de Mayaro. (PAHO 2016). El dengue se destaca como el virus más letal debido a las graves afectaciones a la salud humana y las dificultades en su control. El manejo de esta problemática se ha convertido en una prioridad, ya que los métodos de control actuales, basados en productos químicos, presentan limitaciones, como la resistencia genética de los mosquitos a los insecticidas y el impacto negativo en el medio ambiente y la salud humana. Por ello, se están explorando alternativas más sostenibles y menos contaminantes, entre las cuales destacan las estrategias de prevención, el control cultural y el control biológico (FAO, 2017), otras alternativas incluyen las moléculas biorracionales las cuales tienen promisorios usos bioinsecticidas, bajo o nulo impacto en enemigos naturales de las plagas, mínima o nula toxicidad para los seres humanos y el medio ambiente, así como por su ausencia de residualidad. (Cruz T, et al, 2020). Entre las moléculas biorracionales, los aceites esenciales derivados de plantas se presentan como una opción viable y se han propuesto como bioinsecticidas potenciales. Para predecir la eficacia del modo de acción de estas moléculas, se recurre al enfoque in silico, que emplea herramientas bioinformáticas y técnicas de modelado de proteinas, permite estudiar su interacción molecular con moléculas derivadas de aceites esenciales (Vargas A, et al 2020) y predecir su potencial bioinsecticida. La técnica computacional mencionada se denomina, docking molecular, la cual tiene aplicaciones en la predicción entre una molécula (ligando) y una proteína (receptor).

[1,2]Universidad Católica de Manizales
e-mail: \{Arancibia.jimenez,jmantilla\}@ucm.edu.co

41.2 Objetivos

41.2.1 Objetivo general:

Evaluar in silico las interacciones moleculares entre dianas fisiológicas de *Aedes aegypti* y nuevos compuestos con potencial bioinsecticida derivados de extractos vegetales, aceites esenciales y compuestos sintéticos derivados de lactonas.

41.2.2 Objetivos Específicos

- Caracterizar estructural y funcionalmente las proteínas escogidas, modelar la estructura tridimensional 3D por homología de las proteinas JHBP (juvenile hormone-binding protein), GABA (Gamma-aminobutyric acid) y la enzima AChE (acetylcholinesterase) de *Aedes aegypti*.
- Realizar simulaciones de interacciones ligando-receptor entre psoralen y bergapten, germacrene D, cannabidiol y compuestos derivados de lactonas composto 4 y rac-8, y los receptores JHBP, GABA y ACHE en *Aedes aegypti*.
- Evaluar la selectividad y especificidad de todos los compuestos mediante análisis de las energías de afinidad e identificar candidatos prometedores para futuros bioinsecticidas.

41.3 Materiales y Métodos

Se usaron varios programas bioinformáticos, se utilizó Protparam de *Expasy*[1] para calcular parámetros fisicoquímicos de las proteínas, los dominios, motivos y patrones, se determinaron con la base de datos Prosite usando ScanProsite, se utilizó Hmmer v 3.3.2 para corroborarlos y buscar homólogos. Se descargaron las secuencias fastas de *Aedes aegypti* del NCBI. El modelado de proteínas se realizó usando Phyre 2 y Swiss protein, con Pymol se descargaron las moléculas biorracionales en formato pdb. Usando Autodock Tools 1.5.7.44, se realizaron los docking moleculares, y se generaron estructuras 2D y3D en Discovery Studio, para evaluar sus enlaces.

41.4 Resultados

Entre Begapteno y Rac 8, la energía de afinidad para *Aedes aegypti* es alta. Las diferencias en la energía de afinidad indica una mayor selectividad de estas moléculas hacia *Aedes aegypti*, sugiriendo que podrían ser utilizadas específicamente para el control de este mosquito, reduciendo al mismo tiempo los efectos adversos sobre otras especies de insectos. Se observan enlaces a serina, leucina y alanina que le da estabilidad al complejo molécula-diana, ya que estos aminoácidos proporcionan puntos de anclaje y estabilización en la interfaz de interacción entre la molécula y su ligando o proteína asociada debido a la combinación de interacciones fuertes y específicas con aminoácidos clave.

En estos 2 acoplamientos se observa que la combinación de enlaces alquilo más débiles y el enlace convencional carbono-hidrógeno más fuerte sugiere que la interacción general es estable y significativa. Ya que con las otras moléculas la energía de afinidad es menor de -5. La interacción

[1] (http://web.expasy.org/protparam)

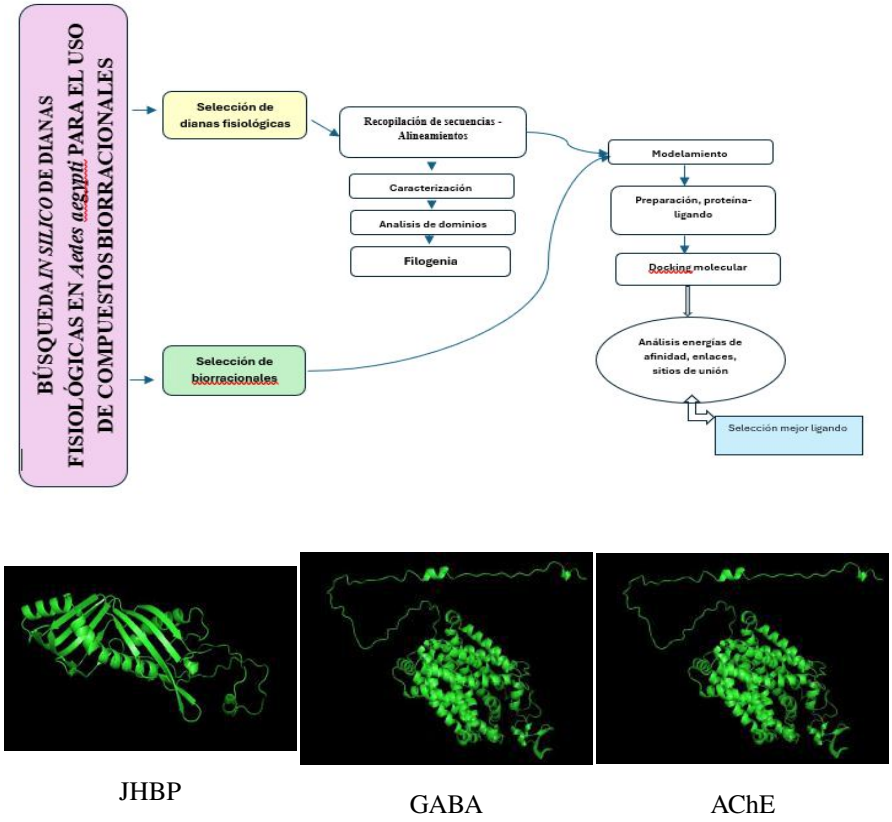


Fig. 41.4: Modelado de proteínas

con aminoácidos como arginina y alanina muestra que la molécula se une en sitios funcionalmente relevantes de la proteína. ACHE-JHBP pueden ser una diana efectiva para moléculas biorracionales.

41.5 Conclusiones

Los estudios in-silico desempeñan un papel crucial en la investigación y desarrollo de insecticidas biorracionales, dan información valiosa y las herramientas computacionales ahorran tiempo y costos para el diseño racional y la optimización de moléculas como posibles insecticidas efectivos, selectivos y respetuosos el medio ambiente. Las moléculas estudiadas se han considerado insecticidas, tienen potencial, pero no todas para *Aedes aegypti*. La Molécula Rac 8, derivada de lactonas, en un estudio previo, ha mostrado un alto potencial como insecticida al unirse de forma óptima a las proteínas AChE y JHBP en *Aedes aegypti*. Este hallazgo es significativo ya que ambas proteínas juegan roles críticos en el desarrollo y la supervivencia de este vector.

Nota: Este trabajo es resultado parcial de una tesis de maestría, que contiene más insectos importantes para el hombre.

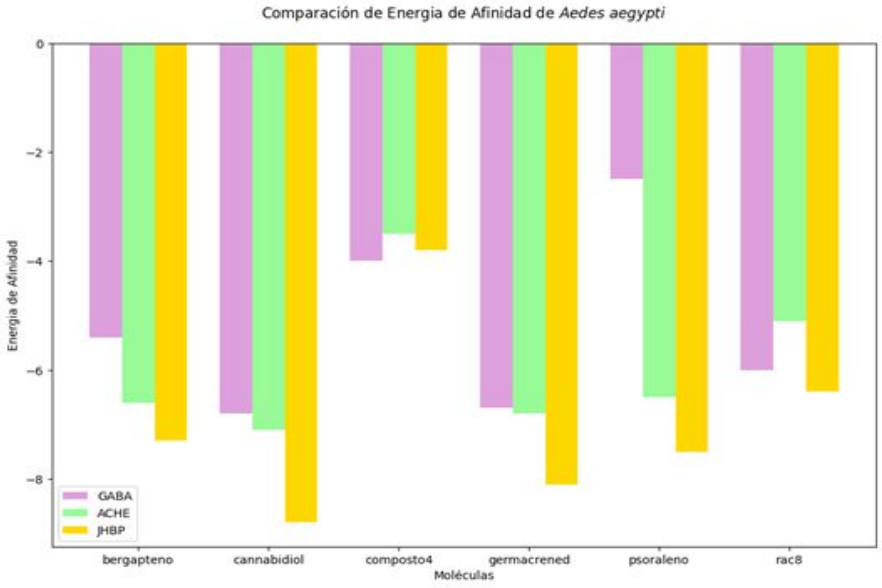


Fig. 41.5: Resultados Docking Molecular

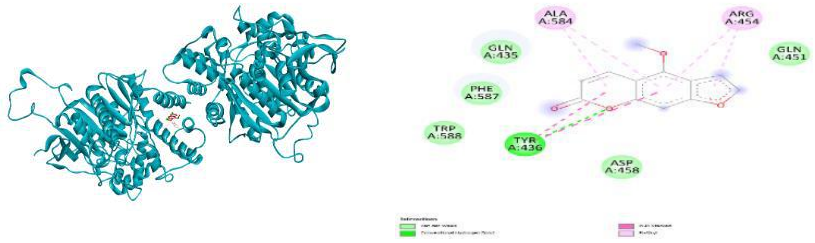


Fig. 41.6: Acoplamiento AChe *Aedes aegypti*- Bergapteno Energia de afinidad -6,6

Referencias

1. Britto Isabella O, Araújo Sabrina H C, Toledo Pedro F S, Lima Graziela D A, Salustiano Iorrana V, Alves Janaína R, Mantilla-Afanador Javier G, Kohlhoff Markus, Oliveira Eugenio E, Leite João Paulo V Potential Of Ficus Carica Extracts Against Euschistus Heros: Toxicity Of Major Active Compounds And Selectivity Against Beneficial • Insects. Pest Manag Sci. . 2021 Oct;77(10):4638-4647. Doi: 10.1002/Ps.6504. Epub 2021 Jun 28.
2. Mantilla Afanador J,G, Araujo S,H,C, Teixeira M,G, Lopes D,T, Cerceau C,I, Andreazza F, Oliveira D,C, Bernardi D, Moura W,S, Aguiar R,W,S, Oliveira A,C,S,S, Santos G,R, Alvarenga ,ES, Oliveira E,E. (2023). Novel Lactone-Based Insecticides and Drosophila suzukii Management: Synthesis, Potential Action Mechanisms and Selectivity for Non-Target Parasitoids. Insects. Aug 9;14(8):697. doi: 10.3390/insects14080697. PMID: 37623407; PMCID:

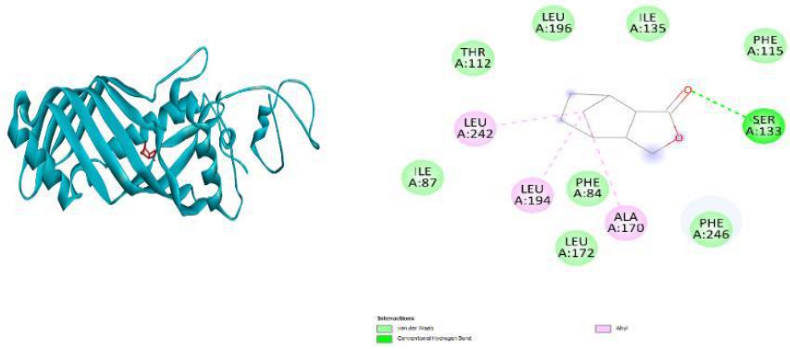


Fig. 41.7: Acoplamiento JHBP *Aedes aegypti*-Rac 8, energia de afinidad -6,40

PMC10455539.

3. Oña Genís, Balant Manica, Bouso José Carlos, Gras Airy, Vallès Joan, Vitales Daniel Y Garnatje Teresa. El Uso De Cannabis Sativa L. Para El Control De Plagas: Del Conocimiento Etnobotánico A Una Revisión Sistemática De Estudios Experimentales. Cannabis And Cannabinoid Research. 2021, Aug;7(4):365-387.
4. Sequeda Wilches Leydi Patricia, Sarmiento Miranda Juan Carlos, Ahumedo Monterrosa Maicol. Estudio In Silico De Moléculas De Origen Natural Con Potencial Actividad Inhibitoria Sobre La Enzima Acetilcolinesterasa. Tesis Facultad De Ciencias Farmacéuticas. Universidad De Cartagena. 2021.

42

Hermes: Un modelo de predicción de bajo rendimiento a partir de alertas tempranas basadas en técnicas computacionales de inteligencia artificial

Lilian Maradiago, [1]
Mariana Cruz [2]
Dillan Asprilla[3]

42.1 Introducción

La deserción estudiantil universitaria en Colombia representa un desafío significativo para el sistema educativo nacional. Este fenómeno, que se refiere al abandono de los estudios antes de la finalización de los programas académicos, tiene profundas repercusiones tanto para las instituciones educativas como para los estudiantes afectados. En particular, la facultad de ingenierías enfrenta una situación alarmante, como lo demuestran estudios que revelan tasas de deserción más elevadas en comparación con otras disciplinas [1], [2], [3]. Esta disparidad se atribuye a la alta exigencia académica y al nivel de abstracción inherente a las disciplinas técnicas y científicas que conforman estos programas. La elevada tasa de deserción en ingenierías no solo impacta negativamente los índices de graduación, sino que también plantea serias interrogantes sobre la optimización de recursos educativos y la retención de talento en el campo de la ingeniería en Colombia. Ante esta problemática, surge la necesidad urgente de desarrollar herramientas efectivas que permitan predecir y mitigar el riesgo de deserción estudiantil [4], [5]. Estas herramientas deben ser capaces de identificar patrones tempranos de bajo rendimiento académico y proporcionar alertas que faciliten intervenciones oportunas por parte de las autoridades educativas y los profesores [6].

42.2 Objetivos

42.2.1 Objetivos Generales

Desarrollar un modelo de alertas tempranas, denominado Hermes, basado en técnicas computacionales de inteligencia artificial, con el fin de disminuir la deserción académicas de los estudiantes.

[1,2,3]Universidad San Buenaventura de Cali
Semillero de Programación, Analítica de Datos e Iteligencia Artificial (PADIA)
e-mail: \{lilian.maradiago,ingmarianacruz,dillansdev\}@gmail.com

42.2.2 Objetivos específicos:

- Analizar el estado actual de los modelos utilizados para reducir la deserción estudiantil a través de alertas tempranas.
- Implementar un modelo de machine learning para predecir las calificaciones de un estudiante a partir del 50% del total de un curso.
- Desarrollar un modelo que permita generar alertas tempranas para apoyar a los estudiantes en riesgo de deserción.

42.3 Materiales y Métodos

Tomando el curso de "Gestión y Análisis de Datos" del programa de Ingeniería de Sistemas en el periodo 2024-1, que cuenta con 18 estudiantes, se analizó cuántos de estos estudiantes, al llegar a la mitad del semestre (semana 6), se encuentran por debajo de la nota aprobatoria (3.0). En la Figura 42.1, se muestra la distribución de estudiantes en relación con sus calificaciones: Para esta

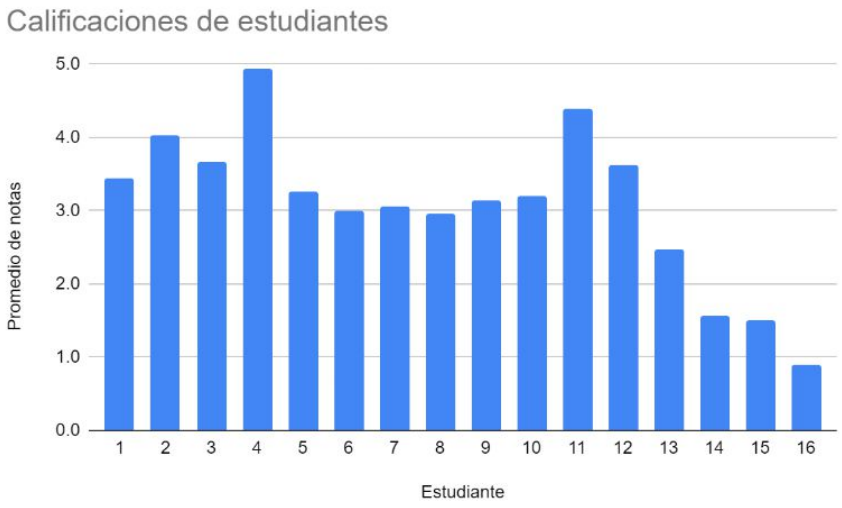


Fig. 42.1

distribución se obtuvo una desviación estándar (s) de 1.1, lo que indica que la mayoría de los datos tienden a estar agrupados cerca de la media del grupo ($\mu = 3.1$)..

42.4 Metodologia

Para desarrollar una herramienta tecnológica de predicción de deserción estudiantil en el programa de Ingeniería de Sistemas, se emplearán las metodologías PRISMA y CRISP-DM. La metodología PRISMA [7] se utilizará para identificar y seleccionar fuentes de datos relevantes, como registros académicos y sistemas de gestión estudiantil. Posteriormente, se evaluará la calidad de estos datos, extrayendo variables clave como calificaciones, asistencia y participación en actividades extracurriculares. Estos datos se sintetizarán y analizarán mediante un algoritmo de aprendizaje automático para generar un modelo predictivo de deserción. Los resultados serán accesibles para administradores y educadores, y la herramienta se implementará en entornos educativos. CRISP-DM (Cross-Industry Standard Process for Data Mining) [8] es una metodología robusta y estructurada para la minería de datos que guía a los equipos a través de seis fases interconectadas: comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue. El proceso comienza con la comprensión del problema y los objetivos del negocio, avanzando luego hacia la exploración y preparación de los datos para el modelado. Durante la fase de modelado, se aplican diversas técnicas para construir y evaluar modelos, culminando en el despliegue y seguimiento de los resultados para asegurar que satisfacen las necesidades iniciales del negocio. Esta metodología fomenta la iteración y la retroalimentación continua, lo que permite mejorar y optimizar los modelos de manera efectiva.

42.5 Resultados

Desde el semillero de investigación PADIA, se desarrolló un modelo de predicción basado en inteligencia artificial utilizando características académicas de los estudiantes para generar alertas tempranas sobre la alta probabilidad de deserción estudiantil. En la Figura 42.2 se muestran los algoritmos utilizados, destacando que AdaBoost resultó ser el mejor modelo, con una métrica MSE de 0.04 y un R^2 Score de 0.97.

42.6 Conclusiones

La implementación de modelos de aprendizaje automático, como AdaBoost en nuestro caso, en los sistemas de predicción de deserción estudiantil permite generar alertas tempranas sobre los estudiantes que están en riesgo de abandonar un curso. Estos enfoques no solo mejoran la precisión en la identificación de factores críticos que influyen en el rendimiento académico, sino que también facilitan intervenciones oportunas por parte de educadores y administradores. Al predecir de manera efectiva qué estudiantes pueden necesitar apoyo adicional, es posible implementar estrategias preventivas y personalizadas que mejoren las tasas de retención y éxito académico. En conjunto, el uso de técnicas avanzadas de machine learning como AdaBoost, combinado con metodologías sólidas de gestión de datos, contribuye significativamente a la mejora de los resultados educativos y a la reducción de la deserción estudiantil.

Agradecimientos Universidad de San Buenaventura Cali.

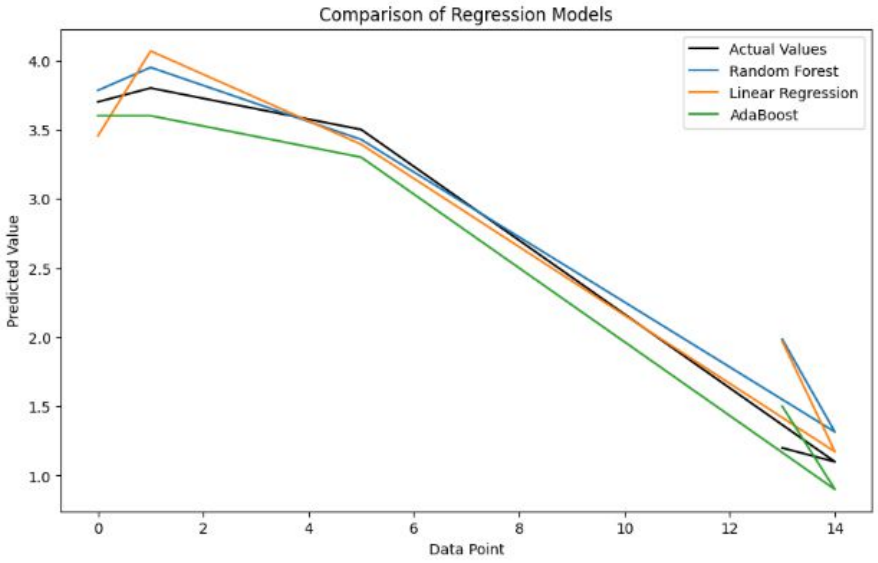


Fig. 42.2

Referencias

1. M. B. Castillo and A. M. Giraldo, “Los retos de la educación superior en Colombia: una reflexión sobre el fenómeno de la deserción universitaria,” Rev. Educ. En Ing., vol. 5, no. 10, Art. no. 10, Dec. 2010, doi: 10.26507/rei.v5n10.101.
2. A. Pérez-Gama, G. Hoyos, L. Parra-Espitia, M. Ortegón, L. G. Rozo-Pardo, and B. Perez-Gutierrez, “Education software architecture: Facing student desertion in Colombia higher education with an intelligent knowledge based coaching system,” in 2010 IEEE ANDESCON Conference Proceedings, ANDESCON 2010, IEEE, 2010, pp. 1–5. doi: 10.1109/ANDESCON.2010.5633139.
3. J. Bennedsen and M. E. Caspersen, “Failure rates in introductory programming - 12 years later,” ACM Inroads, vol. 10, no. 2, pp. 30–35, 2019, doi: 10.1145/3324888.
4. A. Ali and D. Smith, “Teaching an Introductory Programming Language in a General Education Course,” J. Inf. Technol. Educ. Innov. Pract., vol. 13, pp. 57–67, 2014.
5. A. A. Alsulami, A. S. A.-M. AL-Ghamdi, and M. Ragab, “Enhancement of E-Learning Student’s Performance Based on Ensemble Techniques,” Electronics, vol. 12, no. 6, Art. no. 6, Jan. 2023, doi: 10.3390/electronics12061508.
6. R. J. García et al., “DESARROLLO DE UNA HERRAMIENTA COMPUTACIONAL DE ALERTAS TEMPRANAS DE POSIBLES CASOS DE DESERCIÓN ESTUDIANTIL DEL PROGRAMA DE INGENIERÍA DE SISTEMAS DE LA UNIVERSIDAD DEL SINÚ UTILIZANDO ALGORITMOS DE MACHINE LEARNING,” null, 2018, doi: null.
7. “PRISMA.” Accessed: Oct. 17, 2019. [Online]. Available: http://prisma-statement.org/PRISMAStatement/FlowDiagram.aspx
8. R. Wirth and J. Hipp, “CRISP-DM: Towards a standard process model for data mining,” Proc. 4th Int. Conf. Pract. Appl. Knowl. Discov. Data Min., Jan. 2000, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.198.5133

43
Gemas Educativas: aplicación web para detección y atención de dificultades específicas de aprendizaje

Paula Andrea Taborda Montes [1]
Julián Pachón Castrillón[2]
Manuela Marulanda Aguirre[3]
Néstor Darío Duque Méndez [4]
Valentina Tabares Moralese[5]
María José Giraldo Gutiérrez[6]

43.1 Introducción

En un mundo donde la educación se esfuerza por ser cada vez más personalizada e inclusiva, la atención temprana a las necesidades específicas de los estudiantes cobra una mayor importancia. Entender que cada uno aprende de forma distinta les permite a sus docentes abordar los procesos educativos con un enfoque más efectivo, fomentando el análisis y la aplicación de estrategias pedagógicas adecuadas1. Las dificultades en el aprendizaje de las aptitudes académicas esenciales, como lo son la lectura, escritura y matemáticas, se han definido como Dificultades Específicas de Aprendizaje, y pueden llegar a afectar los procesos esenciales en el ámbito académico sin otras causas educativas, ambientales o culturales aparentes2.

43.2 Objetivos

- Desarrollar una herramienta tecnológica que integre los recursos para detección y atención de dificultades específicas de aprendizaje creados en el marco del proyecto de investigación.
- Implementar las interfaces diseñadas según criterios de accesibilidad, usabilidad y experiencia de usuario.
- Brindar a docentes y acudientes información sobre las necesidades educativas y el progreso de los estudiantes con respecto a las dificultades específicas de aprendizaje identificadas.

43.3 Materiales y métodos

- Tecnologías para la implementación del backend: Laravel, PHP y MySQL.
- Tecnologías para la implementación del frontend: NextJS, TypeScript y Tailwind CSS.

[1,2,3,4,5,6]Universidad Nacional de Colombia, sede Manizales
Grupo de Ambientes Inteligentes Adaptativos (GAIA)
e-mail: \{ptabordam,jupachon,mamarulandaa,ndduqueme,vtabaresm,mgiraldogu\}@unal.edu.co

- Metodología de desarrollo: tanto para la implementación del backend como para la implementación del frontend se siguió un desarrollo en espiral.

43.4 Resultados

Backend desplegado, con implementación de los requerimientos principales identificados para la herramienta y una colección de endpoints. Desde el frontend, se han implementado una serie de interfaces que hacen uso de los endpoints disponibles y permiten la interacción de usuarios finales con la plataforma. Se han desarrollado 218 actividades de detección de discalculia, 28 actividades de detección de dislexia, 61 actividades para atención de dislexia, 7 actividades de detección de disgrafía y 7 actividades para atención de disgrafía. Se han integrado 20 actividades de detección de discalculia y 28 actividades de detección de dislexia en Gemas Educativas. El 15 de noviembre

Interfaz de presentación de la prueba de un estudiante

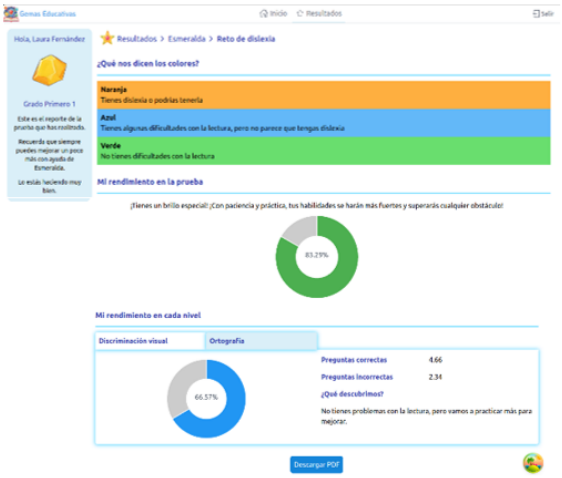


Fig. 43.1: Interfaces de visualización de los resultados obtenidos por un estudiante

de 2023 se realizaron pruebas para detección de dislexia en el Instituto Manizales, con una versión inicial de la plataforma.

43.5 Conclusiones

- La implementación de Gemas Educativas proporciona un medio centralizado para identificar y abordar dificultades en las matemáticas y la lectura.
- Los criterios de accesibilidad, usabilidad y experiencia de usuario para el diseño de las interfaces de la herramienta fueron clave en la implementación final para garantizar una experiencia inclusiva para todos los usuarios.
- La plataforma brinda a docentes y acudientes información que permite ajustar los enfoques pedagógicos y proporcionar un apoyo adaptado a las necesidades de cada estudiante.

Agradecicmientos El trabajo de investigación presentado fue financiado parcialmente por el proyecto titulado: ”Construcción de una herramienta tecnológica para detección y atención de estudiantes con dificultades específicas de aprendizaje” de la Universidad Nacional de Colombia, con Código 55965.

44
Impulsando la innovación aeroespacial desde el suroccidente colombiano: astra aether en el curso-concurso mundial de cansat 2024

Angela L. Mutiz P.[1]
Victor M. Macias M.[2]
Santiago Chaves C.[3]
Sarah A. Cabeza L. [4][ORCID]
Sebastián C. Guerrero E[5][ORCID]
Fabián A. Ramírez M.[6][ORCID]

44.1 Introducción

En la última década, el desarrollo de pequeños satélites tipo CanSat ha emergido como una herramienta educativa clave para la formación en tecnologías espaciales. Estos satélites enlatados, diseñados para replicar las funciones de los satélites convencionales dentro de un volumen de aproximadamente una lata de refresco, ofrecen a los estudiantes una plataforma práctica para aplicar conceptos avanzados de ingeniería aeroespacial. El equipo Astra Aether de la Universidad del Cauca participó en el Curso/Concurso Mundial de CanSat 2024 organizado por el Programa Espacial Universitario de la Universidad Nacional Autónoma de México con el objetivo de diseñar un CanSat optimizado para la recolección y transmisión de datos durante su ascenso y descenso en la misión. A lo largo del curso/concurso, el equipo demostró una notable capacidad técnica, destacando en el top diez mundial en seis de las siete etapas del evento. La competencia incluyó fases rigurosas de diseño, integración y pruebas, evaluando aspectos críticos como la aerodinámica, la funcionalidad del sistema de recuperación y la integridad de la comunicación de datos.

44.2 Objetivos

El equipo Astra Aether de la Universidad del Cauca estableció los siguientes objetivos para su participación en el Curso/Concurso Mundial de CanSat 2024:

1. *Diseñar y Construir un Satélite Enlatado Funcional:* Crear un Satélite Enlatado que pueda transmitir datos precisos sobre presión, temperatura, orientación y aceleración durante todo el trayecto de ascenso y caída. Este objetivo incluye la integración de todos los componentes necesarios para el funcionamiento y la transmisión de datos.
2. *Implementar un Sistema de Autogiro Efectivo:* Desarrollar e implementar un sistema de autogiro que se despliegue a una altitud de 200 metros para reducir la velocidad de caída de la carga primaria. El sistema debe ser capaz de estabilizar el descenso y proteger la carga durante el impacto.

[1,2,3,4,5,6]Universidad del Cauca
e-mail: \{angelamutiz,vmmacias,santchaves,scabeza,secguerrero,framirezm\}@unal.edu.co

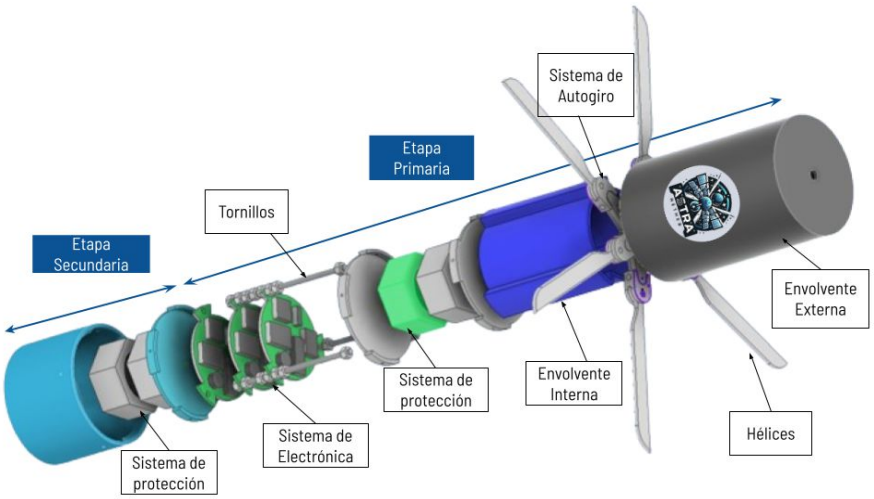


Fig. 44.1: CanSat Astra Aether: representación en CAD del modelo final

3. *Proteger la Carga Útil:* Diseñar un sistema de protección para dos huevos de gallina, asegurando que sobrevivan al impacto del aterrizaje sin romperse. Esto implica la creación de estructuras y mecanismos que absorban y distribuyan las fuerzas durante la caída.
4. *Garantizar la Transmisión Continua de Datos:* Asegurar que el Satélite Enlatado continúe transmitiendo datos durante al menos 10 segundos después del aterrizaje, proporcionando información completa y útil sobre el desempeño del CanSat durante la misión.
5. *Implementar un Mecanismo de Desacoplamiento Eficiente:* Diseñar e implementar un mecanismo de desacoplamiento que permita la separación controlada de la carga secundaria de la carga primaria en el momento adecuado, facilitando la caída libre de la carga secundaria.
6. *Lograr Precisión en el Aterrizaje:* Asegurar que la carga primaria aterrice lo más cerca posible del centro de un objetivo circular de 3 metros de diámetro desde una altura de 450 metros. Este objetivo se enfoca en la precisión del aterrizaje para evaluar la exactitud del sistema de navegación y control.

44.3 Materiales y Métodos

44.3.1 Materiales

El desarrollo del CanSat para la competencia incluyó una cuidadosa selección de materiales y componentes tanto electrónicos como mecánicos, con el objetivo de garantizar la funcionalidad y la integridad del sistema durante toda la misión. A continuación, se describen los materiales y componentes clave utilizados en el proyecto:

1. Componentes Electrónicos:
   - *Tarjeta de desarrollo ESP32S:* Actúa como la computadora de a bordo, proporcionando conectividad WiFi, Bluetooth y Bluetooth de Bajo Consumo (BLE) para la comunicación entre dispositivos y el procesamiento de datos.

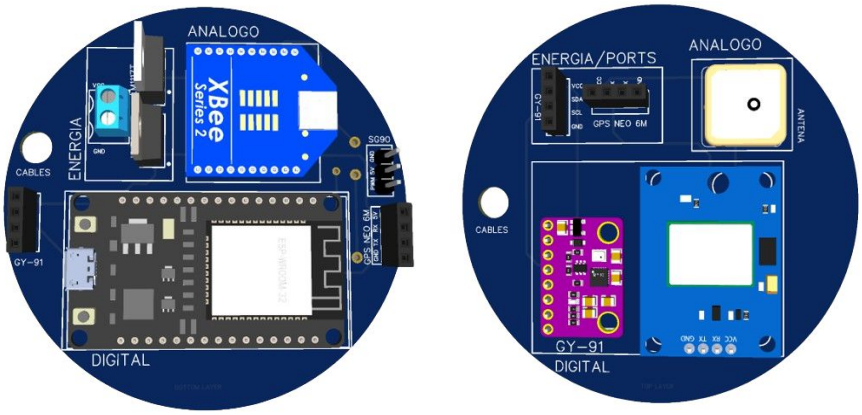


Fig. 44.2: Diseño de PCB´s

- *IMU GY-87:* Sensor combinado de presión, temperatura y giroscopio que recopila datos ambientales y de movimiento, cruciales para el cálculo de altura y velocidad del CanSat.
- *Módulo U-blox NEO-M8N:* GPS que determina la posición del CanSat en tiempo real, esencial para el seguimiento de la misión.
- *Módulo XBee S2 S2C Zigbee:* Facilita la comunicación bidireccional entre el CanSat y la estación terrena mediante radiofrecuencia.
- *Batería Beston 9V:* Fuente de alimentación que suministra energía a todos los dispositivos a bordo del CanSat.
- *Reguladores de Voltaje (L805CV y LM1117T):* Regulan el voltaje de la batería para proporcionar $5V$ y $3.3V$ a los dispositivos según sus requerimientos específicos.
- *Servomotor SG90:* Activa el mecanismo de liberación del sistema de autogiro, permitiendo el desprendimiento de la carga secundaria.

2. Componentes Mecánicos:

- Envolvente (PLA): Estructura impresa en 3D en PLA, proporcionando una estructura ligera pero resistente para proteger la electrónica interna del CanSat.
- Columnas (Nylon): Componentes de Nylon, que contribuyen a la estabilidad y soporte interno del CanSat.
- Circuitos (PCB's): Placas de circuitos impresos, que forman la base de la electrónica del CanSat.
- Rotor (Acero): Componente del sistema de autogiro, esencial para la desaceleración del CanSat durante el descenso.
- Hélices (PLA con Fibra de Carbono): Impresas en 3D con PLA reforzado con fibra de carbono, y forman parte del sistema de autogiro.
- Cápsula del Huevo (Espuma): Proporciona protección adicional a la carga útil, asegurando su integridad durante el impacto.

Fig. 44.3: Diseño del autogiro en modelo CAD.

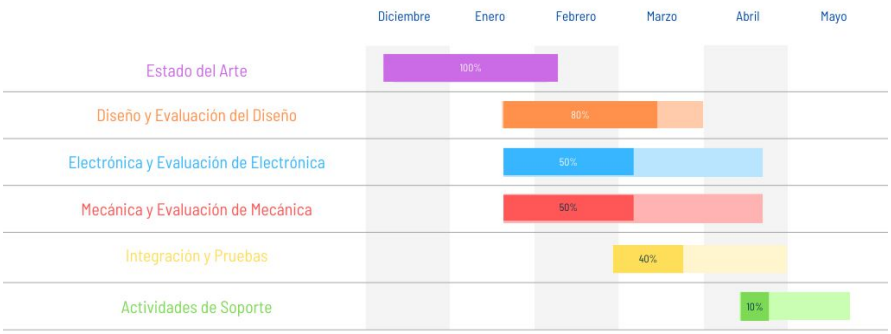


Fig. 44.4: Diagrama de Gantt de atividades

44.4 Métodos

El desarrollo del CanSat se llevó a cabo siguiendo una metodología estructurada en varias fases para asegurar el éxito de la misión. A continuación, se describen los métodos utilizados en cada etapa:

Fase 1: Diseño y Planificación

1. Análisis de Requisitos: Se realizó un análisis exhaustivo de las especificaciones del concurso, incluyendo restricciones de peso, tamaño y funcionalidad.
2. Selección de Materiales y Componentes: Se eligieron materiales como PLA, Nylon y Acero para la estructura, y componentes electrónicos como la tarjeta ESP32S y los sensores, garantizando la compatibilidad y el cumplimiento de los requisitos de la misión.
3. Diseño de la Estructura: Utilizando software CAD, se diseñó una estructura que combina ligereza y resistencia, capaz de proteger la electrónica y soportar las condiciones del vuelo.

Fase 2: Integración y Pruebas

Montaje del Prototipo: Se ensamblaron los componentes mecánicos y electrónicos en la estructura diseñada. Se realizaron pruebas preliminares para verificar la integridad del sistema.

1 Pruebas Funcionales y Mecánicas: Se llevaron a cabo pruebas para asegurar el funcionamiento correcto de todos los subsistemas, incluyendo el autogiro, el sistema de desacople y la transmisión de datos, bajo condiciones simuladas de vuelo.

2 Validación del Sistema de Recuperación: Se realizaron pruebas específicas para verificar la eficacia del sistema de autogiro y las hélices en la reducción de la velocidad de descenso y la protección de la carga útil.

Fase 3: Implementación y Evaluación

1 Lanzamiento y Monitorización: Durante el lanzamiento, se monitorizó la transmisión de datos y el comportamiento general del CanSat. A pesar de un fallo en el sistema de desacople, la electrónica y la estructura demostraron su eficacia.

2 Análisis Post-Lanzamiento: Se analizaron los datos recolectados y se evaluaron los resultados del vuelo, identificando áreas de mejora para futuros proyectos.

44.5 Resultados

A lo largo de la competencia CanSat 2024, el equipo Astra Aether demostró un rendimiento destacado en la mayoría de las etapas, logrando resultados que reflejan tanto el rigor del diseño como la efectividad de las pruebas previas al lanzamiento. A continuación, se presentan los principales resultados obtenidos durante el desarrollo del proyecto:

44.5.1 Desempeño en las Etapas del Concurso

Progreso Consistente: El equipo Astra Aether logró avanzar exitosamente en cada una de las etapas del concurso, obteniendo resultados que permitieron cumplir con los objetivos de la misión establecidos por la convocatoria. Esto incluyó la superación de pruebas de diseño, integración y pruebas funcionales, lo que llevó al equipo a estar dentro del top 10 mundial en 6 de las 7 etapas del concurso.

44.5.2 Comportamiento del CanSat Durante el Lanzamiento

Electrónica Resiliente: A pesar del fallo en el sistema de desacople durante el lanzamiento, la electrónica del CanSat permaneció intacta, demostrando que tanto el diseño como los componentes utilizados eran adecuados para misiones de este tipo. Este resultado es especialmente importante, ya que valida la robustez del sistema frente a condiciones adversas.
Protección Estructural Efectiva: La envolvente del CanSat, fabricada en PLA mediante impresión 3D, cumplió con su función de proteger los componentes internos. Además, la estructura interna, diseñada para ser ligera y fácilmente manipulable, también se comportó de manera óptima, salvaguardando la integridad de la electrónica.
Sistema de Autogiro: Aunque no se pudo probar el sistema de recuperación directamente el día del lanzamiento, las pruebas previas realizadas desde alturas similares a las de la competencia confirmaron que el diseño de los dos rotores con tres hélices cada uno era más que suficiente para proteger el sistema durante el descenso. *Transmisión de Datos:* La transmisión de datos desde el CanSat

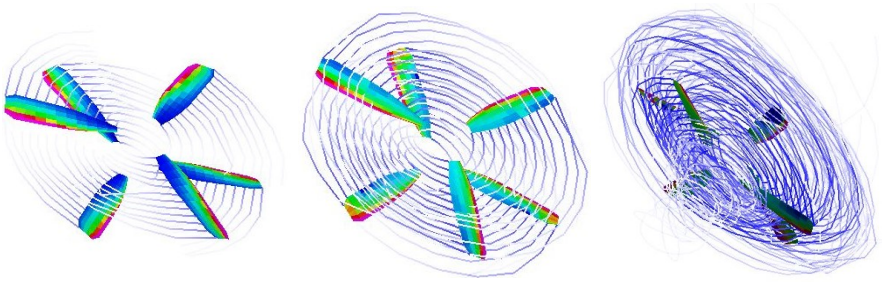


Fig. 44.5: Simulación de rotación de las hélices en las condiciones de CDMX.

hacia la estación terrena se realizó de manera adecuada. Los datos recopilados y transmitidos antes del fallo en el desacople respaldan la eficiencia del sistema de comunicación y el procesamiento de la información a bordo.

44.5.3 Impacto y Aprendizajes

Reconocimiento Internacional: El equipo Astra Aether, con su destacada actuación, contribuyó significativamente al desarrollo tecnológico espacial desde el suroccidente colombiano. A través de este proyecto, se demostró que es posible desarrollar prototipos de satélites desde esta región, impulsando la exploración espacial y posicionando al equipo como uno de los mejores.
Lecciones Aprendidas: La experiencia adquirida en esta competencia es invaluable. Los resultados obtenidos, tanto en las etapas previas al lanzamiento como durante el mismo, proporcionan una base sólida para mejorar en futuras misiones. La identificación de áreas de mejora, como el sistema de desacople, es crucial para avanzar en la evolución del diseño y garantizar el éxito en futuras misiones.

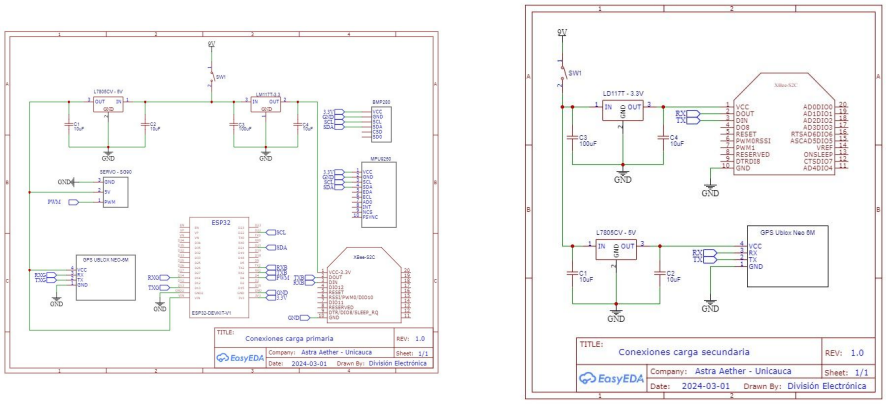


Fig. 44.6: Esquema de conexiones carga primaria (izquierda) y carga secundaria (derecha).

44.6 Conclusiones

El equipo Astra Aether en la competencia CanSat 2024 representó un logro significativo en el ámbito de la tecnología espacial, destacando la capacidad de los estudiantes para abordar desafíos complejos y alcanzar resultados sobresalientes. A pesar de los obstáculos enfrentados, como el fallo en el sistema de desacople durante el lanzamiento, el equipo demostró la solidez de su diseño electrónico y estructural, validando la eficacia de los materiales y técnicas de manufactura seleccionados. Los aprendizajes obtenidos, especialmente en cuanto a la necesidad de pruebas más exhaustivas y la consideración de redundancias en sistemas críticos, serán fundamentales para futuras misiones.

Este proyecto no solo consolidó al equipo Astra Aether como un referente en la Universidad del Cauca, sino que también subrayó el potencial del suroccidente colombiano para contribuir al desarrollo de la exploración espacial. La experiencia adquirida proporcionará una base sólida para futuras iniciativas, asegurando que el equipo esté mejor preparado para enfrentar y superar los retos en la continua búsqueda de la innovación tecnológica en el campo aeroespacial.

Agradecimientos Queremos agradecer a la Universidad del Cauca por su apoyo institucional, que nos permitió llevar a cabo parte de este proyecto en el Concurso Mundial CanSat 2024. A la UNAM, le agradecemos por la excelente organización de este evento, que nos brindó la plataforma para poner a prueba nuestras habilidades en un entorno internacional. A nuestra mentora Eliana Aguilar Larrarte, le debemos un especial agradecimiento. Su guía y aliento fueron vitales en cada etapa del proceso, inspirándonos a dar lo mejor de nosotros mismos. Por último, queremos reconocer y agradecer a todos los miembros del equipo Astra Aether. Su dedicación, pasión y esfuerzo incansable hicieron posible que este proyecto se convirtiera en una realidad. Fue un privilegio trabajar juntos y enfrentar, como equipo, cada desafío que se nos presentó.

Referencias

1. ARLISS. (1999). A rocket launch for international student satellites.. Available from http://www.arliss.org//p/ about Accessed December 2016
2. Ayala, S.,Colin, A., González, A.R.,Rivas,R., Mona,J., Avilés, A., et al.(2016). Site testing for the Observatorio Astronómico UANL. 20.DA10: Research and teaching in Astra Aether 9 astrophysics in Guanajuato. In Proceedings of a Conference Held at Guanajuato, GTO.
3. 2do. Concurso CANSAT. (2016). Segundo Concurso Nacional de Pico Satélites Educativos en la Universidad Autónoma de Nuevo León.. Available from http://www.fcfm.uanl.mx/es/noticia/2302 Accessed December 2016
4. CLTP-6. (2015). 6th CanSat Leader Training Program.. Available from: http://cltp.info/cltp6.html. Accessed December 2015
5. UNAM. (2024). Guía de misión: Concurso Mundial CanSat 2024. Recuperado de http://peu.unam.mx/files/CANSAT/2024/GUIA_DE_MISION_MUNDIAL_CANSAT_2024_Esp.pdf

45
Automatización del Aprendizaje: ¿Una Amenaza para la Educación Virtual en Colombia?

Rodriguez Isabela [1]
Angulo Oscar[2]
Lopez Caterine[3]

45.1 Introducción

El E-learning ha revolucionado nuestro acceso al conocimiento desde cualquier lugar. Sin embargo, surge una preocupación significativa cuando el sector de la Inteligencia Artificial (IA) transforma los métodos educativos y los procesos de enseñanzaaprendizaje. Esta tendencia plantea desafíos complejos, como la automatización con baja dependencia, que lleva a los estudiantes a buscar gratificación instantánea y comprensión superficial en lugar de un aprendizaje significativo. Un ejemplo paradigmático lo encontramos en primer lugar en El Servicio Nacional del Aprendizaje SENA), modelo virtual que actualmente presenta un riesgo inminente, para empresarios y el estado Colombiano, Donde los aprendices o estudiantes, altamente motivados por finalizar su proceso de formación, emplean diversos modelos de IA, para acelerar los procesos de entrega de actividades en la plataforma Zajuna, Cabe mencionar que al entrenar adecuadamente los modelos de IA y brindar acceso a las guías de aprendizaje, rubricas y demás componentes de formación, la IA, logra entregar la totalidad de actividades de un programa técnico o tecnólogo que puede tener una duración entre 12 y 18 meses en un periodo no superior a 30 días, no solo en la formación SENA sino a su vez en cualquier institución que ofrezca formación Virtual.

45.2 Objetivos

Demostrar cómo el entrenamiento adecuado de un modelo de inteligencia artificial permite la entrega completa de las actividades de un programa de formación virtual, sin lograr un aprendizaje significativo en el aprendiz o estudiante. Sin embargo, esta automatización plantea un riesgo tanto en el sector empresarial como en el estatal, teniendo como alcance cualquier programa de formación virtual.

Corporation Universitario Comfacauca – SENA
Regional Cauca Sennova
e-mail: oangulo@unicomfacauca.edu.co

45.3 Materiales y métodos

Se emplea una metodología Mixta: La metodología cualitativa se emplea para comprender la profundidad de los desafíos éticos, las ventajas y las preocupaciones relacionadas con la IA en la formación virtual La cuantitativa para identificar la eficacia de las actividades generadas por la IA, evaluando la tasa de aprobación, el tiempo de finalización y la calidad percibida.
Combinadas las dos, se realiza la triangulación de los datos para validar los hallazgos obteniendo una comprensión completa del impacto de la IA en la formación virtual.

45.4 Resultados

Figura 1. Modelo entrenado para realizar actividades

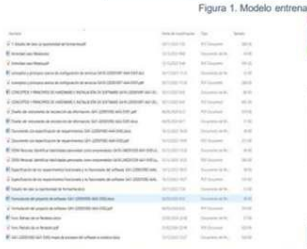


Figura 2. Archivos generados por IA para presentar

Figura 3. Registro de actividades plataforma Zajuna

Figura 3. Reporte de resultados de aprendizaje Aprobado

Figura 4. Barras de Actividades Presentadas y Aprobadas

Figura 5. Tiempo de la IA para generar Actividades

45.5 Conclusiones

Este estudio destaca la necesidad de un análisis minucioso del impacto de los métodos transformadores de enseñanza aprendizaje en el contexto de la inteligencia artificial (IA). El objetivo es garantizar que los enfoques innovadores de eLearning se basen en principios democráticos, sean informados, responsables y éticos. Para lograrlo, se debe explorar tanto los riesgos asociados con la automatización como las oportunidades para fomentar la alfabetización digital y una pedagogía

inclusiva. Estos desafíos afectan directamente nuestra sociedad basada en el conocimiento, y es crucial abordarlos de manera equilibrada y reflexiva, en especial en el ambiente empresarial y la inversión del estado colombiano.

Referencias

1. S. Petit-Renaud, T. Denoeux: Nonparametric regression analysis of uncertain and imprecise data using belief functions. Int. J. Approx. Reason. 35, 2004, 1–28.
2. G. Shafer: Dempster's rule of combination. Int. J. Approx. Reason. 79, 2016, 26–40.
3. B. Panay, N. Baloian, J.A. Pino, S. Peñafiel, H. Sanson, N. Bersano: Feature selection for health care costs prediction using WeightedEvidential Regression. Sensors 20(16), 2020, 4392.
4. B.Panay,N.Baloian,J.A.Pino,S.Peñafiel,J.Frez,C.Fue.

46
PhinGPT: Un gran modelo de lenguaje hecho en casa para el sector financiero

John Robert Gómez Pachón [1]
Tuli Peña Melo[2]
Francisco José Gutiérrez Peña[3]

46.1 Introducción

Los grandes modelos de lenguaje están revolucionando diversos campos del conocimiento. El objetivo de este trabajo fue entrenar un modelo especializado para el sector financiero, buscando optimizar al máximo los costos computacionales pero sin perjudicar la efectividad del entrenamiento. Para esto usamos QLora, un método de fine-tuning de modelos que permite reducir costos computacionales y generar muy buenos resultados. Además de esto evaluamos el rendimiento del modelo entrenado con este método y lo comparamos con el modelo original, enfocándonos en la tarea de análisis de sentimientos y en el reconocimiento de entidades nombradas (NER).

46.2 Objetivos

- Reentrenar el modelo Phi-3-mini-4k-instruct para la tarea de análisis de sentimientos en textos financieros y para la tarea de reconocimiento de entidades nombradas (NER) en textos financieros.
- Evaluar y comparar el desempeño de nuestro modelo reentrenado en las dos tareas seleccionadas con el de Phi-3-mini-4k-instruct sin ningún reentrenamiento.

46.3 Materiales

Como modelo base para esta investigación, utilizamos Phi-3-mini-4k-instruct, un modelo de lenguaje de 3.8 mil millones de parámetros desarrollado por Microsoft. Para satisfacer las demandas computacionales, empleamos Google Colab, un servicio en la nube que proporciona acceso a GPUs y un entorno para ejecutar código. Para el proceso de fine-tuning, utilizamos dos conjuntos de datos, RobertGomezDP/fingpt-cls-sharegptstyle para el análisis de sentimientos en textos financieros, y

Universidad Nacional de Colombia
Semillero de Estudios formales de los modelos generativos
e-mail: \{jrgomezp,tpenam,frgutierrezp\}@unal.edu.co

RobertGomezDP/fingpt-cls-sharegpt-style para el reconocimiento de entidades nombradas (NER) en textos financieros. Ambos datasets se pueden encontrar en la plataforma HuggingFace.

46.4 Métodos

46.4.1 Cuantización

La cuantización de modelos de lenguaje es una técnica de compresión que busca reducir la precisión numérica de los parámetros del modelo sin afectar significativamente su rendimiento. Esta reducción del tamaño de los parámetros brinda ventajas en el proceso de entrenamiento y también en el de inferencia, pues al tener un modelo más liviano es posible entrenarlo y ejecutarlo con un menor costo computacional, lo que permite usar hardware menos especializado o reducir el tiempo de ejecución [2] Formalmente, la cuantización es el proceso de discretizar un conjunto de datos,

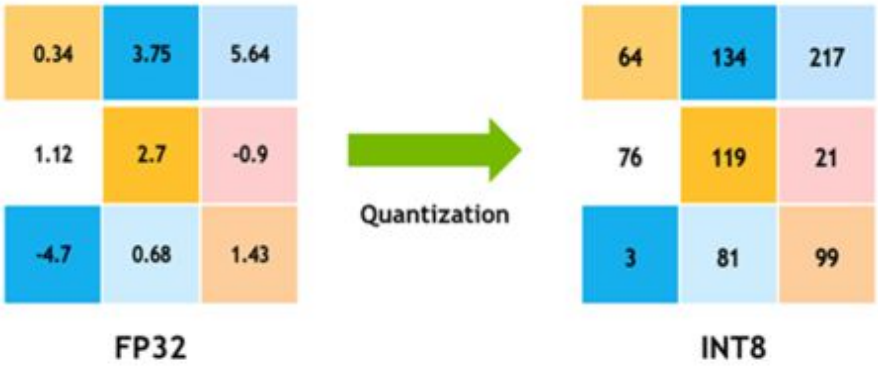


Fig. 46.1: (NVIDIA [5]

transformando una representación que contiene más información en una con menos información. Esto a menudo implica convertir un tipo de dato con más bits en otro con menos bits, por ejemplo, convertir números de coma flotante de 32 bits a enteros de 8 bits [1]. La cuantización realizada en esta investigación consistió en convertir los parámetros de Phi-3, que son números de coma flotante de 16 bits [4], a enteros de 4 bits.

46.4.2 Modelos fundacionales

Los modelos fundacionales son modelos de inteligencia artificial que se caracterizan por su capacidad para aprender representaciones generales del conocimiento a partir de grandes cantidades de

datos no etiquetados y pueden usarse como base para entrenar modelos más específicos. La utilidad de los modelos fundacionales radica en su capacidad de abordar problemas diversos mediante técnicas como el aprendizaje por transferencia y el fine-tuning, lo que permite usar un modelo fundacional en múltiples contextos sin necesidad de entrenar modelos específicos desde cero para cada tarea. Para esta investigación usamos el modelo fundacional Phi-3-mini-4k-instruct.

46.4.3 LoRA

Modelos como los LLMs (Large Language Models) o los modelos de difusión pueden requerir una enorme cantidad de recursos y tiempo para ser entrenados. Por otro lado, pueden ser excesivamente genéricos. Si se desea utilizar uno de ellos en una tarea específica, como un chat-bot o un clasificador, pueden resultar decepcionantes. Sin embargo, se puede tomar un modelo fundacional y reentrenarlo en un conjunto de datos específico, a este proceso lo llamamos fine-tuning. Uno de los métodos más famosos para realizar fine-tuning es LoRA[3]. LoRA congela los pesos preentrenados y, durante el entrenamiento, en lugar de actualizar toda la matriz de parámetros ΔW , solo se actualizan un par de matrices, A y B. Sea ΔW una matriz de dimensiones $d \times d$ y r un parámetro, entonces A tendrá las dimensiones $d \times r$ y B será $r \times d$. Esto nos ayudará a reducir la memoria de la GPU. Si ΔW es una matriz de 1000×1000, ΔW tendrá 1,000,000 de parámetros. Sin embargo, si configuramos r en 2, nuestras matrices A y B tendrán 2000 parámetros cada una, lo que significa que el número de parámetros se reduce a 4000. Al reducir el tamaño de los parámetros con LoRA se reduce a su vez el costo computacional que requiere entrenar y ejecutar el modelo.

46.4.4 QLoRA

Podemos optimizar el proceso de fine-tuning aún más utilizando QLoRA (Dettmers, 2023). QLoRA está construido sobre LoRA pero, además de reducir la matriz de pesos a reentrenar, hace una cuantización de los pesos del modelo a 4-bits, lo que permite ahorrar memoria sin sacrificar rendimiento, y hace también una doble cuantización que nos ayuda a ahorrar aún más memoria. En este método los optimizadores paginados permiten cambiar entre la RAM, VRAM y el disco duro cuando no se tiene memoria disponible. QLoRA es el método que usamos en nuestro proceso de fine-tuning, pues resulta más efectivo por su arquitectura y por la cuantización de los parámetros.

46.5 Resultados

Para la evaluación se considera la veracidad y la estructura de la respuesta. Es correcta si cumple ambos criterios. En base a esto, obtenemos los siguientes resultados.

	Phi-3-mini-4k-instruct	**PhinGPT**
Análisis de Sentimientos	0.45	0.98
Reconocimiento de entidades nombradas	0.13	0.99

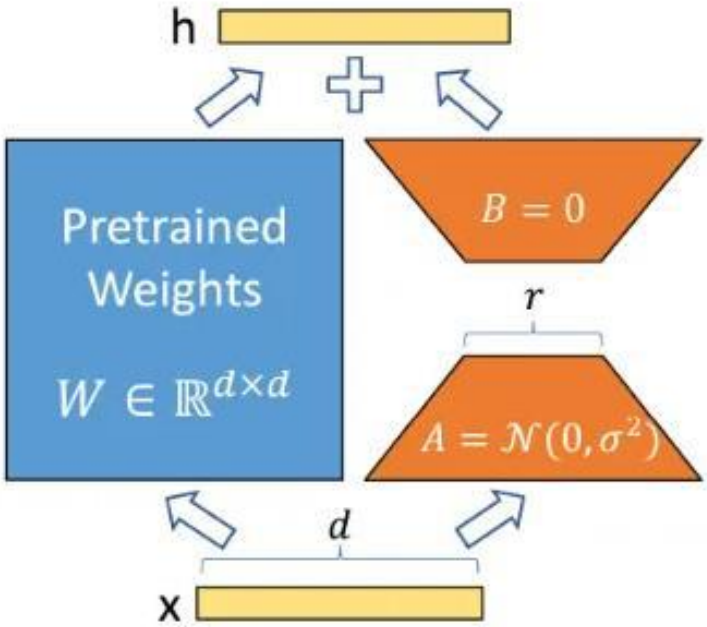


Fig. 46.2: (NVIDIA, 2021 [5])

46.6 Conclusiones

En conclusión, el uso de métodos de optimización como QLoRA ha mostrado un buen resultado a nivel de entrenamiento y una alta eficiencia computacional, lo que lo convierte en una opción computacionalmente viable para diversas aplicaciones, sobre todo para modelos con una gran cantidad de parámetros. Por otro lado, los resultados obtenidos han sido sumamente prometedores, tanto en la tarea de análisis de sentimientos como en la de reconocimiento de entidades nombradas, pues el reentrenamiento mostró ser efectivo para mejorar las capacidades del modelo y también para establecer la estructura deseada de respuesta.

46.7 Trabajo futuro

En el futuro planteamos no solo limitarnos a técnicas de reentrenamiento como LoRA y QLoRA, sino también considerar métodos como el Reinforcement Learning from Human Feedback (RLHF) o el Direct Preference Optimization (DPO) para buscar resolver tareas diversas. También apuntamos a la creación de un Mixture-of-Experts, una arquitectura que integra varios agentes especializados en tareas distintas para mejorar su versatilidad.

Referencias

1. Dettmers, T: QLoRA: Efficient Finetuning of Quantized LLMs (2023).
2. Gholami, A.: A Survey of Quantization Methods for Efficient Neural Network Inference (2021).
3. Hu, E.: Lora: Low-Rank Adaptation Of Large Language Models (2021).
4. Microsoft: Phi-3 Technical Report: A Highly Capable Language Model Locally on Your Phone (2024).
5. NVIDIA, https://developer.nvidia.com/blog/achieving-fp32-accuracy-for-int8-inference-using-quantization-aware-training-with-tensorrt/, last accessed 2024/06/01.

47

Interactive Voice Response in an Emergency Medical Dispatch Center: Is it efficient and safe for the elderly population?

Anna Ozguler[1]
Marin Boyet[2]
Margot Cassuto[3]
Jeremie Boutet[4]
Guillaume Douge[5]
Michel Baer[6]
Thomas Loeb[7]

47.1 Introduction

In the first wave of COVID-19 pandemic, Emergency Medical Dispatch Center faced an influx of calls, many concerning light COVID symptoms, overwhelming its capacity of response in due time. An Interactive Voice Response (IVR) was activated to handle emergency calls more quickly with limited human resources.

47.2 Objectives

This study's objective was to identify if the implementation of IVR was efficient, safe and used by elderly people.

47.3 Methodology

From March 19 until April 26, 2020, an IVR was activated every day from 8 AM to 12 PM. IVR offered the caller 2 options:

- either press the 'zero' key for COVID or Flu symptoms with no respiratory failure or
- stay on line for any other complaint.

If the caller pressed "zero", the call was transferred to an emergency medical dispatcher (EMD) specially trained to handle light COVID cases (COVID EMD). If he stayed on line, his call was placed in the queueing list of all emergency calls and was hung up by a "conventional EMD". All medical dispatch files with a single incoming call picked up during IVR activation period

[1,2,3,4,5,6,7]Ecole Polytechnique, Palaiseau- France
SAMU 92 – APHP
e-mail: anna.ozguler@aphp.fr

were included and classified in 2 groups: “IVR Yes” if caller pressed ‘zero’ and “IVR No” if not. Patient’s age and gender, profile of the caller (patient or other) were collected. EMD assessed the patients’ severity by the way of 3 decisions: sending an Advanced Life Support (ALS) ambulance, a Basic Life Support (BLS) ambulance or no transport. Data were compared between the 2 groups with Chi-square tests. Results Two thousand eight hundred and forty six callers were in the group “IVR Yes” and 12111 in “IVR No”. Elderly patients used IVR less often: 12.7for younger (p¡10-3). Results are in table 1. IVR allowed transferring almost 15longer on line to be hung up. This process decreased the number of calls handled by the conventional EMD and allowed more time to respond to severe emergency calls.

47.4 Discussion

Only 0.07% “IVR Yes” calls needed an ALS ambulance, so we may assume that the use of IVR in this COVID background is safe. The population over 65 is under-represented in “IVR Yes”, and it is likely that the elderly may not be able to use IVR with ease at least during those COVID times. Moreover, the number of ambulances (mainly BLS) in the “IVR yes elderly group” is high, reflecting severe cases and the difficulty for this population to use safely this technology.

47.5 Conclusion

IVR in COVID times was an effective tool, allowing safe triage of less serious patients and freeing up time to respond to severe calls. The elderly may be reluctant to use this technology or sometimes misuse it.

Table 47.1: Comparisons of patients’ age, gender, caller profile and decision taken at dispatch center among IVR-Yes (n=2846) and IVR-No (n=12111) callers *ALS: Advanced Life Support ambulance : **BLS: Basic Life Support ambulance

	65 **years old**			< 65 **years old**		
Gender			$< 10^{-3}$			$< 10^{-3}$
- Male	121 (32.5%)	1075 (42.1%)	¡	850 (34.3%)	3895 (40.8%)	
- Female	189 (50.8%)	1370 (53.6%)		1095 (44.3%)	5011 (52.4%)	
- Not collected	62 (16.7%)	110 (4.3%)		529 (21.4%)	650 (6.8%)	
Caller (patient her/himself)	252 (67.9%)	1067 (41.8%)	$< 10^{-3}$	2088 (84.5%)	5870 (61.4%)	$< 10^{-3}$
Assessment of severity: Decision			$< 10^{-3}$			$< 10^{-3}$
- ALS* ambulance	1 (0.3%) 60	(2.3%)		1 (0.1%)	81 (0.8%)	
- BLS ambulance**	67 (18.1%)	1045 (40.0%)		114 (4.6%)	1614 (16.9%)	
- Non transport	302 (81.6%)	1450 (56.7%)		2353 (95.3%)	7859 (82.3%)	

48

Sensores de bajo costo para monitoreo ambiental comunitario en sistemas de alertas comunitarios ante crecientes súbitas y avalanchas

Maria J. Henao Salgado [1]
Aldemir Vargas Eudor[2]
Jorge J. Vélez Upegui[3]
Néstor Dario Duque Méndez[4]

48.1 Introducción

Los sistemas de alerta temprana comunitarios (SATC) ante inundaciones en regiones montañosas son estrategias participativas y territoriales que reconocen al ciudadano como protagonista de la gestión del riesgo ante inundaciones [1]. El monitoreo participativo de las variables hidro-climatológicas como las variaciones de niveles de las quebradas puede ayudar a mejorar la eficacia y toma de decisiones del SATC ante inundaciones en la quebrada Manizales (Manizales). En 2011, dos crecientes súbitas de la quebrada Manizales causaron impactos negativos significativos en viviendas, empresas y afectaron a las comunidades de Maltería, Juanchito y Verdum, dejando a estas zonas expuestas a un alto riesgo de futuros eventos de crecientes súbitas.

48.2 Objetivos

- Desarrollar una solución tecnológica asequible para el monitoreo del nivel de quebradas en áreas montañosas.
- Diseño y construcción de sensores de bajo costo que funcionen de manera remota.
- Capacitar al colegio IES Maltería y la comunidad de Maltería sobre programación y uso de sensores de bajo costo.
- Implementar niveles de alerta con sonidos de acuerdo con umbrales comunitarios para avisar a los observadores ciudadanos, ubicados en las comunidades afectadas Maltería, Juanchito y Verdum.

[1,2,3,4]Universidad Nacional de Colombia, sede Manizales. Colombia
[2,4]Grupo de Ambientes Inteligentes Adaptativos. (GAIA)
[1,3]Grupo de Trabajo Académico en Ingeniería Hidráulica y Ambiental. (GTA)
[1]Doctorado en Ingeniería Civil.
[2]Profesor Administración de Sistemas Informáticos.
[3]Profesor titular Departamento de Ingeniería Civil.
[4]Profesor titular Departamento de Informática y Computación.
e-mail: \{mjhenaos,avargase,jjvelezu,ndduqueme\}@unal.edu.co

48.3 Materiales y Métodos

Componentes de la estación (ver Tabla 48.1):

Table 48.1: Componentes de la estación con el sensor de bajo costo

Componente	Referencia Tienda	Precio
Placa de desarrollo FireBeetle ESP32 IoT	[1]SKU:DFR0478	US$8,90
Sensor de distancia ultrasónico impermeable con sonda separada (20 600 cm)	[1]SKU:SEN0208	US$16,00
Administrador de energía solar con Panel (5V 1A)	[1]SKU:DFR0559-1	US$25,40
Módulo de radio LoRa 433Mhz	[1]SKU:TEL0114	US$39,90
Módulo reloj de tiempo real	[2]RTC-I2C-TINY	US$1,50
Módulo adaptador para USD microSD	[2]TARUSD-CN-01	US$1,13
[1]Tienda de electrónica DFRobot	[2]Tienda Didácticas Electrónicas	

- Investigación y análisis: Revisión de tecnologías existentes (LoRa, Bluetooth, Zigbee, Wi-Fi, GPRS, etc) [2][3][4][5] para el monitoreo de nivel analizando las características técnicas y los costos.
- Prototipado: Diseño de una solución con componentes de bajo costo. Se emplearon PCB (Printed Circuit Board) de fabricación propia para ensamblar los componentes. Una vez ensamblado el prototipo se realizaron pruebas de medición de distancia y pruebas de comunicación, lo que permitió identificar y corregir posibles fallos.

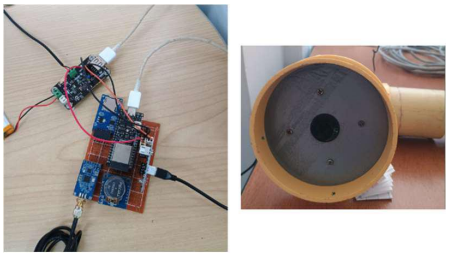


Fig. 48.1: De izquierda a derecha: ensamble en la PCB y Ubicacion del sensor de ultrasonido en tuberia PVC.

- Puesta en marcha (prueba piloto): Actualmente se está llevando a cabo una prueba en campo, en la que se ha instalado la estación de nivel en la zona de Maltería (Figura 48.3). Esta prueba tiene como objetivo evaluar el rendimiento y la efectividad del equipo en condiciones reales de operación.
- Capacitación y colaboración con la comunidad: Realización de talleres con la comunidad local sobre como instalar y mantener los sensores y como interpretar los datos recopilados.

48.4 Resultados

En la Figura 48.2, se tiene la representación del sistema completo para la captura de nivel (estación propuesta) y el envío de los datos para almacenamiento en un motor de base de datos SQL a través de conexión a internet del módulo receptor. Se ha llevado a cabo la instalación de la estación de

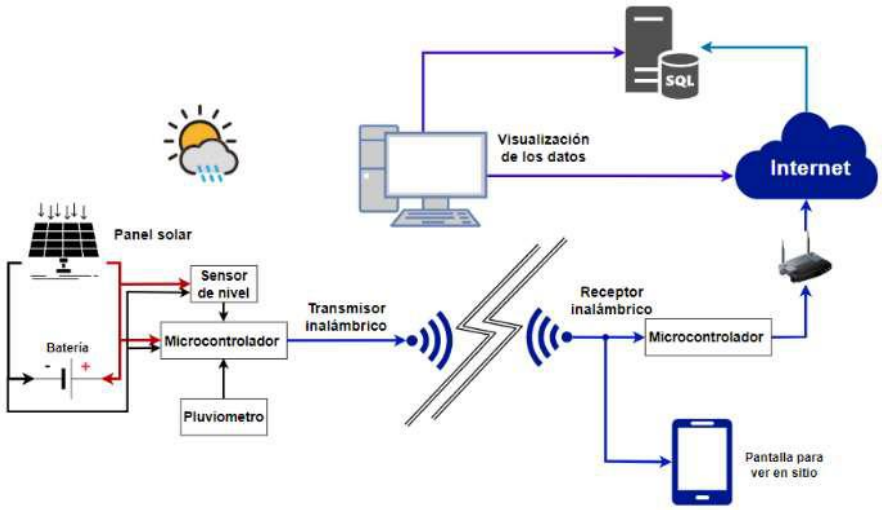


Fig. 48.2: Representación del sistema propuesto.

nivel en la quebrada Manizales, en un punto del barrio Maltería con el fin recopilar datos sobre el funcionamiento en realizar ajustes y optimizaciones al sistema propuesto.

48.5 Conclusiones

Las quebradas en zonas montañosas son propensas a crecidas repentinas por lluvias intensas, y la falta de monitoreo adecuado, debido a los altos costos, aumenta el riesgo de desastres que afectan a las comunidades cercanas. La reducción de costos en el monitoreo permite ampliar la cobertura de estaciones y son una solución viable para abordar los desafíos relacionados con el monitoreo de quebradas en áreas montañosas y la creación de sistemas de alerta comunitaria que ayuden a prevenir pérdidas materiales y humanas. La implementación de estaciones de monitoreo de bajo costo y la colaboración de la comunidad local, ofrece una solución efectiva para establecer sistemas de alerta temprana. Estas estaciones pueden suministrar datos en tiempo real sobre niveles de agua, precipitación y otros indicadores clave y al involucrar la comunidad en el uso de estas estaciones, se fomenta la conciencia y la acción preventiva ante posibles desastres por inundaciones. Finalmente, involucrar la participación ciudadana en el manejo y soporte de las estaciones de monitoreo crea conciencia, resiliencia, adaptación al cambio climático y promueve la acción preventiva.

Agradecimientos Esta investigación hace parte del proyecto “Proyecto de monitoreo comunitario y ciencia ciudadana de la quebrada Manizales para fortalecer el sistema de alertas tempranas ante

Fig. 48.3: Representación del sistema propuesto.

inundaciones y creciente". Se agradece especialmente al grupo GAIA y GTA por su compromiso y esfuerzo en este proyecto y en el desarrollo y puesta en marcha de los sensores.

Referencias

1. SIATA. (2014). Sistema de Alerta Temprana de Medellín y el Valle de Aburrá. http://www.siata.gov.co/newpage/index.php, last accessed 2024/06/15
2. R. A. Sadek and H. M. Elbadawy, "Towards IoT Era with current and Future Wireless Communication Technologies: An Overview," 2022 39th National Radio Science Conference (NRSC), Cairo, Egypt, 2022, pp. 343-353
3. O. Albayari, A. Yousef, R. Albawab, B. Jibrini and M. Salah, "Low-Cost WirelessmMonitoring and Control: A Case Study for Industrial Implementation," 2023 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), Amman, Jordan, 2023, pp. 171-176
4. 4. R. Srinivasan, B. Mallikarjuna, M. Kavitha, R. Kavitha and B. B. Naib, "A Comparative Study: Wireless Technologies in Internet of Things," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, pp. 675-679
5. I. D. Sumitra, G. Medhav Kumar and T. Khaerunisa, "Smart Farming Based Low-Cost and Energy Efficient on Wireless Sensor Networks," 2023 International Conference on Informatics Engineering, Science & Technology (INCITEST), Bandung, Indonesia, 2023, pp. 1-6

49

Modelo de Interoperabilidad (IO) de Historia Clínica Electrónica (HCE) para la Red Pública de Prestadores de Servicios de Salud en el nivel territorial departamental

Julio Cesar Chavarro Porras [1[0000-0001-8876-8855]]

49.1 Introducción

La ley 2015 de 2020 establece un marco normativo para garantizar el intercambio de datos clínicos entre las instituciones que prestan servicios de salud en el territorio colombiano; la resolución 866 de 2021, determina las variables y datos que deben ser intercambiados; y finalmente, los lineamientos diagnóstico, formulación, planificación y directrices técnicas y de gobernanza para entrar en producción a nivel nacional han sido establecidos desde el pasado 23 de noviembre de 2023. El MSPS y MinTIC han adoptado los estándares XROAD y HL7-FHIR, como las restricciones que deben ser tenidas en cuenta en la construcción de la solución de IO. En este marco normativo se desarrolla una propuesta de interoperabilidad de la HCE, iniciando por prestadores que constituyen la red pública de salud en un territorito, constituida por los hospitales públicos con sus diferentes niveles de complejidad.

49.2 Objetivos

Proponer un modelo de Interoperabilidad de la HCE para la red pública de prestadores de un territorito departamental sujeto a los estándares tecnológicos y los elementos normativos establecidos.

49.3 Materiales y Métodos

Se desarrollaron las siguientes etapas:

1. Identificación del cuerpo normativo aplicable en HC e HCE.
2. Propuesta de un modelo arquitectural de microservicios para los diferentes componentes que se integran el proceso de IO de la HCE.

Universidad Tecnológica de Pereira.
Pereira Risaralda, Colombia.
e-mail: jchavar@utp.edu.co

3. Construcción de BSS, Desambiguador, Gestor de Eventos clínicos, visualizador, como los componentes básicos del sistema.
4. Construcción de un modelo territorial.
5. Prueba de la IO en instituciones
6. Como restricción adicional se plantea que deben ser utilizadas tecnologías open source.

49.4 Resultados

Se construyó un modelo de IO soportado en una arquitectura de microservicios, controlada por un BSS, un modelo distribuido de persistencia, y un conjunto de microservicios que soportan cada uno d ellos elementos básicos del sistema de IO

49.5 Conclusiones

1. El modelo de IO propuesto satisface los lineamientos propuestos por el MSPS y el MinTIC.
2. El modelo es extensible a los prestadores privados de servicios de salud vinculados a los territorios departamentales.
3. Las tecnologías open source han sido suficientes para garantizar el modelo de IO.

Agradecimientos Interoperabilidad, Sistemas de información para entidades de salud en Colombia, Arquitectura Historia Clínica Electrónica.
Presentado por Grupo de Investigación en Inteligencia Artificial de la UTP.