

A futuristic white and blue robot with a large, dark visor-like opening on its head. The robot is positioned in the center-right of the frame, facing slightly left. The background is a dark blue, digital space filled with glowing, semi-transparent text fragments such as "tran", "foll", "frien", "lis", "fav", "time", and "geo". The overall aesthetic is high-tech and futuristic.

# *Horizontes en Computación: Estudios Sobre Inteligencia Artificial, Ciencia de Datos y Seguridad*

*María del Carmen Santiago Díaz*

**HORIZONTES EN COMPUTACIÓN:  
ESTUDIOS SOBRE INTELIGENCIA ARTIFICIAL, CIENCIA  
DE DATOS Y SEGURIDAD**

María del Carmen Santiago Díaz

Gustavo Trinidad Rubín Linares

Ana Claudia Zenteno Vázquez

Judith Pérez Marcial

**(Editores)**

María del Carmen Santiago Díaz

**(Coordinadora)**

*María del Carmen Santiago Díaz, Gustavo Trinidad Rubín Linares, Ana Claudia Zenteno Vázquez, Judith Pérez Marcial*  
(editores BUAP)

*María del Carmen Santiago Díaz*  
(coordinadora BUAP)

*Nicolás Quiroz Hernández, Rogelio González Velázquez, Luis Enrique Colmenares Guillén, Pedro García Juárez, Maya Carrillo Ruiz, María del Carmen Santiago Díaz, Beatriz Beltrán Martínez, José Andrés Vázquez Flores, Paola Eunice Rivera Salas, José de Jesús Lavalle Martínez, Julio César Díaz Mendoza, Abraham Sánchez López, Ana Luisa Ballinas Hernández, Meliza Contreras González, Luis Carlos Altamirano Robles, Germán Ardul Muñoz Hernández, Gustavo Trinidad Rubín Linares, María Claudia Denicia Carral, David Eduardo Pinto Avendaño, Carmen Cerón Garnica, Alba Maribel Sánchez Gálvez, María Eugenia Narciza Sully Sánchez Gálvez, Gilberto Lorenzo Martínez Luna, Mario Rossainz López, Pedro Bello López, Guillermo De Ita Luna, Mariano Larios Gómez, Miguel Angel Peña Azpiri, Ana Claudia Zenteno Vázquez, Hermes Moreno Álvarez, Raúl Antonio Aguilar Vera, Osslan Osiris Vergara Villegas, Victor Manuel Morales Rocha, Carina Toxqui Quitl, Alfonso Padilla Vivanco, Heidy Marín Castro, Judith Pérez Marcial, Carlos Soubervielle Montalvo, Abelardo Gómez Andrade.*

(revisores)

Primera edición: 2024

ISBN: 978-607-8975-38-9

**Montiel & Soriano Editores S.A. de C.V.**

15 sur 1103-6 Col. Santiago Puebla, Pue.

BENEMÉRITA UNIVERSIDAD AUTÓNOMA DE PUEBLA

*Rectora:*

Dra. María Lilia Cedillo Ramírez

*Secretario General:*

Mtra. José Manuel Alonso Orozco

*Vicerrector de Investigación y Estudios de Posgrado:*

Dr. Ygnacio Martínez Laguna

*Directora de la Facultad de Ciencias de la Computación:*

M.I. María del Consuelo Molina García

## Contenido

<b>Prefacio.....</b>	<b>6</b>
<b>Sistema inmersivo virtual auxiliar para la capacitación en materia de seguridad laboral</b>	
<i>Jesus Horacio Pacheco Ramirez</i>	
<i>David Rafael Vidal Corona</i>	
<i>Victor Hugo Benitez Baltazar</i>	
<i>Pratik Satam.....</i>	<b>7</b>
<b>Implementación y supervisión de sensores de temperatura en centros de datos mediante IoT</b>	
<i>Bárbara Emma Sánchez Rinza</i>	
<i>José Rodolfo Rosas Lezama</i>	
<i>Carlos Ignacio Robledo Sánchez</i>	
<i>Mario Rossainz López.....</i>	<b>20</b>
<b>Una arquitectura escalable basada en microservicios para la toma de decisiones empresariales</b>	
<i>Hugo Alexis Chiquito Onofre</i>	
<i>Abraham Sánchez López</i>	
<i>Alfredo Toriz Palacios.....</i>	<b>33</b>
<b>Sentiment analysis for Cmiique iitom language</b>	
<i>Jesé Urrutia Romero</i>	
<i>Pedro González Zamora</i>	
<i>Jesús Horacio Pacheco Ramírez</i>	
<i>Victor Hugo Benítez Baltazar.....</i>	<b>44</b>
<b>Aplicación de realidad virtual como apoyo en la terapia conductual a la acrofobia</b>	
<i>Ana Luisa Ballinas Hernández</i>	
<i>Daniel Machorro de Jesús</i>	
<i>María Claudia Denicia Carral</i>	
<i>Gustavo Manuel Miquiz Xolo</i>	
<i>Héctor Medina Cruz.....</i>	<b>52</b>
<b>Recuperación de la Información efectiva y selectiva a través de inteligencia artificial: Una revisión sistemática</b>	
<i>Daniel Arellano García</i>	
<i>Diana Margarita Córdova Esparza.....</i>	<b>63</b>

**Tecnología y derechos humanos: cómo las apps transforman la educación infantil**

*Juan Carlos Cruz Armas*  
*María del Carmen Santiago Díaz*  
*Judith Pérez Marcial*  
*Raúl Antonio Aguilar Vera*  
*Julio César Díaz Mendoza*  
*Ana Claudia Zenteno Vázquez*  
*Gustavo Trinidad Rubín Linares*.....78

**Identificación automática de género a partir de habla en español**

*Brayan Blancas Monsalvo*  
*Maya Carrillo Ruiz*.....90

**Prototipo para la detección de vehículos y motocicletas en puntos ciegos utilizando el modelo YOLO**

*Aldo de Jesus Peña Cuellar*  
*Jesús David Treviño Gandarilla*  
*Abraham Isai Cantu Sanchez*  
*Jose Guadalupe Martinez Herrera*  
*Marco Aurelio Nuño Maganda*.....101

**Uso de diccionarios en Python e interpolación para conocer las tendencias en X sobre el aumento de feminicidios en México**

*María Beatriz Bernábe Loranca*  
*Melissa Isaaly Mendoza Bernábe*  
*Rogelio González Velázquez*.....113

**Reconocimiento de signos del lenguaje de señas mexicano utilizando redes neuronales convolucionales**

*Alejandra Carolina Rodríguez Porras*  
*Luis Antonio Raga Reyes*  
*María Fernanda Parras Pecina*  
*Agustín Alejandro Mota Hinojosa*  
*Marco Aurelio Nuño Maganda*  
*José Hugo Barrón Zambrano*.....124

**Análisis comparativo de herramientas de recuperación de contraseñas en entornos de seguridad**

*Kevin Josaphat Bañuelos Camaño*  
*Judith Pérez Marcial*  
*María del Carmen Santiago Díaz*  
*Hermes Moreno Álvarez*  
*Gustavo Trinidad Rubín Linares*  
*Ana Claudia Zenteno Vázquez*.....135

## **Prefacio**

La inteligencia Artificial se encuentra en un crecimiento tan acelerado que cuesta trabajo imaginarnos todo el potencial que tiene y lo que se está desarrollando para los siguientes años, como el chatgpt, que dicho sirva de paso mucha gente hace unos años dudaba que se podría tener un chat que interactuara con humanos y que además le sirviera no solo para entablar conversaciones inteligentes, sino también generarle documentos que le ayuden en sus tareas cotidianas y que no solo sean las típicas que le ayuden en las actividades escolares de niños y adolescentes, como responder preguntas de geografía, matemáticas, etc. lo cual realmente constituye una gran ayuda porque en nuestra sociedad, en lo general, aunque se tenga un título universitario nunca ha sido garantía de eficiente preparación, pero para finalizar con el chatgpt, en este momento se cuenta con una versión que nos permite generar presentaciones profesionales, edición de imagen y video, etc., cosas realmente increíbles y que aunque se encuentra en la cima del desarrollo de la inteligencia artificial, la competencia es muy fuerte y mientras escribimos esta líneas Google está realizando el lanzamiento de BARD que aunque son herramientas diferentes no dejan de ser una competencia. Pero el desarrollo de estos chats no ha sido trabajo de un par de meses, se sabe que ha llevado varios años, y otros más liberarlo, no solo por aquellas ideas románticas de que pueden destruir el mundo, sino porque se deben limitar las aplicaciones a fin de no crear confusión y mal uso e interpretación de la información que genera, lo cual claramente nos marca una total complicidad con la Ciencia de Datos y la Ciberseguridad, y es que no se conciben tecnologías que estén libres de estos paradigmas, es claro que manejar 170 billones de parámetros y atender a casi 200 millones de usuarios al mes no se puede hacer sin estas 3 tecnologías. Pero afortunadamente todos estos cambios tecnológicos se dan ante nuestros ojos y sentidos, por lo cual ahora requerimos de un acelerado aprendizaje, así es, las tecnologías que nos ponen al alcance no solo requieren sino necesitan de una capacitación para utilizarlas en su máximo potencial a fin de realmente obtener los frutos para lo que han sido diseñadas, y es en este punto que no hay esfuerzos vanos, todos los esfuerzos deben orientarse en aplicar las nuevas tecnologías, desarrollarlas, difundirlas y sobre todo ponerlas incluso en una encrucijada cuestionándolas, para eso son, y es que para nosotros como usuarios de estas tecnologías y parte de una comunidad académica y científica, es una responsabilidad mayor pues se espera que nuestra visión objetiva oriente y potencialice estas herramientas para nuestros estudiantes y en general a la sociedad que los utilizará.

Aunque hemos hablado en estas líneas del tan acertado desarrollo realizado por OpenIA del chatgpt, y en donde los reflectores se encuentran ahora, sobre la IA aún hay muchísimo más, como aplicaciones de IA en medicina, en bienes y servicios, en la industria, en la educación, etc. realmente está en todos lados, no se puede imaginar un sector donde no se encuentre la Inteligencia Artificial y la Seguridad Informática, por lo que este libro es nuestra aportación y en el cual hemos puesto un especial cuidado de los trabajos que se presentan buscando la calidad que se requiere a fin de que sean referencia de otras aportaciones y fuente de conocimiento que despierte la creatividad de estudiantes e investigadores.

**María del Carmen Santiago Díaz  
Gustavo Trinidad Rubín Linares**

## **Sistema inmersivo virtual auxiliar para la capacitación en materia de seguridad laboral**

### **Immersive virtual assistive system for training in occupational safety**

Jesús Pacheco<sup>1</sup>, David Rafael Vidal Corona<sup>1</sup>, Victor H Benitez<sup>1</sup> and Pratik Satam<sup>2</sup>

<sup>1</sup>Universidad de Sonora, Blvd. Luis Encinas y Rosales S/N. Col. Centro. C.P. 83000.  
Hermosillo, Sonora, México.

<sup>2</sup>The University of Arizona, 1200 E University Blvd, Tucson, AZ 85721, USA

jesus.pacheco@unison.mx, david.vidal@unison.mx, victor.benitez@unison.mx,  
pratiksatham@arizona.edu

#### **Abstract**

This work introduces a virtual reality system designed to improve immersion in occupational safety training, focusing on enhancing knowledge retention about proper PPE use. The methodology includes defining system requirements, developing virtual and interactive elements, implementing a data acquisition process, and conducting a proof of concept. An adaptive algorithm adjusts difficulty based on real-time user performance, creating a customized training experience. Results show that the system effectively boosts user engagement and understanding of PPE usage, with positive feedback and increased correct application of PPE observed during training. The adaptive difficulty settings maintained an optimal challenge level, improving the overall learning experience without overwhelming users. These findings underscore the potential of virtual reality as a valuable tool in occupational safety training, with opportunities for further refinement of algorithms and expanded immersive scenarios to cover a broader range of safety protocols.

#### **Resumen**

Este trabajo presenta un sistema de realidad virtual para mejorar la capacitación en seguridad laboral, enfocado en la correcta utilización del equipo de protección personal (EPP). La metodología abarca la definición de requisitos, desarrollo de elementos virtuales, adquisición de datos y prueba de concepto. El sistema utiliza un algoritmo adaptativo que ajusta la dificultad según el rendimiento del usuario en tiempo real, ofreciendo una experiencia personalizada. Los resultados muestran un aumento significativo en el compromiso y comprensión del uso del EPP, con mejoras en la correcta aplicación durante la capacitación. El nivel de dificultad se adapta de manera óptima, favoreciendo el aprendizaje sin sobrecargar a los usuarios. Estos resultados subrayan el potencial de la realidad virtual en la capacitación en seguridad, con posibilidades de mejorar algoritmos y expandir los escenarios a más protocolos de seguridad.

---

Keywords and phrases: Virtual Reality, Dynamic Difficult Adjustment, Automaton, Industrial Security, Protective Personal Equipment.

---

## **1 Introduction**

Virtual reality (VR) implementations in industry have been boosted by progressive technological advances in recent years and are of great research interest in care applications and promoting a better culture of industrial safety. In [1] it is proposed a method of training and evaluating employees in the industrial sector for cultural strengthening in care and safety in work environments, carrying out an evaluation of the impact of said system compared to other methods.

With a similar approach, measurements of its effectiveness and validation of the benefits of similar implementations were carried out as an effective method to improve performance in police training compared to intensive training with conventional practices without the presence of this kind of systems [2].

Another application is proposed by [3], presenting a virtual reality trainer for fires considering the evaluation of international safety standards regarding such events. His work evaluates and exhibits the performance of participants immersed in the system compared to others who do not use the equipment. In addition to the results in favor of the immersive training system, of which there was no need to place users in risky conditions, it is shown that more biometric tests can be evaluated to measure the effectiveness of the system.

This work contributes significantly to the field by developing a virtual reality system tailored for occupational safety training, enhancing user immersion and facilitating improved retention of safety protocols. A key aspect of this project is the implementation of NOM017-STPS-2008, the Mexican standard for the selection, use, and maintenance of personal protective equipment, marking it as the first of its kind to integrate these regulatory guidelines into a VR training environment. The novel data acquisition framework introduced in the system allows for real-time adjustment of parameters based on user feedback, providing a more dynamic and responsive training experience that goes beyond traditional methodologies. By aligning with NOM017-STPS-2008, this work not only sets a benchmark for compliance-based safety training but also paves the way for future VR applications that can improve safety outcomes and adherence to legal standards in industrial settings.

The remainder of the document is structured as follows: Section 2 reviews the related literature; Section 3 details the methodology, including system design and data collection; Section 4 presents a discussion and analysis of results; and Section 5 concludes with implications, limitations, and suggestions for future research direction.

## **2 Literature review**

Immersive technology emulates the physical world through a simulation or digital world by creating a sensation of perception of the environment and, consequently, a sensation of immersion, allowing the user to interact with the virtual environment [4]. These characteristics make virtual reality useful for the industrial sector, as presented in the work developed by [5], where a literary review of virtual reality applications in the industrial sector is carried out, emphasizing its use for the automotive and aviation industries.

Usually, intelligent systems help with the assistance of specialized duties in professional areas such as the medical field, as well as tedious or high-risk tasks for humans [6]. Having an interest in coupling these technologies in occupational safety training, it must be kept in mind that the causes of accidents during the working day may be due to both individual issues and characteristics of the environment or behaviors external to the worker in question. It is mentioned by [7] that accidents in the workplace can originate from immediate causes, which generate them instantly and derive from unsafe or inappropriate acts by the worker or from unsafe conditions. That is why it is important that users keep these issues in mind to carry out their corresponding tasks.

## **3 Methodology**

The methodology adopted for this project builds upon the approaches used in developing AR applications as proposed by [8] and the stages outlined in the initial study for predicting player skills using Dynamic Difficulty Adjustment (DDA) by [9]. This methodology is divided into four main stages, as shown in Figure 1.

### **Stage 1: Preliminary definition of the system**

In the first stage, we establish the foundational aspects of the system by defining its primary characteristics. This involves identifying the specific case study for the project, which in this instance focuses on training for the correct use of personal protective equipment (PPE) in an industrial setting. We also outline the key activities that users will engage in, such as inspecting virtual workplaces and evaluating the PPE usage by virtual workers. Additionally, this stage specifies the necessary interactive elements (e.g., PPE selection and adjustment tools) and environmental components (e.g., simulated industrial scenarios) required to achieve a realistic and immersive training experience.

### **Stage 2: Technological development**

The second stage involves the technological development of the system. During this phase, detailed virtual representations of all system elements are created using advanced 3D modeling software. This includes designing virtual workplaces, PPE items, and virtual workers. The system's interactivity is developed by programming the interactions between the users and the virtual elements, allowing users to navigate the environment, inspect PPE on workers, and make adjustments as needed. This

stage utilizes game development platforms such as Unity to integrate the 3D models with user interaction capabilities, creating a seamless and immersive virtual environment.

#### Stage 3: Data collection and analysis

In the third stage, we focus on data collection and analysis to fine-tune the system's adaptive capabilities. We define several parameters, such as the number of workers in a scenario and the probability of each worker initially wearing correct PPE. These parameters are adjusted to create different difficulty levels within the training simulations. User performance data, including time taken to complete tasks and accuracy in identifying PPE issues, is collected across various difficulty configurations. This data is then analyzed to develop an adaptive algorithm that dynamically adjusts the difficulty settings based on the user's performance, aiming to optimize the training experience for each participant.

#### Stage 4: Proof of concept

The final stage is the proof of concept, where the complete system, including the adaptive algorithm, is rigorously tested. This involves running simulations with users to evaluate how effectively the system adapts difficulty levels in response to user performance. We assess the algorithm's capability to provide appropriate challenges that enhance learning without causing frustration. Feedback from this stage is crucial, as it informs further refinements and ensures that the system's adaptive mechanisms are robust and capable of delivering personalized training experiences that meet the objectives of occupational safety education.

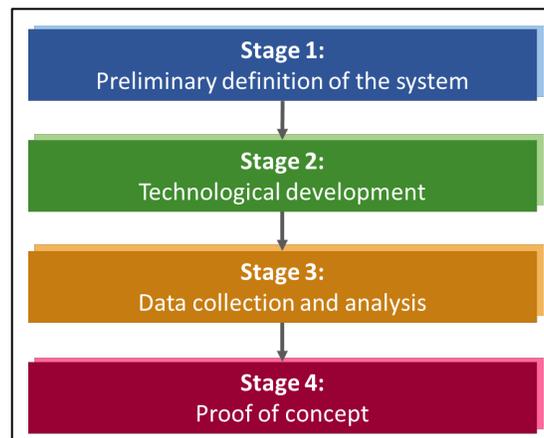


Figure 1. Used methodology for the development of the system.

For the elaboration of this project, the required tools were a computer unit with VR capabilities, being the PC hardware specifications for the used device, a Ryzen 7 3700X processor, 32 GB of Random Access Memory, and a Nvidia GeForce RTX 3050 Graphics Processing Unit. This work used the HTC VIVE VR Headsets with the corresponding sensors, and the system was developed in Unity

version 2021.3.16f1, along with Visual Studio 2022 Community Edition. For three-dimensional representations, Blender 3.1.2 was employed. Python 3.10 was also used for data analysis. The physical space required was a room with a base of 3 square meters and a height of 2.5 meters in dimensions, as shown in Figure 2.

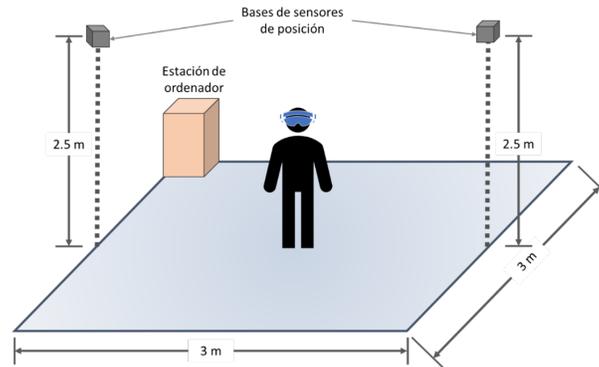


Figure 2. Environmental space for the development of the activity of the system.

#### 4 Discussion and analysis of results

From the methodology described in Figure 1, the first stage for this project defined the case of study of the system as the training in the identification of faults of the use of Personal Protective Equipment (PPE), according to the available information regarding this subject from the Mexican norm NOM017-STPS-2008. The activity consisted of conducting an inspection of the workplaces and evaluating the PPE used by the virtual workers in an industrial environment. If the user witnessed any problems, the user had to select the agent or agents in question and make the necessary adjustments to said equipment. The objective was to facilitate the reinforcement of the theoretical knowledge about the correct use of the PPE taught in previous trainings and capacitation in a practical way with the use of this system. The required elements defined for this project were the interactive elements, which the user needed to interact with to complete the activity, and the remaining elements, whose purpose was to help with the immersion to the system.

With the basic characteristics of the project described, the technological development began with the three-dimensional representation of each of the elements of the system. The industrial representation achieved for the workplace can be seen in Figure 3.

For the first state of the project, four types of industrial workplaces were implemented, being those the virtual representations of the manufacturing, electrical, robotics and warehouse areas. Each type of scenario had four different samples, being a total of 16 scenes where the system was played. In Figure 4 illustrates the machinery and equipment for each of these types of scenarios.

The virtual workers were the main components of the system that were represented, as well as the PPE placed in each corporal section. The interaction with VR technology was also implemented to

select these workers. The representation of these components and the selection that takes place can be found in [5].

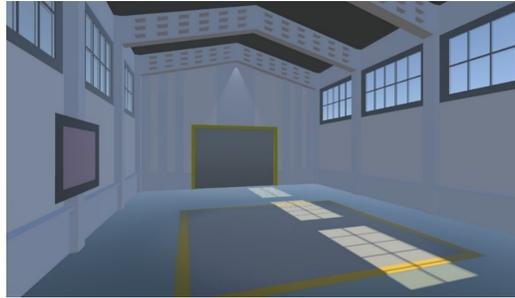


Figure 3. Industrial workspace representation where the user was situated.

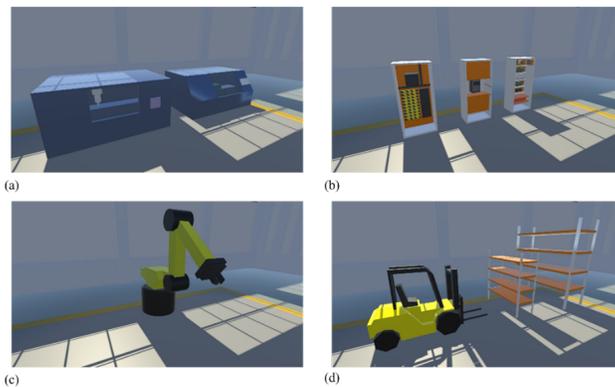


Figure 4. Machinery and equipment for each type of industrial workplace. (a) Manufacture components, (b) Electrical components, (c) Robotic components, and (d) Warehouse components.

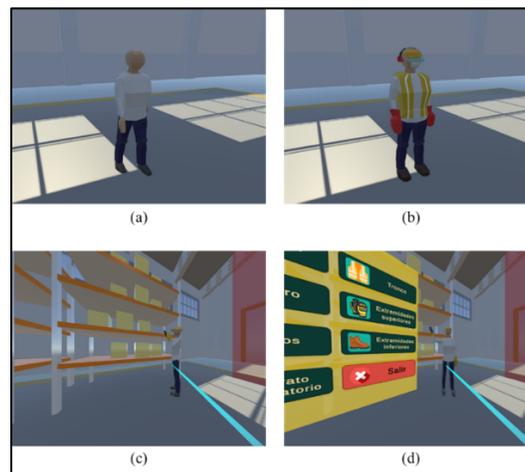


Figure 5. Representation of virtual workers. (a) Virtual worker without equipment, (b) Virtual worker with equipment placed, (c) the worker prior to being selected and (d) the worker after making the selection.

Finally, the interactive panels that allowed the control of certain system operations were represented, such as the control of the initial configuration, the culmination of the activity for the scenario taking place, and the manipulation of the PPE of the virtual workers. These panels are illustrated in Figure 6.



Figure 6. Interactive panels are used to control certain system operations.

With the initial achieved representation of all the required elements and the working VR implementation, the third stage of the methodology was started. The parameters considered to establish the different difficulty configurations of the system were the following:

- $(n_w)$  Number of workers in the workplace. This parameter set the number of workers that would show up at workstations at random locations, which domain was given by the Equation (1):

$$A_n(n_A) = \{n_A \in \mathbb{N} | 1 \leq n_A \leq 5\} \quad (1)$$

Where  $A_n(n_A)$  represents a set of natural numbers associated with the parameter  $n_A$ , a variable representing the number of workers in the workplace.  $\mathbb{N}$  is the set of natural numbers, ranging from 1 to 5.

- $(P_G)$  Probability of having a starting gear per corporal group: This parameter was used in the decision making to place the correct PPE in a determined area of the body at the start of the scene, being an independent decision for each one of the body parts. The domain for the possible values was given by Equation (2):

$$P_E(p) = \{p \in \mathbb{R} | 0 \leq p \leq 1\} \quad (2)$$

Where  $P_E(p)$  represents a set of real numbers associated with the parameter  $p$ , which is a probability value, which belongs to the set of real numbers. However, for this project the possible values were limited to the extreme values and the quartiles in Equation (3):

$$P_E(p) = \{0, \quad 0.25, \quad 0.50, \quad 0.75, \quad 1\} \quad (3)$$

Having defined these parameters, the system was used by a group of 22 volunteers to collect information on the user's performance in different random configurations of difficulty. Each

participant carried out the activity in 3 or 5 scenarios, as shown in Figure 7, collecting a total of 82 samples.

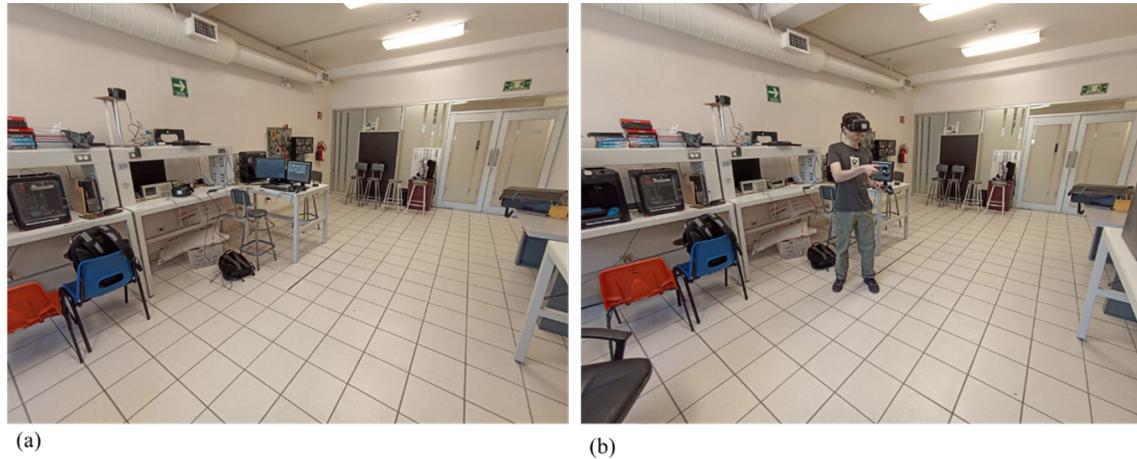


Figure 7. (a) General setup. (b) The system being used by a participant.

The information collected was the time each participant took to complete the activity and the performance on the rate hits for the correct PPE per working agent on the given scenario. Table 1 shows the statistical description of these results.

Table 1. Statistical description of the obtained results in the third stage of the project.

Statistic	Time taken	Hits percentage
Mean ( $\bar{x}$ )	68.93	78%
Standard deviation ( $\sigma$ )	34.93	17%
Minimum value ( $x_{min}$ )	27.00	25%
Maximum value ( $x_{max}$ )	208.00	100%

In Table 2, the information retrieved about the performance of the participants was grouped in sets of difficulty according to the probability of starting with the correct PPE. Both time and rate hits are analyzed. The data regarding the rate hits is also plotted in Figure 8. Even though it is not possible to conclude that there is a trend that relates the difficulty groups with the performance obtained by the users, it can be noted that there are a few variations between each of the mentioned sets of difficulties, in which either the mean value of the rate of hits is decreased, or the standard deviation is increased.

The finite automaton illustrated in Figure 8 was proposed with the previous analysis. The groups previously described are labeled in each of the states of this diagram, starting the system with the second set of difficulties. To make a transition between states, the program needed a certain weight to evaluate which will be the next group. For it, the performance of the user was reviewed in each

of three iterations, adding weight when the rate of hits was above 80 percent, effectuating the transition depending on the final value and the current state.

Table 2. Statistical information of the results obtained in the third stage of the project is grouped by set of difficulty.

	Groups									
	$G_1 \rightarrow P = 1.00$		$G_2 \rightarrow P = 0.75$		$G_3 \rightarrow P = 0.50$		$G_4 \rightarrow P = 0.25$		$G_5 \rightarrow P = 0.00$	
	Time	Hits								
$\bar{x}$	70.28	0.76	72.13	0.72	53.42	0.86	76.94	0.82	67.32	0.78
$\sigma$	40.29	0.24	35.67	0.15	13.85	0.16	46.32	0.13	24.18	0.16

Once the finite automaton was implemented, a second group of 6 participants used the system to observe the behavior of this feature. Each user was placed in 12 different scenarios, for a total of 72 samples. Table 3 describes the general information of the obtained performances.

The results in Table 4 are similar to those in Table 3, show no clear patterns between the difficulty levels and user performance, indicating that the variations in difficulty did not significantly affect how participants performed. This could be due to several factors, including high variability in individual participant skills and adaptability, which may overshadow the effects of difficulty adjustments. Additionally, the adaptive algorithm might lack sufficient sensitivity or might adjust parameters in ways that do not meaningfully impact performance. Furthermore, the difficulty levels may not be distinctly calibrated, resulting in minimal perceived differences by users. Participants could also be quickly adapting to the tasks, effectively neutralizing the impact of varying difficulty levels. Finally, the task complexity and user engagement might play a role, where participants focus more on task completion than on the nuances of difficulty, leading to similar performance outcomes across different sets.

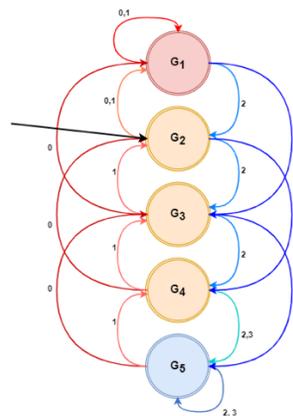


Figure 8. A finite automaton is proposed to adjust the system's difficulty settings according to users' performance.

Table 3. Statistical description of the obtained results in the fourth stage of the project.

Statistic	Time	Hits percentage
Mean ( $\bar{x}$ )	48.94	83%
Standard deviation ( $\sigma$ )	23.89	14%
Minimum value ( $x_{min}$ )	10.00	50%
Maximum value ( $x_{max}$ )	122.00	100%

Table 4. Statistical information of the obtained results in the fourth stage of the project grouped by difficulty sets.

	Groups									
	$G_1 \rightarrow P = 1.00$		$G_1 \rightarrow P = 0.75$		$G_1 \rightarrow P = 0.50$		$G_1 \rightarrow P = 0.25$		$G_1 \rightarrow P = 0.00$	
	Time	Hits								
$\bar{x}$	34.20	0.81	43.27	0.84	42.11	0.78	68.62	0.86	45.17	0.83
$\sigma$	7.03	0.15	12.57	0.14	30.00	0.16	26.32	0.14	19.29	0.16

Finally, the transitions from the system using the algorithm from the finite automaton were analyzed. In Figure 9, the first pair of users had mostly outstanding performance on each stage of the program, and the system adjusted to the maximum group of difficulties for these participants.

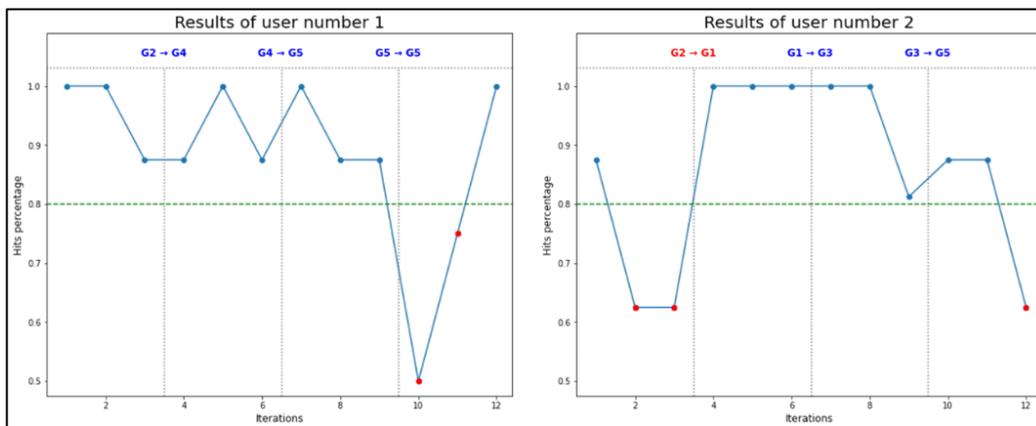


Figure 9. The first pair of graphs will show the performance of the user in the final stage.

In Figure 10, it can be noted that the users had good performance when using the system but also had some setbacks, having the system adapt to such conditions to ensure low levels of frustration from the users by using this software.

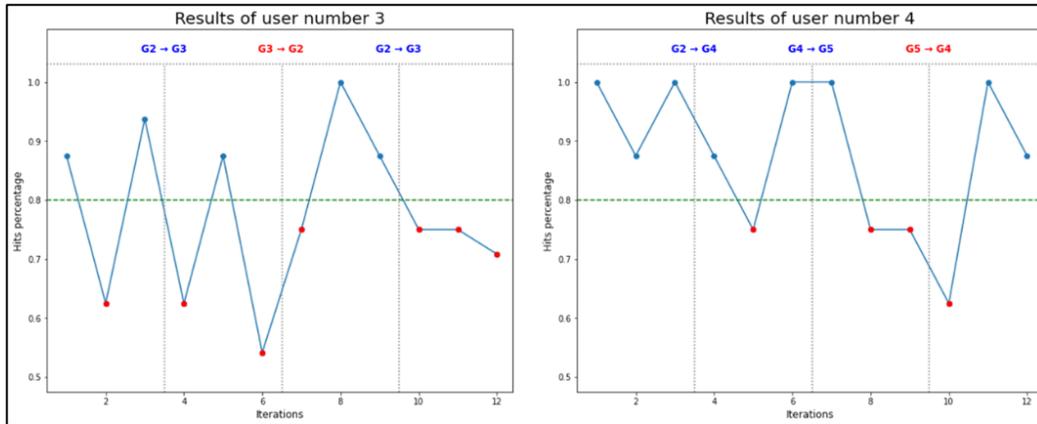


Figure 10. The second pair of graphs shows the performance of the user in the final stage.

Figure 11 shows a few setbacks due to low user performance, meaning that the system had to adapt to lower levels of difficulty for the users to have a better experience.

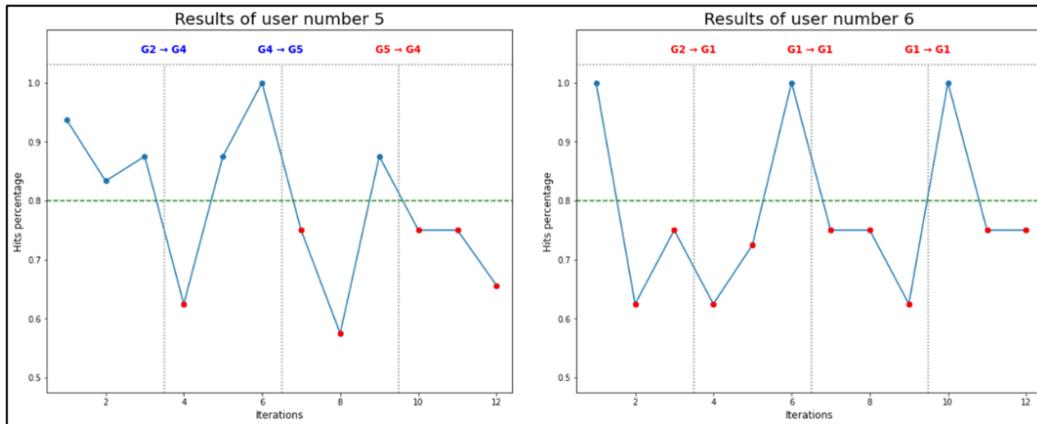


Figure 11. The third pair of graphs shows the performance of the user in the final stage.

## 5 Conclusions

Based on the results obtained from this project, it can be concluded that virtual and augmented reality technologies offer extensive applications beyond the entertainment sector. These tools can significantly enhance pedagogical processes, various training programs, and specialized educational activities by complementing existing methods and improving user learning outcomes.

This work successfully developed and utilized an effective methodology for creating a prototype that demonstrated considerable potential in occupational safety training. A significant aspect of this project is its alignment with NOM017-STPS-2008, the Mexican standard for the selection, use, and maintenance of personal protective equipment, which adds a layer of regulatory relevance and practical applicability to the training. Feedback from participants indicated high levels of engagement throughout the various stages of the methodology, underscoring the system's ability to capture and maintain user attention. The virtual environment effectively simulated different

work areas, leading to a heightened degree of immersion in the training activities. The integration of NOM017-STPS-2008 ensures that the training is not only immersive but also compliant with legal standards, enhancing the credibility and impact of the system. The analysis of participant performance also revealed slight variations, which provided insights into refining the system's parameters for an improved user experience, ensuring that the training is both effective and aligned with industry regulations.

For future implementations and recommendations, it is crucial to consider the feedback and suggestions from participants, particularly regarding the enhancement of performance feedback in each scenario and the improvement of the three-dimensional representations of work environments and their components. Addressing these aspects will likely enhance user immersion and responsiveness of the system to user performance, thereby positively influencing the learning outcomes.

Additionally, improvements in the methodological approach are recommended for future work, including increasing the number of iterations or cycles for all development stages to facilitate continuous improvement. Emphasis should be placed on refining the technological components of the system and expanding user participation to gather more comprehensive feedback. Collecting a larger dataset would enable more robust statistical analysis, considering a broader range of participants and factors such as physiological, socioeconomic, and cultural variables relevant to the target population.

Furthermore, exploring advanced algorithms and adaptive systems, particularly those involving specialized artificial intelligence techniques, could optimize the system's difficulty adjustments and overall performance. Incorporating these advancements with a comparative study of user knowledge acquisition with and without the system will help identify the differences, benefits, and areas of opportunity for the project, guiding future enhancements and broader applications.

## **Acknowledgements**

The authors of this article thank the CONAHCYT for the support provided during the studies carried out that made us able to develop this project, through the program for postgraduate scholarship.

## **References**

- [1] Ján Lacko, "Health Safety Training for Industry in Virtual Reality," 2020.
- [2] Saunders Jonathan, Bayerl Petra Saskia, Davey Steffi, and Lohrmann Philipp, "Validating Virtual Reality as an Effective Training Medium in the Security Domain," in IEEE Conference on Virtual Reality and 3D User Interfaces, 2019.
- [3] Torres-Gerrero Francisco, Sanchez Romeo, and Neira-Tovar Leticia, "Virtual Reality Trainer in the Evaluation of International Safety Standards in Fire Situations," in IEEE, 2020.

- [4] P. A. S. Saraswathi, A. Pavithra, J. Kowsalya, S. K. Priya, G. Jayasree, and T. K. Nandhini, "An Emerging Immersive Technology-A Survey Article in," *International Journal of Innovative Research & Growth*, 2020, [Online]. Available: <https://www.researchgate.net/publication/338819764>
- [5] A. G. da Silva, M. V. M. Gomes, and I. Winkler, "Virtual Reality and Digital Human Modeling for Ergonomic Assessment in Industrial Product Development: A Patent and Literature Review," *Applied Sciences (Switzerland)*, vol. 12, no. 3. MDPI, Feb. 01, 2022. doi: 10.3390/app12031084.
- [6] M. Molina, "What is an intelligent system?," Aug. 2022, [Online]. Available: <http://arxiv.org/abs/2009.09083>
- [7] P. Guerra, D. Viera, D. Beltrán, and S. Bonilla, *Seguridad industrial y capacitación: un enfoque preventivo de salud laboral*. 2021.
- [8] G. Pantoja, L. E. Garza, and E. G. Mendivil, "Augmented reality in pneumatic conveying system: Fuller pump dry material line charger," in *Iberian Conference on Information Systems and Technologies, CISTI*, IEEE Computer Society, 2014. doi: 10.1109/CISTI.2014.6877050.
- [9] R. Rodriguez, "Player Modeling for Dynamic Difficulty Adjustment in Top Down Action Games," 2019.

## **Implementación y supervisión de sensores de temperatura en centros de datos mediante IoT**

### **Implementation and monitoring of temperature sensors in data centers using IoT**

Bárbara Emma Sánchez Rinza<sup>1</sup>, José Rodolfo Lezama<sup>2</sup>, Carlos I. Robledo Sánchez<sup>1</sup>,  
Mario Rossainz López<sup>2</sup>.

<sup>1</sup>Facultad de Físico Matemáticas, Benemérita Universidad Autónoma de Puebla.  
14 sur y avenida San Claudio. C.P. 7200 Puebla, Pue., México.

<sup>2</sup>Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla,  
14 sur y avenida San Claudio. C.P. 7200 Puebla, Pue., México.

barbara.sanchez@correo.buap.mx, 223470502@viep.com.mx,  
carlos.robledo@correo.buap.mx, mrossainzl@gmail.com.

#### **Abstract**

This paper presents the temperature monitoring and control of a "SITE" data center in manual and automatic modes, aimed at the hotel industry. In addition, a ventilation failure notification system is integrated. For this project, the ESP32 module was used, establishing communication through its Bluetooth and Wi-Fi interfaces. This module was programmed to continuously sense the temperature, notify its changes and automatically control the ventilation when the temperature exceeds the preset limit. The ventilation can also be manually activated. The administration of the proposed system is carried out by an Android device that uses the ESP32 module with a generic Bluetooth communication application to manually monitor and control the temperature, in addition to automatic control through the ESP32 module. The real-time alert system, through email, reports on faults detected in the temperature sensor and notifies both the start and recovery of the sensor when it is ready.

#### **Resumen**

En este trabajo se presenta el control y monitoreo de la temperatura de un centro de datos "SITE" en modos manual y automático, orientado a la industria hotelera. Además, se integra un sistema de notificación de fallos en la ventilación. Para este proyecto se utilizó el módulo ESP32, estableciendo comunicación mediante sus interfaces Bluetooth y Wi-Fi. Este módulo fue programado para censar la temperatura continuamente, notificar sus cambios y controlar automáticamente la ventilación cuando la temperatura supera el límite preestablecido. La ventilación también puede ser accionada manualmente. La administración del sistema propuesto se realiza mediante un dispositivo Android

que utiliza el módulo ESP32 con una aplicación de comunicación Bluetooth genérica para monitorear y controlar manualmente la temperatura, además del control automático a través del módulo ESP32. El sistema de alerta en tiempo real, a través de correo electrónico, informa sobre fallas detectadas en el sensor de temperatura y notifica tanto el inicio como la recuperación del sensor cuando esté listo.

---

Keywords and phrases: Arduino, Wifi, Esp-32.

---

## **1 Introducción**

El Internet de las Cosas (IoT) se encarga de la interconexión de diferentes tipos de dispositivos a través de internet, permitiendo la recopilación y el intercambio de datos. Lo que actualmente conocemos como IoT comenzó con la posibilidad de la interconexión de dispositivos para compartir paquetes en la década de 1970. Uno de los primeros dispositivos considerados como IoT fue una máquina expendedora de Coca-Cola que reportaba el estado de los refrescos que contenía. Aunque el término aún no había sido inventado, el dispositivo ya pertenecía a esta categoría. El término "Internet de las Cosas" fue acuñado en 1999 por el Auto-ID Center en el MIT. A partir de 2010, comenzó la expansión de los dispositivos IoT cuando las empresas empezaron a desarrollar productos comerciales y se lanzaron las primeras plataformas IoT. Con la integración de gigantes de la tecnología como Google (Nest Labs) y Amazon (Amazon Echo), se inició la incorporación de IA, permitiendo el aprendizaje y la recolección masiva de datos para el entrenamiento de modelos de IA [1,2].

Uno de los principales objetivos del IoT es eficientar varios tipos de procesos y sectores como energía, recursos, salud, movilidad, entre otros. Los dispositivos IoT están equipados con sensores y tecnologías de comunicación para interactuar entre sí o con sistemas de control centralizados. Estos dispositivos recopilan datos del entorno, como temperatura, movimiento, humedad, y luz. Pueden trabajar en conjunto o de manera independiente, independientemente del fabricante, y pueden ser controlados y monitoreados a distancia a través de aplicaciones móviles u otras interfaces digitales.

Para este proyecto se utilizó el módulo ESP32, una solución de Wi-Fi/Bluetooth que proporciona un procesador integrado con interfaces para conectarse con varios periféricos. Este procesador de dos núcleos tiene frecuencias operativas entre 80 y 240 megahercios (MHz). En este proyecto, el ESP32 se utilizó para implementar un sistema de monitoreo de temperatura y humedad dentro de una sala de servidores (SITE), donde se requiere el monitoreo constante y el control de la temperatura para el correcto funcionamiento de los servidores ubicados en los racks. El módulo ESP32 funcionará como un servidor que proporciona datos, en este caso, las temperaturas y la humedad del SITE

donde se instale el sistema de monitoreo y control de temperatura, comunicándose a través de su interfaz Bluetooth con algún dispositivo de lectura y control.

Los dispositivos IoT admiten tres capas:

- La capa física, que comprende sensores y actuadores.
- La capa de red, que interconecta los dispositivos a través de redes hasta los usuarios.
- La capa de aplicación, que conecta las capas físicas y de red digitalmente y facilita las interacciones entre el usuario y otras máquinas, permitiendo a las personas tomar decisiones a un nivel superior de control o automatizar tareas seleccionadas.

El monitoreo de la temperatura del SITE donde se instale este sistema permitirá a los dispositivos en él (servidores de control, monitoreo y almacenamiento) alcanzar el mejor desempeño posible según las indicaciones de los fabricantes, quienes establecen las condiciones físicas para la instalación de sus equipos de acuerdo a la carga de trabajo que realizarán. Dado que el procesamiento de grandes volúmenes de datos es requerido, los servidores que realizan estas actividades generan grandes cantidades de calor en diferentes instantes de tiempo, por lo cual un control de la temperatura en tiempo real no solo es necesario sino vital.

## 2 Implementación

La implementación se realizó con la tarjeta EPS32 que se muestra en la Figura 1, el módulo de temperatura DTH11 mostrado en la Figura 2, un ventilador genérico (Figura 3) TN4012MY-S1 y el Transistor BC547B-NPN mostrado en la figura 4.

### • Tarjeta EPS32

ESP32 es una plataforma desarrollada por Espressif Systems Company (China), ver figura 1, tiene las siguientes características:

- Un procesador de doble núcleo,
- Conectividad Wi-Fi y Bluetooth integrada,
- Una gran cantidad de pines de entrada/salida de uso general (GPIO) y
- Baja potencia consumo. ESP32 está equipado con un Tensilica de doble núcleo Xtensa Microprocesador LX6, a su vez cuenta con las siguientes interfaces Wi-Fi y Bluetooth integradas que simplifican la conexión y comunicación con otros dispositivos o redes. Es compatible con los protocolos Wi-Fi, 802.11 b/g/n, además de proporcionar Bluetooth Classic y opciones de conectividad Bluetooth de bajo consumo (BLE). Este módulo cuenta con varios pines GPIO que facilitan la conexión, el control de dispositivos y sensores externos, también admite una variedad de interfaces, como: SPI, I2C, UART y PWM.

Para la programación de este módulo utilizaremos C++ utilizando el Arduino IDE o Platform IO. Además, el ESP-IDF (Espressif IoT Development Framework) proporciona un conjunto completo de bibliotecas y herramientas específicas para el desarrollo de ESP32.

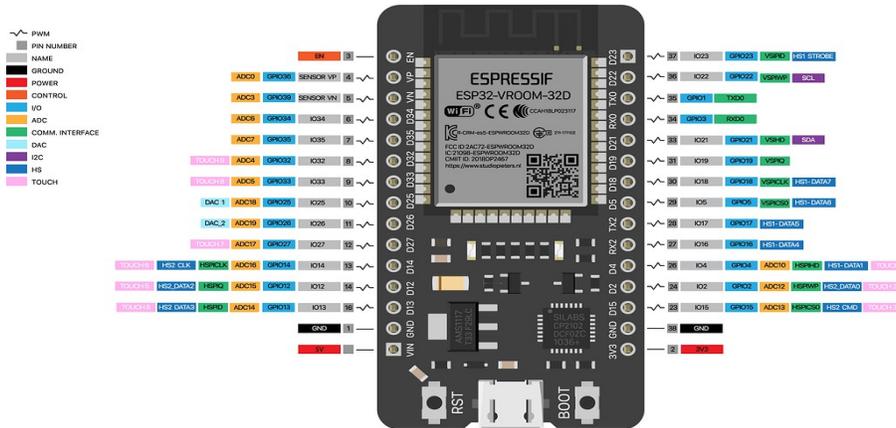


Figura 1. Tarjeta EPS32.

- **Módulo de temperatura DTH11**



Figura 2. Sensor DTH11.

El módulo DHT11 o KY-015 es un sensor de temperatura y humedad de salida de señal digital, de tamaño compacto y bajo consumo de energía, dispone de librerías para Arduino con soporte para el protocolo «Single bus». En cuanto al hardware, solo es necesario conectar el pin VCC de alimentación a 3 o 5V, el pin GND a Tierra (0V) y el pin de datos a un pin digital. Algunas de sus principales características son:

- Voltaje de funcionamiento: 3.5 V a 5.5 V,
- Rango de medición de humedad: 20% a 90% RH (Error de medición de humedad: +-5%),
- Rango de medición de temperatura: 0 °C a 50 °C, Error de medición de temperatura: +-2 °,
- Peso: 3 g.

- **Ventilador**

Este es un ventilador de enfriamiento genérico utilizado generalmente para enfriar fuentes de poder, para este proyecto se utilizó en la demostración del correcto funcionamiento del sistema. Este ventilador funciona con DC 12V 0.10A, 2-Wire y también es conocido con el nombre: TN40L2MY-SL.



Figura 3. Ventilador genérico TN4012MY-S1.

Este es un dispositivo electrónico de estado sólido de unión bipolar BJT (Bipolar Junction Transistor) que utiliza las propiedades del silicio para amplificar señales de voltaje o corriente. Este transistor es de tipo «NPN», formado por dos capas de material tipo «N», separadas por otra de tipo «P». El BC547B NPN está protegido por un encapsulado de plástico color negro conocido como TO-92 y cuenta con 3 pines que son base, colector y emisor, donde el emisor se encarga de inyectar electrones, la base permite transferir o pasar los electrones y el colector se encarga de coleccionar electrones. Este transistor es de propósito general, y de bajo costo, aunque es lo suficientemente robusto. Para este proyecto se utilizó como amplificador de señales para controlar el ventilador.

Características:

- Tipo de Transistor: NPN
- Serie: BC547B
- Encapsulado: TO-92
- Pines: 3
- Voltaje máximo en las uniones:
- Emisor-base: 6V
- Colector-emisor: 45V
- Colector-Base VCBO: 50V
- Voltaje de saturación colector-emisor: 250 mV
- hfe: 330
- Corriente de colector: 100 mA DC
- Capacitancia de entrada: 10 pF
- Capacitancia de salida: 1.7 pF
- Frecuencia Máxima de Funcionamiento: 300 MHz
- Temperatura de Funcionamiento: -55 °C a 150 °C
- Dimensiones: 18 mm x 4.19 mm x 3.45 mm
- Peso de la unidad: 0.18 g

- **Transistor**

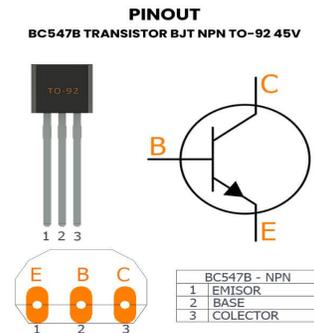


Figura 4. Transistor BC547B-NPN.

### 3 Instalación

Para comenzar con el desarrollo del proyecto es necesario iniciar todo el proceso con la instalación de los drivers del módulo [3] para que la PC reconozca al módulo ESP32.

Después de la instalación de los drivers, instalamos el Ide de Arduino [4] el cual se encuentra disponible desde la página de Arduino. Una vez concluidas estas instalaciones comenzaremos con el desarrollo del programa de control para el módulo ESP32.

Las bibliotecas que utilizamos en este proyecto fueron las siguientes:

- DHT.h
- WiFi.h
- BluetoothSerial.h
- ESP\_Mail\_Client.h

Las cuales son necesarias para la correcta comunicación entre nuestro módulo ESP32 con el sensor de temperatura/humedad y del módulo con el exterior a través de sus interfaces Bluetooth y WiFi.

### 4 Programa

El programa comienza inicializando las variables globales necesarias, incluyendo la salida para el control del ventilador (Pin 23), los datos de conexión Wi-Fi y los datos de correo electrónico (correo de origen asociado al módulo ESP32 mediante un api key, y correo de destino para reportar las fallas del sistema). Luego, se inician los servicios de Bluetooth y Wi-Fi para comunicarse con el módulo, utilizando los parámetros previamente definidos para conectarse a la red Wi-Fi y reportar fallos del sistema, así como para notificar el estado inicial y las reparaciones completas del sistema.

El programa realiza el censado periódico de temperatura y humedad cada 3 segundos. Sin embargo, está diseñado para reportar solo los cambios en estas magnitudes; mientras permanezcan constantes, la consola de control mostrará un '.' para indicar el correcto funcionamiento sin saturar la consola con datos innecesarios. Los cambios se muestran en la consola y se envían a través de

Bluetooth si hay un dispositivo disponible para recibirlos. Además, al iniciar, el sistema enviará automáticamente un correo electrónico de notificación [5].

El programa también incluye el modo de ventilación manual, que permite cambios rápidos de temperatura sin importar la temperatura actual. Este modo solo se activa a través de Bluetooth y tiene prioridad sobre el modo automático, desactivando el control automático de temperatura mientras esté activo.

Cuando se detecta un fallo en el sensor de temperatura, se envía un correo electrónico de notificación a los destinatarios registrados. Del mismo modo, cuando se repara o reemplaza el sensor, se reporta el evento a través del correo electrónico [6].

## 5. Operación

- 1) Para iniciar el sistema basta con alimentar al módulo ESP32 con una corriente de 5v. los cuales pueden provenir de una PC o un convertidor de corriente ac a USB (cargador normal de celular).
  - a) Si el módulo se alimenta con una PC y el IDE de "Arduino" se encuentra operando en su interfaz serial, podremos apreciar los cambios de temperatura y humedad desde esta consola.
  - b) Si el sistema se alimenta con el convertidor AC - DV 5v, comenzará a operar de forma normal.
  - c) En el inicio del sistema éste buscará la red de Wifi definida para el alertamiento. Cuando la encuentra, realizará el logueo en dicha red, cuando este paso se encuentre completado, el sistema manda un correo electrónico anunciando el inicio del sistema.
- 2) Monitoreo y control de temperatura por interfaz bluetooth.
  - a) Es necesario bajar una aplicación del tipo "Serial bluetooth" genérica a nuestro celular, con esta aplicación podremos ver la conexión a nuestro sistema: "ESP32\_SO". En dispositivos Android recomendamos la aplicación "Serial Bluetooth Terminal".
  - b) Para realizar la conexión hay que conectarse a este dispositivo por la interfaz bluetooth "pair".
  - c) Una vez establecida la conexión, comenzaremos a visualizar en la pantalla de la aplicación los mensajes de control para la administración del sistema, así como los mensajes de actualización de temperatura y humedad.
- 3) Control de la temperatura en forma automática
  - a) El sistema contempla un modo automático en el cual guarda la primera temperatura sensada por el sistema y la establece como la temperatura base que debe mantener.
  - b) Si el sistema no se encuentra en modo manual, el sistema por default será el sistema automático.

- c) Cuando el sensor detecta un incremento  $\geq 2$  grados centígrados, el sistema de ventilación se activará de forma automática.
  - d) El funcionamiento del sistema de ventilación continuará funcionando mientras la temperatura no baje a la temperatura base.
- 4) El sistema contempla un modo de alertamiento de fallas detectadas en el mismo, por lo cual, si el sensor de temperatura/humedad perdiera conexión con el sistema (se desconectaría, rompiera alguno de los pines de conexión o si este se des calibra) este evento desencadena una subrutina en la cual, se manda un correo electrónico notificando a los clientes definidos en el sistema con el evento detectado.

La rutina del alertamiento a través del correo electrónico depende de que el sistema se encuentre conectado a la red de Wi-Fi definida en el mismo al inicio de este. Esta rutina se encuentra definida en el sistema como una función, por lo cual ésta se puede llamar a voluntad, estableciendo una conexión con el servidor cada vez que un correo electrónico se necesite enviar [7,8].

## 6 Estado del arte

En la literatura no se reportan trabajos de IOT para centros de datos, solo se venden ya estos dispositivos, para utilizarse como usos netamente comerciales, y obviamente se ven como cajas negras no se muestra el desarrollo que se debe de hacer para llegar a tal fin. Además, los precios son elevados. Y lo que se muestra en este trabajo es que mediante una tarjeta Esp-32 y sensores se puede realizar este trabajo, con bajo costo y llega a funcionar perfectamente.

## 7 Resultados

Este proyecto consta del siguiente diagrama ver figura 5

El sistema se encuentra en estado de reposo inicialmente, estableciendo la temperatura inicial que debe mantener constante en todo momento (figura 6).

En la figura 7, se puede apreciar cómo se maneja de forma manual la activación del sistema de ventilación (se manda el 1 desde la aplicación). Cuando se activa el modo manual, el modo automático queda en segundo plano, por lo cual será ignorado hasta la desactivación del modo manual [9,10].

En la figura 8, se puede apreciar cómo se maneja de forma manual el apagado del sistema de ventilación (se manda el 2 desde la aplicación) con lo cual el sistema sabe que desactiva la ventilación y se activa nuevamente el modo automático.

En la figura 9, se puede apreciar cómo se maneja de forma automática la activación del sistema de ventilación cuando la temperatura original del sistema se ve rebasada por más de 2 grados centígrados

En la figura 10, se puede apreciar cómo se maneja de forma automática el apagado del sistema de ventilación, el cual sucede cuando la temperatura regresa a la temperatura base del sistema.

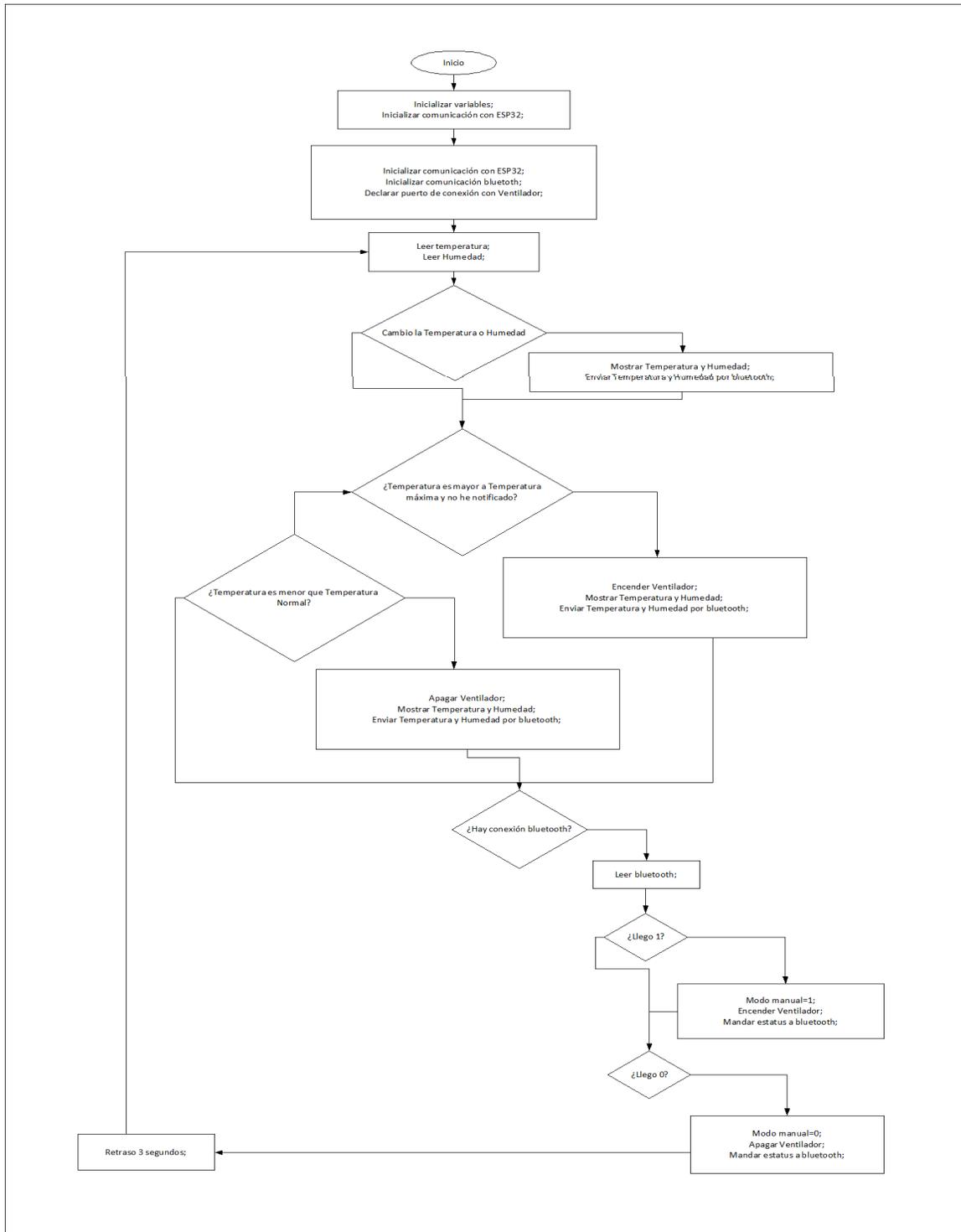


Figura 5 Diagrama con cada uno de los pasos que se siguió

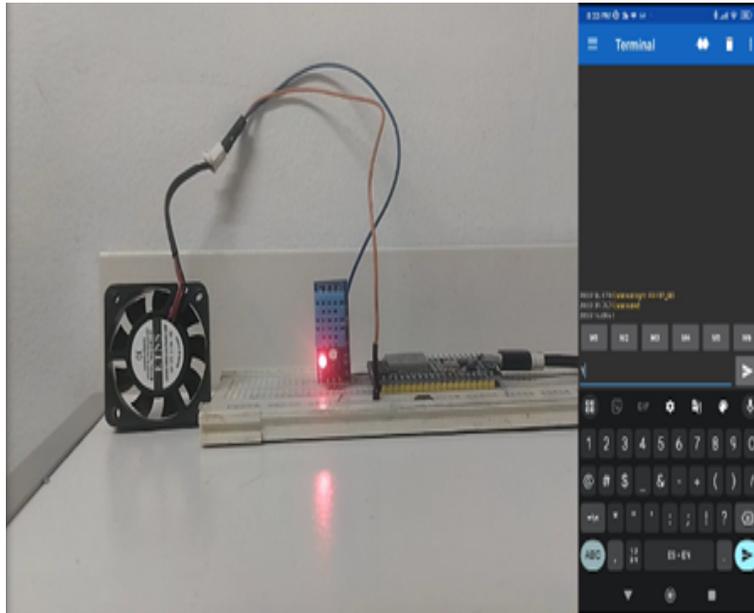


Figura 6. Imagen ilustrativa del sistema en su estado inicial.

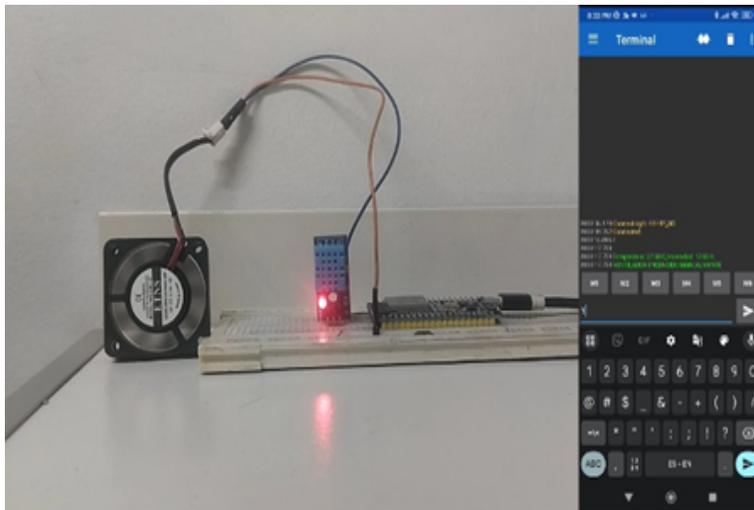


Figura 7. Imagen ilustrativa del sistema ejecutando el comando de activación de ventilación (modo manual).



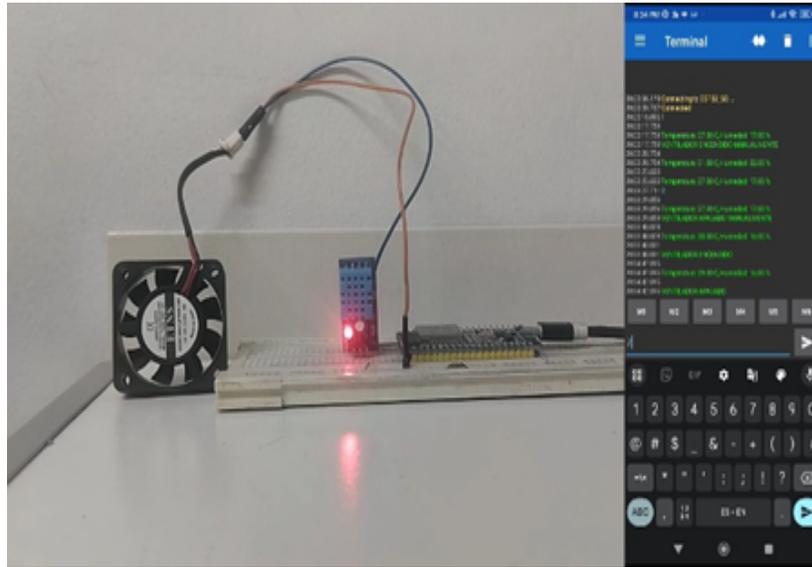


Figura 10. Imagen ilustrativa del sistema ejecutando el apagado de ventilación (modo Automático).

## 8. Conclusiones

Como se pudo apreciar en las imágenes y el video mostrado durante la exposición del proyecto, el sistema funciona de manera correcta tanto en los modos manual como automático. Al inicio del proyecto, surgieron algunos problemas debido a que el módulo ESP32 solo enviaba un pulso de activación y desactivación según fuera necesario, lo que provocaba que el ventilador no estuviera siempre activo. Esto se solucionó integrando un transistor al sistema, el cual proporcionó la alimentación necesaria para mantener el ventilador activo hasta que la temperatura volviera a la ideal definida en nuestro sistema.

Otro problema detectado durante las pruebas fue la falta ocasional de notificación de fallas a través del correo electrónico. Esto se resolvió consultando los requisitos de la librería utilizada para enviar correos electrónicos, que especifica la conexión del sistema a un servidor NTP para obtener el tiempo de internet y así conectar correctamente con el servidor de correo electrónico de Google. Una vez resueltos estos problemas, se logró notificar al cliente de manera oportuna sobre cualquier alerta o cambio en el estado del sistema.

## Referencias

[1] Cherry, N. (2006). Linux Smart Homes For Dummies. For Dummies.

[2] Miller, M. (2015). The internet of things: How Smart TVs, Smart Cars, Smart Homes, and Smart Cities are Changing the World. Pearson Education.

[3] Bárbara Emma Sánchez Rinza, Liosbel Cabrera Hernandez, Mario Rossainz López-(2022) 1ra edición. Editorial Ingeniería y Desarrollo de la Nueva Era. Editorial instituto Antioqueño de Investigación.

[6] Bárbara Emma Sánchez Rinza, Kevin Arturo Ramirez Zavalza, Mario Rossainz López. (2019). Modelado, Tic y Siste,más Distribuidos: avances y aplicaciones. Monitoreo de sensors de humedad y temperatura con Nodemcu. 1ra edición, editorial fondo de cultura.

[7] Barbara Emma Sánchez Rinza, Nallely Morales Lozada, Alberto Jaramillo Núñez. (2020) Electro. Monitoreo de incendios por medio de internet de las cosas, vol42, pag 107-112.

[8] Bonilla-Fabela, Islas et 2016, IOT, el internet de las cosas y la innovación de sus aplicaciones,vincula técnica, junio N 1, issn 2448-5101

[9] David, N., Chima, A., Ugochukwu, A., Obinna, E. (2015). Design of a home automation system using arduino. International Journal of Scien- tific Engineering Research, 6(6), 795-801.

[10] Wildi, T. (2007). Máquinas eléctricas y sistemas de potencia. México, DF: Pearson Educación.

## **Una arquitectura escalable basada en microservicios para la toma de decisiones empresariales**

### **A scalable microservices-based architecture for business decision making**

Hugo A. Chiquito Onofre, Abraham Sánchez López, Alfredo Toriz Palacios  
Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla,  
Av. 14 Sur y San Claudio. Col. San Manuel. C.P. 72570. Puebla, Puebla, México.

hugo.chiquito@outlook.com, abraham.sanchez@correo.buap.mx,  
alfredo.toriz@correo.buap.mx.

#### **Abstract**

Software architecture is a fundamental area within software engineering, which has been constantly evolving to adapt to new trends and needs in software development. This work proposes a scalable architecture based on microservices that provides a scalable environment that allows intensive data processing, which would allow greater efficiency, simplification and fluidity for data analysis. Within the architectural proposal, we apply the fundamental principles that govern architectures: such as scalability, modularity, autonomy and agility. This architecture meets the strategic objectives demanded by today's companies.

#### **Resumen**

La arquitectura de software es un área fundamental dentro de la ingeniería de software, que ha ido evolucionado constantemente para adaptarse a las nuevas tendencias y necesidades del desarrollo de software. Este trabajo propone una arquitectura escalable basada en microservicios que proporcione un ambiente escalable que permita el procesamiento intensivo de datos, lo que permitiría una mayor eficiencia, simplificación y fluidez para el análisis de datos. Dentro de esta propuesta arquitectural, aplicamos los principios fundamentales que rigen las arquitecturas: como son escalabilidad, modularidad, autonomía y agilidad. Dicha arquitectura cumple con los objetivos estratégicos demandados por las empresas actuales.

---

Keywords and phrases: Arquitectura de Software, Arquitectura Orientada a Servicios, Microservicios, Análisis de datos.

---

## **1 Introducción**

Tradicionalmente, el término ‘arquitectura’, se refiere a la disciplina que trata con la planificación y diseño de diferentes construcciones. Similarmente, la arquitectura de software nos guía en el proceso de planificación y diseño del desarrollo de software gracias a sus principios y patrones que gobiernan el desarrollo, y todo esto garantiza una estructura sólida, que a menudo permite eficiencia y adaptabilidad.

En comparación la arquitectura ‘tradicional’ hay diferentes procesos, planos y diseños para cada tipo de estructura, en la arquitectura de software se han propuesto una gran cantidad de arquitecturas y elementos, que se adaptan a situaciones y/o problemas diferentes.

Es por ello que recalcamos la importancia de conocer los diferentes tipos que existen, al igual que sus características, elementos, ventajas y los casos de éxito en su implementación. Al conocer todos estos detalles podemos elegir, adaptar e inclusive crear nuestra propia arquitectura de software [1].

En el entorno actual del desarrollo de sistemas empresariales, sobre todo los que se especializan en la recolección y análisis de información como sistemas financieros o comerciales, un punto crítico a considerar es el manejo eficiente de grandes volúmenes de información. Estos sistemas también se deben adaptar a diferentes niveles de carga y ajustarse gradualmente a las nuevas necesidades del flujo de información empresarial. La adecuada implementación de estos elementos aporta una importante ventaja a las empresas en la toma de decisiones y el cómo hacer frente a nuevos retos y desafíos. Dentro de esta propuesta, se plantea una propuesta de arquitectura de software orientada a resolver los desafíos antes citados y con ello ofrecer una solución altamente flexible, eficiente y escalable para sistemas empresariales de alta demanda.

La propuesta adoptada utiliza principalmente las arquitecturas de microservicios y la arquitectura orientada a eventos, los que no lleva en base a sus características proponer componentes de granularidad fina que separan las responsabilidades y manejan eventos gracias a tecnologías novedosas como Kafka. La propuesta se respalda fuertemente con una serie de elementos adicionales, como patrones de diseño y estilos arquitecturales, que mejoran de forma muy importante el rendimiento de los sistemas.

## **2 Principios de diseño de la propuesta arquitectural**

Más adelante destacaremos elementos importantes de nuestra propuesta, pero primero comentaremos de forma breve, los principios de diseño que soportan a la arquitectura presentada en este trabajo.

### **2.1 Modularidad**

La modularidad permite descomponer el sistema en componentes pequeños, independientes y flexibles, facilitando la gestión y el uso del principio de responsabilidad única en cada microservicio [3].

### **2.2 Escalabilidad**

La arquitectura es capaz de adaptarse a diferentes cargas, mejorando el rendimiento mediante la adición de servidores o instancias para escalar horizontalmente, o incrementando los recursos del servidor para escalar verticalmente [10].

### 2.3 Resiliencia

La resiliencia permite al sistema recuperarse de fallos sin afectar su operación global, gracias al desacoplamiento y la separación de responsabilidades entre los microservicios [4].

### 2.4 Separación de responsabilidades

Este principio organiza los componentes de manera que cada uno esté enfocado en una función específica, optimizando la claridad y gestión del sistema.

### 2.5 Encapsulamiento

El encapsulamiento oculta la lógica interna del sistema, aplicándose en las interfaces para la comunicación entre servicios y usuarios, mejorando la interacción y seguridad del sistema.

## 3 Aplicación de los patrones arquitecturales

### 3.1 Estilos arquitectónicos aplicados

Estilo de microservicios: Este estilo tiene como concepto principal la implementación de microservicios, los microservicios son pequeños módulos autónomos que normalmente suelen delimitarse en base a la división natural de la empresa, por ejemplo, basándose en funcionalidades de algún proceso de negocio [2].

Estilo basado en eventos: Tal como su nombre lo indica, la base de este estilo son los eventos. Para adoptarlo, se requiere al menos un componente productor y un componente consumidor. El componente productor es responsable de generar el flujo de información basado en alguna regla o evento específico, mientras que el componente consumidor recibe esta información y realiza ciertas acciones en función de ella.

### 3.2 Patrones de arquitectura aplicados

Patrón de puerta de enlace (Gateway): El Gateway actúa como intermediario entre los clientes y los microservicios, mejorando la interacción y manteniendo el principio de encapsulación al ocultar los detalles internos y exponer solo los puntos de acceso necesarios [7].

Patrón de disyuntor (Circuit Breaker): Este patrón mejora la resiliencia al detectar fallos en los módulos y suspender las peticiones hacia ellos hasta que se restablezcan, evitando ciclos de errores que afecten al sistema [6].

CQRS: El patrón CQRS separa las responsabilidades de comandos y consultas, optimizando el rendimiento y la escalabilidad. Esto permite que cada operación escale de forma independiente, evitando los problemas que surgen en sistemas donde ambas acciones están combinadas [5].

### 3.3 Patrones de diseño

Strategy: La implementación de este patrón trae consigo ventajas como facilitar el mantenimiento o la implementación de nuevas estrategias y en nuestro caso específico, eventos [11].

### 3.4 Tecnologías utilizadas

Java Spring Boot Framework: Este framework es popular en el desarrollo de microservicios por su inyección de dependencias, facilitando la modularidad. Ofrece soporte para repositorios de datos, integración con la nube, seguridad, gestión de recursos y manejo de excepciones [9].

MySQL: MySQL es el sistema de gestión de bases de datos seleccionado. Aunque no se probaron otras opciones, se recomienda elegir la base de datos que mejor se adecúe al proyecto. En arquitecturas de microservicios, cada servicio puede utilizar una base de datos diferente.

Kafka: Apache Kafka se utiliza para la transmisión de datos entre microservicios, permitiendo la publicación, procesamiento y gestión eficiente de eventos. Además, se integra bien con contenedores y Kubernetes, lo que lo hace ideal para arquitecturas distribuidas [8].

## 4 Comparación con arquitecturas existentes

La arquitectura basada en microservicios propuesta en este trabajo ha sido diseñada no solo para superar las limitaciones de las arquitecturas tradicionales, sino también para aprovechar las mejores características de enfoques arquitectónicos bien establecidos, como la modularidad de los microservicios, la eficiencia de la arquitectura orientada a eventos y la optimización del rendimiento con patrones como CQRS. Además, ha sido pensada específicamente para entornos altamente escalables como la nube, donde se pueden gestionar y desplegar microservicios con diferentes tecnologías según las necesidades del sistema.

En comparación con las arquitecturas monolíticas tradicionales, que presentan dificultades para escalar y mantener a largo plazo debido a su fuerte interdependencia entre componentes, la arquitectura basada en microservicios permite una mayor flexibilidad y escalabilidad. Esto es gracias a la modularidad, que permite que cada servicio se desarrolle, despliegue y escale de manera independiente. Además, la arquitectura orientada a eventos implementada en esta propuesta se basa en el uso de tecnologías como Apache Kafka, lo que permite una comunicación eficiente y asíncrona entre los componentes. Este enfoque se alinea con los principios de los sistemas distribuidos modernos, que optimizan la capacidad de resiliencia y adaptación ante fallos.

Además, se ha integrado el patrón CQRS (Command Query Responsibility Segregation) para separar las operaciones de lectura y escritura, lo que permite optimizar el rendimiento en sistemas con

diferentes niveles de carga. Este enfoque se vuelve aún más eficaz en un entorno en la nube, donde cada operación puede escalarse de manera independiente, utilizando tecnologías específicas según sea necesario. Por ejemplo, un microservicio de lectura podría aprovechar una base de datos NoSQL optimizada para consultas rápidas, mientras que un microservicio de escritura podría utilizar una base de datos relacional que garantice la consistencia de las transacciones.

En resumen, esta arquitectura ha sido diseñada para sacar el máximo provecho de las características clave de diferentes enfoques arquitectónicos, combinando la flexibilidad y escalabilidad de los microservicios, la eficiencia de la arquitectura orientada a eventos, y la optimización del rendimiento mediante patrones como CQRS. Esta combinación no solo es ideal para entornos en la nube, sino que también permite la integración de tecnologías heterogéneas, adaptándose a los diferentes requisitos de los sistemas empresariales modernos.

### 5. Elementos destacables de la arquitectura propuesta

La arquitectura que se propone se constituye de 4 componentes esenciales que son (ver la figura 1):

- Microservicio CUDS
- Microservicio de visualización
- Microservicio CRUD reactivo
- Microservicio de análisis de datos

Cada uno de estos componentes se implementaron combinando los estilos de arquitectura, los patrones de arquitectura y los patrones de diseño antes descritos.

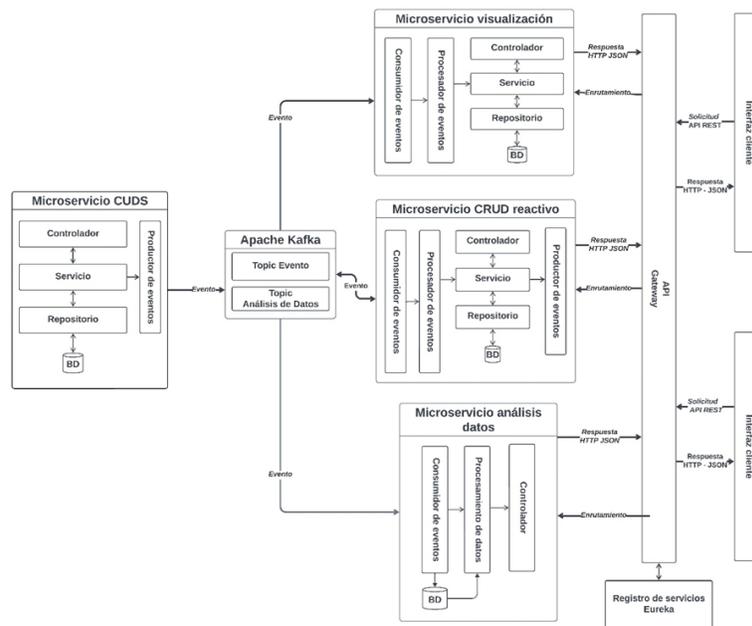


Figura 1. Diagrama general de la arquitectura.

Cada elemento tiene un enfoque y responsabilidades diferentes, estos comparten elementos clave importantes que se deberían detallar.

En primer lugar, los componentes de microservicios desarrollados comparten una estructura similar, todos utilizan 3 entidades o módulos para el correcto manejo de peticiones, es decir:

- Controlador: Encargado de recibir las peticiones REST.
- Servicio: Encargado de procesar la información (lógica de negocio).
- Repositorio: Encargado de persistir y gestionar el almacenamiento de datos.

Una vez clarificado esto, explicaremos con mayor detalle cada uno de los componentes propuestos, enfatizando en sus responsabilidades y características únicas.

Microservicio C.U.D.S.: El acrónimo C.U.D.S. fue creado específicamente para esta propuesta, y derivado del conocido CRUD (Create, Read, Update, Delete), pero con la diferencia de que reemplazamos la 'R' (Read) por una 'S' (Share). Según estas siglas, el componente C.U.D.S. se encarga de crear, actualizar, eliminar y compartir entidades del sistema.

La estructura de este componente se compone de 4 módulos, el flujo de trabajo que sigue es el siguiente:

1. El controlador recibe las peticiones REST validando los datos recibidos.
2. Los datos son procesados dentro del módulo de servicios
3. Se invoca al módulo de repositorio para persistir la operación.
4. Se verifica si la operación se efectuó y envió al módulo de eventos.
5. El módulo de eventos creará y publicará un evento con la información proporcionada.

Microservicio de visualización: La función principal de este componente es la manipulación de las operaciones de lectura, de hecho tiene una fuerte relación con los demás componentes dado que este microservicio es el responsable de consumir los eventos publicados por el resto de los microservicios.

De manera similar este componente cuenta con 4 módulos o capas y su flujo de trabajo es el siguiente:

1. El módulo consumidor de eventos recibe un evento.
2. La información es enviada y procesada en el módulo de procesamiento.
3. Dentro del módulo de procesamiento, se invocan los servicios necesarios para realizar la operación.
4. Se almacena el resultado haciendo uso del módulo de persistencia.

Microservicio CRUD reactivo: Las responsabilidades principales se refieren a la manipulación de los datos y a la capacidad de consumir y publicar eventos. En situaciones donde no se vislumbra una gran diferencia entre el número de solicitudes para las peticiones de lectura y actualización de datos,

este elemento es la mejor elección. Por lo tanto, este componente puede generar eventos y publicarlos, así como recibir información no solo de los controladores, sino que también de otros componentes.

**Microservicio análisis de datos:** Este componente tiene como tarea principal recopilar, procesar, almacenar y generar información a través del análisis de datos. Para realizar esto, se necesita consumir datos de los eventos de interés, y posteriormente procesarlos y almacenarlos en los repositorios adecuados. Para este microservicio se propone utilizar Python, debido a la gran popularidad y herramientas que tiene este lenguaje en el manejo de los datos.

**Conectores:** Para la comunicación de los componentes, proponemos utilizar alternativamente la comunicación asíncrona y síncrona haciendo uso del intercambio de peticiones a través de peticiones HTTP REST y Kafka con el manejo de eventos.

**Interfaces:** En lo referente a la interfaz, se integra un API Gateway en conjunto con APIs REST en cada microservicio implementado. Este enfoque permite una interacción más ad hoc entre los usuarios y el sistema, así como entre los componentes internos dado que el punto de entrada está unificado, estandarizado y centralizado.

**Almacenamiento de datos:** En nuestra propuesta modular basada en microservicios, cada microservicio debe tener su propia base de datos. La elección entre bases de datos relacionales y no relacionales depende de las operaciones que realizará cada microservicio. Para almacenar archivos grandes, se recomienda utilizar sistemas de archivos externos.

**Middleware:** Kafka Broker: encargado de distribuir y gestionar los eventos hacia y entre los microservicios. Spring Gateway: encargado de la gestión del enrutamiento hacia las instancias de los microservicios. Eureka Server: se encarga del seguimiento y registro de las instancias disponibles, y los microservicios.

## **6 Implementación y resultados**

Con el fin de validar la propuesta arquitectónica, se implementó un sistema que simula funciones esenciales de un entorno empresarial moderno. La arquitectura está compuesta por cuatro microservicios, cada uno alineado con los componentes propuestos, los cuales están diseñados para manejar diferentes aspectos del flujo de información. Este enfoque permite observar la concurrencia de datos, así como el procesamiento de eventos en un contexto de alta demanda y escalabilidad.

El microservicio de creación de productos (CUDS) actúa como el productor inicial dentro del flujo de datos, generando eventos cada vez que se crea un nuevo producto en el sistema. Estos eventos son consumidos por el microservicio de ventas (CRUD Reactivo), que utiliza la información de los productos para realizar transacciones de venta. Además de consumir eventos, el microservicio de ventas genera sus propios eventos relacionados con las transacciones, que posteriormente son

procesados por otros microservicios. La integración entre los microservicios garantiza que los datos fluyan de manera eficiente y sincrónica, cumpliendo con los principios de modularidad y autonomía característicos de los sistemas basados en microservicios.

Instances currently registered with Eureka			
Application	AMIs	Availability Zones	Status
MICROSERVICE-CATALOG	n/a (1)	(1)	UP (1) - <a href="#">microservice-catalog:7cd06db6495fae9262721f39aa528991</a>
MICROSERVICE-GATEWAY	n/a (1)	(1)	UP (1) - <a href="#">ddc32657ae7a:microservice-gateway:8090</a>
MICROSERVICE-PRODUCTS	n/a (1)	(1)	UP (1) - <a href="#">microservice-products:a2861ea560460128d3b496f0794ff431</a>
MICROSERVICE-SALES	n/a (1)	(1)	UP (1) - <a href="#">microservice-sales:1d3540d925b57289447b8a20a18a2e38</a>
SALES-DATA-MICROSERVICE	n/a (1)	(1)	UP (1) - <a href="#">192.168.65.254:sales-data-microservice:5000</a>

Figura 2. Microservicios en ejecución reconocidos por Eureka.

Para evaluar el rendimiento y la capacidad de estos microservicios, se llevaron a cabo pruebas de carga utilizando JMeter. Las pruebas se centraron en medir la capacidad de respuesta y la estabilidad del sistema bajo una carga simulada de usuarios concurrentes. La configuración de las pruebas incluyó 1000 usuarios, con un Ramp-Up Period de 10 segundos y un Loop Count de 1 iteración, lo que permitió simular escenarios realistas en un entorno controlado.

El propósito de las pruebas fue identificar la capacidad óptima del sistema, es decir, el punto en el que las latencias comenzaban a incrementarse significativamente. A continuación, se presentan los resultados clave para cada microservicio evaluado, detallando el comportamiento del sistema en términos de latencia, error, y desviación estándar, lo que permitió identificar su rendimiento bajo condiciones de carga:

- Microservicio de productos: Actuando como el punto de entrada de eventos, el rendimiento de este componente fue clave para garantizar la fluidez de todo el sistema.
- Microservicio de catálogo: Encargado de la visualización de productos, este microservicio demostró tiempos de respuesta rápidos y consistentes bajo la carga simulada.
- Microservicio de ventas: Este componente, que consume y genera eventos de transacciones, fue sometido a pruebas rigurosas para asegurar su capacidad de manejar un alto volumen de transacciones simultáneas sin afectar el rendimiento global del sistema.

Los resultados reflejan la capacidad de los microservicios para manejar cargas moderadas de usuarios sin afectar de manera significativa la latencia ni la estabilidad del sistema. Sin embargo, debido a limitaciones de hardware, no fue posible realizar pruebas de escalabilidad a gran escala. A pesar de ello, los resultados obtenidos sugieren que el sistema tiene un alto potencial de

escalabilidad, lo que podría ser confirmado en un entorno de nube con mayor capacidad, como AWS.

Tabla 1. Tabla de resultados de pruebas

Microservicio	Usuarios	Tiempo promedio	Tiempo máximo	Error %	Desviación estándar
Productos (creación de un producto)	1000	10	37	0	2.84
Catalogo (visualizar producto)	1000	3	7	0	0.65
Ventas (crear una compra)	1000	11	34	0	2.63

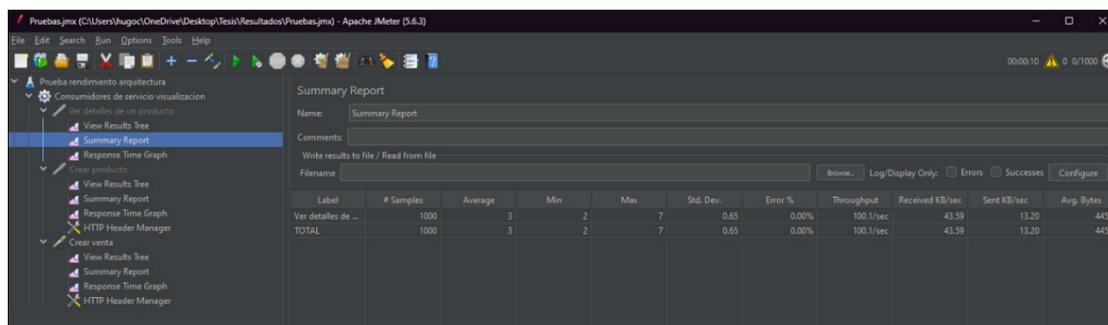


Figura 3. Ejemplo de resultado de pruebas.

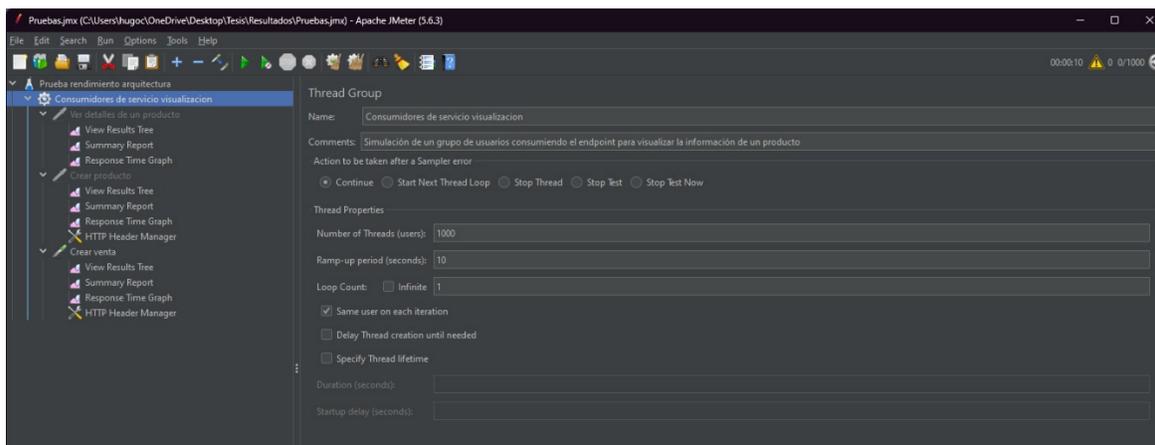


Figura 4. Configuración del grupo de clientes.

El microservicio de análisis de datos desempeña un papel clave en consolidar la información generada por los microservicios de productos y ventas, procesando los eventos recibidos para generar reportes útiles. Aunque no es el único responsable de garantizar la propagación correcta de los eventos, su función consiste en consumir los datos ya procesados por otros microservicios y

ofrecer análisis detallados, confirmando que las transacciones de ventas se han registrado y procesado adecuadamente.

## **7 Conclusiones y trabajo a futuro**

La arquitectura propuesta está diseñada para ofrecer una flexibilidad robusta, escalabilidad eficiente y una comunicación ágil entre microservicios. La implementación valida cómo los conceptos, métodos y tecnologías se integran eficazmente, cumpliendo con los principios clave de la propuesta. Esto garantiza que los microservicios faciliten el flujo de información de manera eficiente entre los componentes del sistema, optimizando la respuesta a las demandas empresariales actuales.

Si bien las pruebas de rendimiento realizadas con JMeter bajo un entorno local ofrecen una evaluación preliminar del comportamiento del sistema bajo condiciones controladas, es importante señalar que la escalabilidad de la arquitectura aún no ha sido validada en entornos de nube, donde se podrían aprovechar completamente las capacidades de auto escalado y la distribución eficiente de recursos. Debido a limitaciones de hardware, no se realizaron pruebas con una mayor cantidad de usuarios simultáneos, lo cual será abordado en futuros trabajos. Se planea realizar pruebas en entornos como AWS, lo que permitirá validar de manera exhaustiva el rendimiento y la estabilidad del sistema bajo condiciones de carga más exigentes, asegurando que la arquitectura pueda escalar de manera efectiva a miles de usuarios concurrentes sin comprometer la eficiencia del procesamiento de datos.

## **Referencias**

[1] Bass, Len, Paul Clements, and Rick Kazman. *Software Architecture in Practice*. Addison-Wesley, 2012.

[2] Newman, Sam. *Building Microservices: Designing Fine-Grained Systems*. O'Reilly Media, 2015.

[3] IBM Documentation. (2023a, marzo 24). Recuperado el 15 de enero de 2024, de [IBM.com](https://www.ibm.com/docs/es/aix/7.3?topic=processes-modularity). <https://www.ibm.com/docs/es/aix/7.3?topic=processes-modularity>.

[4] IBM Documentation. (2023b, abril 11). Recuperado el 15 de enero de 2024 [IBM.com](https://www.ibm.com/docs/es/i/7.5?topic=availability-application-resilience). <https://www.ibm.com/docs/es/i/7.5?topic=availability-application-resilience>.

[5] Young, Greg. *CQRS Documents*. 2010.

[6] Nygard, Michael. *Release It!: Design and Deploy Production-Ready Software*. Pragmatic Bookshelf, 2018.

[7] Patrón Gateway Routing. (s/f). Microsoft.com. Recuperado el 29 de febrero de 2024, de <https://learn.microsoft.com/es-es/azure/architecture/patterns/gateway-routing>

[8] Kreps, J., Narkhede, N. and Rao, J. (2011) Kafka: A Distributed Messaging System for Log Processing.

[9] Moiseev, D., Carnell, J. Learn Microservices with Spring Boot: A Practical Approach to RESTful Services using an Event-Driven Architecture, Cloud-Native Patterns, and Containerization. Apress, 2021.

[10] Coulouris, George, Jean Dollimore, y Tim Kindberg. Distributed Systems: Concepts and Design. Pearson Education, 2005.

[11] Gamma, Erich, Richard Helm, Ralph Johnson, and John Vlissides. Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley, 1994.

## **Análisis de sentimiento en la lengua Cmiique Iitom**

### **Sentiment Analysis for Cmiique Iitom Language**

Jesé Urrutia Romero, Pedro González Zamora. Jesús Horacio Pacheco Ramírez, Víctor Hugo Benítez Baltazar

Universidad de Sonora, División de ingeniería. Edificio 50 segundo piso, Blvd. Luis Encinas y Rosales S/N, Col. Centro, Hermosillo, Sonora, México.

jeseurrutiaro@gmail.com, pedro.gonzalez@unison.mx, jesus.pacheco@unison.mx, victor.benitez@unison.mx

#### **Abstract**

The study explores the creation of a model capable of performing sentiment analysis in Cmiique Iitom, an endangered language spoken by the Comcáac community in northwest Mexico. This document aims to describe the methodology used in data collection and analysis and present the preliminary results of a project that seeks to provide tools to help native speakers in the preservation of their language.

#### **Resumen**

El estudio explora la creación de un modelo capaz de realizar un análisis de sentimiento en cmiique iitom, un idioma en peligro de extinción hablado por la comunidad comcáac en el noroeste de México. Este documento tiene como objetivo describir la metodología empleada en la recolección y análisis de datos, y presentar los resultados preliminares de un proyecto que busca aportar herramientas que ayuden a los hablantes nativos en la preservación de su lengua.

---

Keywords and phrases: Sentiment Analysis, Natural Language Processing. Artificial Intelligence

---

### **1 Introduction**

The Cmiique Iitom, spoken on the coast of the state of Sonora, in northwest Mexico, is the language of the Comcáac, also known as the Seris, the native inhabitants of the region. This community has experienced significant changes in its social and linguistic context, leading to the displacement of Cmiique Iitom by dominant languages such as Mexican Spanish and American English. This displacement phenomenon threatens the vitality of Cmiique Iitom, putting it at risk of extinction [1].

Given this scenario, there is a need to explore sentiment analysis of Cmiique litom to better understand the emotional expressions and feelings of the Comcáac community.

The linguistic landscape across Mexican territory is vast and expansive, with 69 living languages coexisting within the country, 68 of which are indigenous, according to the National Institute for Geography, Statistics and Informatics [2]. Most of these languages are part of linguistic families, with Yuto-Nahua being one of the largest. However, Cmiique litom stands apart as an isolated language [3], meaning it does not share a relationship with any other language systems in the region. This uniqueness is a rare phenomenon, and given its status as an endangered language, it raises concerns about the potential loss of an entire cosmovision from the world.

This study aims to identify and analyze the emotions expressed in Cmiique litom, using natural language processing (NLP) methods [4] adapted to the unique features of this language. The goal is to develop a tool that not only allows for a deeper understanding of the emotional richness of Cmiique litom but also contributes to its preservation and revitalization.

Focusing on sentiment analysis represents an important step in the field of NLP, especially in the context of minority and poorly documented languages. By better understanding how feelings and emotions are expressed in Cmiique litom, we can contribute not only to linguistic preservation but also to strengthening the cultural identity of the Comcáac community. Furthermore, this analysis can provide valuable insights for the development of more inclusive language technologies tailored to the needs of diverse linguistic communities.

With this study, we aim to lay the groundwork for future research in sentiment analysis in indigenous languages, paving the way for a broader understanding of human linguistic and emotional diversity. By integrating Cmiique litom into the realm of NLP, we advocate not only for its preservation but also reaffirm the importance of including all languages in the global digital narrative.

Building upon the foundational insights laid out in the introduction, this paper delineates its structure as follows: The second section unfolds the methodology employed in this investigation, detailing the journey from the initial data collection phase through to the algorithmic approaches devised to train the model. In the ensuing third section, we pivot to an exposition of the preliminary results garnered thus far. The fourth and concluding section of this paper synthesizes our findings into a conclusion, wherein we not only reflect on the implications of our work but also chart a course for future endeavors. This forward-looking perspective encompasses an approximation of forthcoming works, setting the stage for subsequent inquiries into the sentiment analysis of indigenous languages and the broader ambit of natural language processing.

## **2 Methodology**

The methodology of this project is divided into 5 phases, as shown in figure 1:



Figure 1. Project's phases

Phase 1 involves extracting phrases from the Cmiique litom language. For this task, we utilize the only dictionary available [5], from which each phrase is extracted along with its corresponding translation into Spanish. Subsequently, each phrase is categorized into three groups: Positive, Negative, and Neutral. This step involved careful examination and categorization of each phrase by three humans that independently classified the same set of phrases without any knowledge of the others' classifications. This independent classification ensured that personal biases were minimized, and each phrase received an unbiased evaluation. Finally, a consensus agreement was reached: if a phrase was classified identically by all three individuals, it was accepted into the corresponding category. This rigorous process ensured the accuracy and reliability of the classification.

The second phase involves analyzing the data structure to extract its main characteristics. During this step, stop words are identified for removal. To identify the stop words, we created a dataset of 100 phrases extracted from the dictionary and other texts. Each word in these phrases was counted, and the 20 most frequently occurring words were included in the stop words dictionary. This is a solid and common method to discover stop words for a given text or language. As Ferilli suggests, it is possible to automate stop word identification using a small corpus [6]. In our context, we did not automate the process, but we did use frequency counts to discover the stop words for Cmiique litom.

The third phase, known as Tokenization, involves extracting individual words from the phrases. This is conducted to measure the frequency of each word appearing in a single phrase. Table 1 provides an example of the data set that is constructed after tokenizer.

Table 1. Frequency table for individual words

	Ziix	Cooha	Tintica	Haa
Ziix cooha tintica haa yahaaco.	1	1	1	1
Cói ziix chaá oo z itcmá teete, yoque.	1	0	0	0
Ziix zo ntizi, meye.	1	0	0	0

The fourth phase involves training a model to classify the phrase into positive or negative categories. For this purpose, the Naive Bayes algorithm was utilized. This algorithm consists of computing the probability of a word being part of a specific class. Naive Bayes employs the equation shown in Equation (1) [7].

$$P(c|z) = \frac{p(Z|C) \times p(c)}{p(z)} \quad (1)$$

where  $c$  corresponds to the class,  $z$  is the attribute applied, while  $P(c)$  and  $P(x|c)$  are the previous probabilities of the class. The model is tested in the phase 5 and the preliminary results will be shown in next section.

The rationale behind our decision to employ the Naive Bayes algorithm is anchored in the algorithm's robustness and its proven reliability across numerous domains of natural language processing and sentiment analysis. The Naive Bayes algorithm's strength lies in its simplicity and effectiveness, particularly in dealing with classification tasks involving text data. It allows for a computationally efficient approach to predict the sentiment category of phrases. This assertion is supported by the findings of Wongkar & Angdresey [8], and also validated by Surya & Subbulakshmi [9].

While the Naive Bayes algorithm serves as an effective initial approach due to its simplicity and computational efficiency, we recognize that it has certain limitations, particularly when applied to smaller datasets like the one used in this study. As our work progresses and the dataset grows, we intend to explore more advanced algorithms that may yield higher accuracy and handle greater complexity in text classification tasks.

### 3 Preliminary results

One of the first achievements of this project is the creation of a functional stop word dictionary in Cmique litom. We believe this comprehensive list of common but less meaningful words will be essential for filtering out text during analysis. This resource will significantly support various future projects related to natural language processing in Cmique litom.

Upon analyzing the sentences from our corpus and tallying the lexemes, we discovered that pronouns, articles, and postpositions are the most frequently occurring types of words, as shown in table 2. This makes them ideal candidates for our stopwords dictionary. Their prevalence in the corpus aligns with our expectations and is indicative of their suitability for the task at hand, reinforcing their selection as foundational elements for our analysis.

This finding is somewhat anticipated and serves as a positive indicator in the development of our sentiment analysis model. The identification of these common lexemes as the most suitable for the stopwords dictionary not only aids in refining the model but also in streamlining the processing of Cmique litom text.

Following the establishment of a stopwords dictionary based on the analysis of the Cmiique litom corpus, we progressed to the preliminary evaluation of our sentiment analysis model. The construction of this model was informed by the stopwords dictionary, aiming to accurately discern the sentiment expressed in Cmiique litom texts. For this evaluation, we utilized a manually classified dataset to benchmark the model's predictions against human judgment.

Table 2. Candidates for the stopwords dictionary

1.'quih': 193	11.'cop': 38
2.'ma,': 74	12.'ano': 38
3.'aha': 57	13.'hax': 35
4.'hant': 57	14.'pac': 35
5.'quij': 52	15.'iti': 34
6.'ziix': 51	16.'com': 34
7.'he': 42	17.'x,': 32
8.'coi': 41	18.'iha': 29
9.'hac': 39	19.'oo': 26
10.'zo': 38	20.'iiqui': 25

Following the establishment of a stopwords dictionary based on the analysis of the Cmiique litom corpus, we progressed to the preliminary evaluation of our sentiment analysis model. The construction of this model was informed by the stopwords dictionary, aiming to accurately discern the sentiment expressed in Cmiique litom texts. For this evaluation, we utilized a manually classified dataset to benchmark the model's predictions against human judgment.

The dataset consisted of ten text samples from the Cmiique litom corpus, each manually classified as either positive or negative. These classifications served as a ground truth against which the model's performance was measured. The sentiment analysis model then predicted the sentiment of each text sample, assigning both a positive and a negative score to reflect the intensity of sentiment detected.

The results of this preliminary evaluation are detailed in Table 2. Each row represents a single text sample, detailing its manual classification, the sentiment predicted by the model, and the corresponding positive and negative scores. These scores are indicative of the model's confidence in its prediction, with higher absolute values suggesting a stronger detected sentiment.

This evaluation demonstrates the sentiment analysis model's capacity to align with human judgment in many instances, as evidenced by the congruence between manual classifications and predicted sentiments, achieving 80% effectiveness in the initial phase. However, as shown in Figure 2, when tested on a larger dataset of 100 records, the model's success rate dropped to 56.1%, with an error rate of 43.9%, which is only marginally better than random guessing. This decline in performance is expected given the limitations of working with such a small dataset. While the 80% accuracy from the initial experiment represents a promising start, the performance on the expanded

dataset highlights the urgent need for more comprehensive data to adequately train the model. The results underline that while our early efforts show potential, significantly more data is required to reduce the impact of randomness and improve the model's ability to accurately classify sentiments in the Cmiique litom language.

Table 3. Scores and predictions

N°	Manual Classification	Predicted Sentiment	Positive Score	Negative Score
1	Negative	Negative	-11.87	-11.48
2	Positive	Positive	-0.71	-6.5
3	Negative	Negative	-6.03	-5.65
4	<b>Positive</b>	<b>Negative</b>	-22.84	-17.35
5	Negative	Negative	-12.09	-11.82
6	<b>Negative</b>	<b>Positive</b>	-0.71	-18.96
7	Negative	Negative	-6.95	-0.68
8	Positive	Positive	-8.92	-12.04
9	Positive	Positive	-11.69	-17.87
10	Positive	Positive	-0.71	-6.91

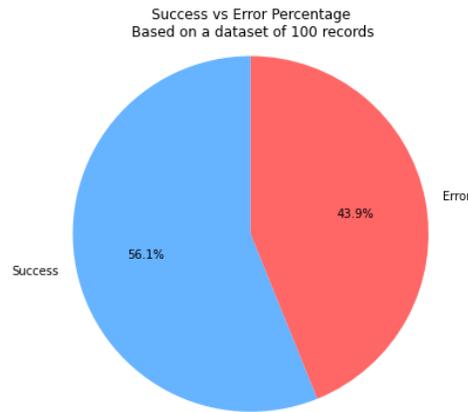


Figure 2. Results with a dataset of 100 records.

#### 4 Conclusions and future works

The sentiment analysis model tailored for the Cmiique litom language has shown potential in its preliminary stages. While the initial experiments yielded promising results with 80% effectiveness in aligning with human sentiment classification, the performance on a larger dataset of 100 records was less satisfactory, with the success rate dropping to 56.1%. This lower accuracy underscores the challenges inherent in working with a limited and sparsely populated dataset. Despite this setback, the results reflect the model's potential, but they also highlight the urgent need for more data to adequately train and refine it.

It is important to emphasize that this project is an ongoing effort, and we are only beginning to uncover its full potential. The current results reflect a work in progress, and we anticipate significant improvements as we continue refining the model and expanding the dataset. While the early

outcomes are encouraging, the work is far from complete, and ongoing iterations will be critical for enhancing both accuracy and robustness.

A critical factor in the model's ongoing development is the need for an expansive and diverse dataset. Enriching the model with more data is essential for its evolution, aiming to make it more robust and adept at handling the complexities of sentiment analysis in an endangered language. This process is not just about quantity but also about the quality and diversity of the data, which will help in minimizing biases and improving the model's comprehension of nuanced sentiments.

Furthermore, collaboration with native speakers of Cmiique litom remains essential. Their involvement ensures that the model benefits from direct insights into the language's sentiment, nuances, and cultural contexts that are difficult to grasp without deep linguistic and cultural knowledge. We are fortunate to have willing collaborators among the Cmiique litom community, who are eager to contribute to this project. Their participation will not only enhance the model's accuracy but also contribute to the preservation and digital representation of their language.

In the future, the focus will be on refining the sentiment analysis model through continuous feedback loops involving native speakers, expanding our dataset, and incorporating these insights into the model's development. This inclusive and collaborative approach promises not only to advance the field of computational linguistics for endangered languages but also to foster a deeper understanding and appreciation of the rich linguistic heritage of the Cmiique litom people. By bridging technology and traditional knowledge, we aim to create a tool that is not only technologically advanced but also culturally informed and respectful, paving the way for future works that respect and uplift the linguistic diversity of our world.

## **Acknowledgements**

We express our sincere gratitude to our university for the support and resources provided throughout the course of this ongoing research. Our heartfelt thanks also go to the native speakers who generously shared their time and insights,

## **References**

[1] J. L. Moctezuma, "Vitalidad y desplazamiento de las lenguas originarias en el norte de México," *Ichan Tecolotl*, vol. 356, Jan. 2022.

[2] Instituto Nacional de Geografía, Estadística e Informática, "Estadísticas a propósito del día internacional de los pueblos indígenas." Aug. 08, 2022. [Online]. Available: [https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2022/EAP\\_PueblosInd22.pdf](https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2022/EAP_PueblosInd22.pdf)

[3] S. A. Marlett, "Las relaciones entre las lenguas 'hokanas' en México: ¿cuál es la evidencia?," in *Clasificación de las lenguas indígenas de México: Memorias del III Coloquio Internacional de*

*Lingüística Mauricio Swadesh*, Primera edición., México, D.F.: Universidad Nacional Autónoma de México, 2007, pp. 165–192.

[4] S. Bandyopadhyay, E. Cambria, D. Das, and A. Feraco, Eds., *A Practical Guide to Sentiment Analysis*, 1st ed. 2017. in *Socio-Affective Computing*, no. 5. Cham: Springer International Publishing : Imprint: Springer, 2017. doi: 10.1007/978-3-319-55394-8.

[5] M. Moser and S. A. Marlett, *Comcáac quih yaza quih hant ihíip hac: Diccionario Seri - Español - Inglés*. Universidad de Sonora, Plaza Valdéz Editores., 2010. doi: 10.1086/521732.

[6] S. Ferilli, “Automatic Multilingual Stopwords Identification from Very Small Corpora,” *Electronics*, vol. 10, no. 17, p. 2169, Sep. 2021, doi: 10.3390/electronics10172169.

[7] Samsir *et al.*, “Naives Bayes Algorithm for Twitter Sentiment Analysis,” *J. Phys. Conf. Ser.*, vol. 1933, no. 1, p. 012019, Jun. 2021, doi: 10.1088/1742-6596/1933/1/012019.

[8] M. Wongkar and A. Angdresey, “Sentiment Analysis Using Naive Bayes Algorithm Of The Data Crawler: Twitter,” in *2019 Fourth International Conference on Informatics and Computing (ICIC)*, Semarang, Indonesia: IEEE, Oct. 2019, pp. 1–5. doi: 10.1109/ICIC47613.2019.8985884.

[9] P. P. Surya and B. Subbulakshmi, “Sentimental Analysis using Naive Bayes Classifier,” in *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)*, Vellore, India: IEEE, Mar. 2019, pp. 1–5. doi: 10.1109/ViTECoN.2019.8899618.

## **Aplicación de realidad virtual como apoyo en la terapia conductual a la acrofobia**

### **Application of virtual reality as support in behavioral therapy for acrophobia**

Ana L. Ballinas Hernández, Daniel Machorro de Jesús, M. Claudia Denicia Carral, Gustavo M. Minquiz Xolo, Héctor Medina Cruz

Complejo Regional Centro, Benemérita Universidad Autónoma de Puebla.  
Calle 2 Sur S/N, Ciudad Modelo. C.P. 75012. San José Chiapa, Puebla, México.

analuisa.ballinas@correo.buap.mx, daniel.machorro@alumno.buap.mx, {claudia.denicia, gustavo.minquiz, hector.medinacruz}@correo.buap.mx.

#### **Abstract**

Virtual reality has been used as a support in exposure therapies for phobias, particularly acrophobia, which is the fear of heights. In this project, a mobile application with virtual reality is developed to support behavioral therapies for acrophobia. The proposed methodology includes the following phases: context analysis, scenario selection, prototype creation, virtual environment modeling, application implementation, and testing. The result obtained is a mobile application that includes equitable breathing exercises and height exposure therapies with virtual tours that immerse users in an immersive experience of heights integrated into three scenarios. The application was run on smartphones mounted on virtual reality glasses and the user experience was evaluated by a sample of 20 users, obtaining positive results from the experience.

#### **Resumen**

La realidad virtual ha servido de apoyo en las terapias de exposición a las fobias, en particular en la acrofobia, que es el miedo a las alturas. En este proyecto se desarrolla una aplicación móvil con realidad virtual como apoyo a las terapias conductuales a la acrofobia. La metodología propuesta incluye las fases de: análisis del contexto, selección de escenarios, creación de prototipo, modelado de entornos virtuales, implementación de la aplicación y realización de pruebas. El resultado obtenido es una aplicación móvil que incluye ejercicios de respiración equitativa y terapias de exposición a las alturas con recorridos virtuales que sumergen a los usuarios en una experiencia inmersiva de las alturas integrados en tres escenarios. La aplicación se ejecutó en teléfonos inteligentes montados en gafas de realidad virtual y la experiencia de uso fue evaluada por una muestra de 20 usuarios obteniendo resultados positivos a la experiencia.

---

Keywords and phrases: Acrofobia, Realidad Virtual, Terapia de Exposición, Técnicas de Relajación.

---

## **1 Introducción**

En los últimos años se ha empleado la realidad virtual (RV) como herramienta de apoyo al tratamiento de fobias. En este contexto, se expone gradualmente a las personas a entornos que detonen ataques de pánico para que puedan habituarse y aprender a manejar el miedo. Al comparar la eficacia entre la terapia de exposición de RV frente a la exposición en vivo se han encontrado resultados positivos para la mayoría de las fobias [1].

El miedo y la ansiedad son las principales características que ocurren cuando las personas sufren trastornos de ansiedad fóbica ante ciertas situaciones. Las principales fobias estudiadas con RV son la agorafobia y la fobia social, muy comunes en las personas [2].

Otro trastorno muy común es la acrofobia que se experimenta como un miedo irracional cuando las personas están en alturas elevadas, tales como, un puente, en lo alto de un edificio o en un balcón. Existen tres diferentes estados de exposición a las alturas: desequilibrio fisiológico resultante de un control visual deficiente, intolerancia angustiante y la acrofobia en el extremo más severo [3]. La acrofobia puede afectar significativamente la vida de quienes la padecen debido a que pueden experimentar varios síntomas como mareos, sudoración excesiva, sensación de vértigo, dolor de cabeza, taquicardia, tensión, nervios y dificultad para respirar. Todo esto puede suceder incluso en situaciones en donde la altura sea mínima, como subir escaleras o estar en un balcón.

Hay varias formas de tratar las fobias, pero muchas no pueden adentrar a los pacientes al miedo ya que solo se usan imágenes o su imaginación [4]. El tratamiento conductual suele implicar una exposición gradual a las alturas, pero no siempre es posible para las personas por cuestiones como, la falta de recursos financieros, el tiempo para salir a esos lugares o porque su temor es mayor, lo que puede retrasar la superación del trastorno.

Otra forma de tratamiento conductual consta del uso de la imaginación, pero este proceso muchas veces no logra sumergir al paciente en una experiencia realista para superar el trastorno. Por eso la realidad virtual puede ofrecer una solución mucho más inmersiva, siendo una herramienta útil en psicología [5].

La psicoterapia cognitivo-conductual para el tratamiento de la acrofobia tiene como objetivo que el paciente entienda que su miedo no es real, sino irracional. La terapia conductual se centra en exponer al paciente a su miedo de manera progresiva, utilizando ya sea la imaginación o utilizando dispositivos de realidad virtual [1].

En este proyecto se propone una aplicación de realidad virtual basada en la terapia conductual que proporcione entornos seguros para que las personas con acrofobia puedan enfrentar su miedo sin arriesgarse. Inicialmente, la aplicación proporciona al usuario ejercicios repetitivos de relajación para reducir el estrés antes de iniciar con la exposición y la terapia no sea contraproducente. La aplicación tiene escenarios virtuales diseñados para simular situaciones realistas que se repiten periódicamente según el avance del usuario. Para las personas que trabajan en edificios de grandes alturas, ya sea en oficinas o en el sector de construcción, es importante atender este trastorno para que su rendimiento laboral sea óptimo.

## **2 Trabajo relacionado**

La acrofobia puede abordarse con diferentes terapias, algunas apoyadas con la tecnología. En el trabajo de Ruiz y Valero (2021) se explora con terapias de exposición utilizando elementos multimedia de forma progresiva, es decir, primero fotos, después vídeos y finalmente un entorno en vivo, los resultados disminuyeron la ansiedad y en la fobia [6]. Una de las claves principales para una terapia exitosa es la forma y el momento en el que el paciente es expuesto a las situaciones que causan la fobia, algunos trabajos se han valido de la realidad virtual para simular un entorno de exposición controlado pero combinándolo con otras terapias, por ejemplo, en [7] se utiliza una Terapia de Aceptación y Compromiso (ACT) apoyada RV a través de un casco de realidad virtual Oculus Rift y controladores táctiles Oculus así como de la plataforma “Safemind”. Janecky et al., utilizaron HoloLens 2 con modelos virtuales y realidad aumentada para mostrar las fobias en el entorno real del paciente [8].

Sin embargo, el uso de la tecnología Oculus Rift es costosa y las aplicaciones comerciales utilizadas para las terapias no son modificables fácilmente, por tal motivo, otros trabajos han optado por utilizar gafas de realidad virtual y aplicaciones móviles. Se ha aplicado la terapia cognitivo-conductual a través de realidad virtual, empleando gafas, los resultados muestran la factibilidad de bajo costo para el tratamiento de la acrofobia [9]. La terapia cognitivo-conductual mediante gafas y aplicaciones móviles ha demostrado tener potencial de desarrollo, se ha documentado como fácil de usar, aceptable y tolerable, por lo que da la pauta para seguir usando estos recursos [10]. Ureña y Álvarez (2023) mencionan que el uso de realidad virtual a través de gafas es más frecuente en las terapias de acrofobia, debido a que su uso y costo son más asequibles para pacientes y terapeutas, además de que pueden ser más fácilmente adaptados a las necesidades de cada paciente [11].

Los diseños utilizados para las terapias de exposición se han desarrollado en diferentes herramientas y recreando distintos escenarios, Muñoz et al., desarrollaron un prototipo con 8 modelos virtuales que van desde puentes hasta montañas y grúas de construcción por medio de Blender y Unity [12]. Otras aplicaciones se valen de la realidad virtual para el tratamiento de fobias y pueden mejorarse para brindar sensaciones más realistas.

En este trabajo la aplicación propuesta también incorpora una terapia de relajación que complementa la terapia conductual y que es incluida en la aplicación, lo que hace que sea más

completa comparada con las aplicaciones antes presentadas que solamente involucran la realidad virtual.

### 3 Desarrollo de la aplicación

En el presente trabajo se desarrolla una aplicación móvil con realidad virtual como apoyo a las terapias conductuales para la acrofobia, principalmente en personas que se exponen a trabajar en las alturas. Para el desarrollo de la aplicación se siguen las fases metodológicas mostradas en la Figura 1.

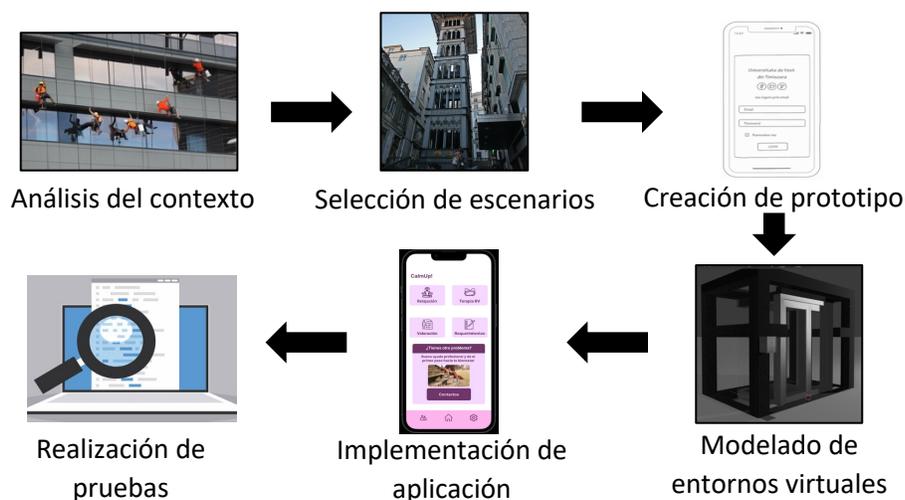


Figura 1. Fases metodológicas para el desarrollo de la aplicación.

#### 3.1 Análisis del contexto

Esta etapa se centra en comprender las distintas terapias de exposición existentes y en la identificación de la terapia más adecuada para construir una aplicación de la realidad virtual que sirva de apoyo para personas con acrofobia. Para ello, se realizó una investigación de las terapias centradas en la modificación cognitiva y aquellas orientadas en una exposición gradual de los estímulos que provocan ansiedad. Además, se analiza cómo los psicoterapeutas realizan las terapias de exposición en escenarios reales donde es crucial realizar una sesión de respiraciones profundas previas a la terapia para que el usuario se encuentre relajado.

#### 3.2 Selección de escenarios

Se propone la creación de tres escenarios virtuales con apariencia realista para sumergir a los usuarios una experiencia inmersiva muy similar a la exposición del entorno real. Los escenarios seleccionados son los siguientes:

- Puente peatonal: la selección de este escenario se debe a su frecuencia de uso, específicamente, ya que son muy usados y con alto tránsito de personas.

- Elevador: aunque existan escaleras para subir varios pisos de un edificio, los elevadores son mucho más usados por su rapidez sin esfuerzos.
- Azotea de un edificio: la cima de un edificio es un entorno clásico asociado con el miedo a las alturas, por lo que este escenario ofrece una vista panorámica que es la más desafiante.

### 3.3 Creación de prototipo

En esta etapa se propone el maquetado y diseño de cada una de las interfaces que componen la aplicación utilizando la herramienta *Figma*. Se hace una selección de colores y tamaños para los iconos, botones y fondo, para garantizar que los usuarios no sufran fatiga visual ni molestias al usar la aplicación. De la misma forma, se propone la creación de los flujos de interacción que contiene la aplicación para lograr un diseño intuitivo y atractivo. La Figura 2 muestra el diseño de la ventana principal de la aplicación móvil para terapias conductuales a la acrofobia.



Figura 2. Prototipo de aplicación móvil.

### 3.4 Modelado de entornos virtuales

En esta etapa se construyen los escenarios virtuales propuestos en la etapa de selección de escenarios (puente peatonal, elevador y azotea). Para ello, se hace uso de la herramienta *Blender* que permite modelar cada escenario añadiendo diversos elementos visuales a su alrededor para lograr que los usuarios tengan una experiencia inmersiva. La Figura 3 muestra un ejemplo del modelado de un puente virtual.

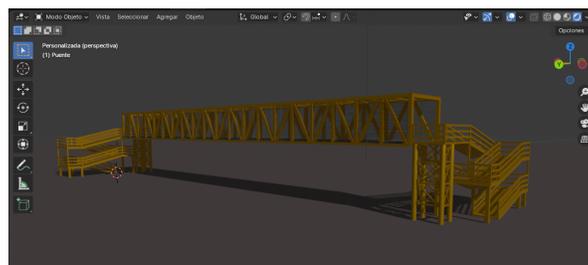


Figura 3. Modelo de un puente virtual.

### 3.5 Implementación de la aplicación

En esta etapa se desarrolla la aplicación móvil con realidad virtual usando el software *Unity 3D* usando el lenguaje de programación *C#* para darle funcionalidad en cada escena. Además, se integra un controlador en primera persona a los recorridos para que el usuario puede interactuar con su entorno. Los recorridos de realidad virtual son integrados a la aplicación para que los usuarios realicen repeticiones de exposición a las alturas. La aplicación funciona solo para dispositivos móviles con sistema operativo Android. El mapa de navegación de la aplicación es mostrado en la Figura 4. La aplicación contiene un menú principal con opciones de: ejercicios de respiración equitativa para la relajación guiada del usuario (con cuatro temáticas diferentes), escenarios de realidad virtual donde el usuario deberá exponerse a cada uno con 10 repeticiones (puente peatonal, elevador y azotea de edificio), valoración del uso de la aplicación donde el usuario contesta un cuestionario sobre su experiencia con la misma y requerimientos técnicos sobre sistema operativo y recursos empleados para su ejecución.

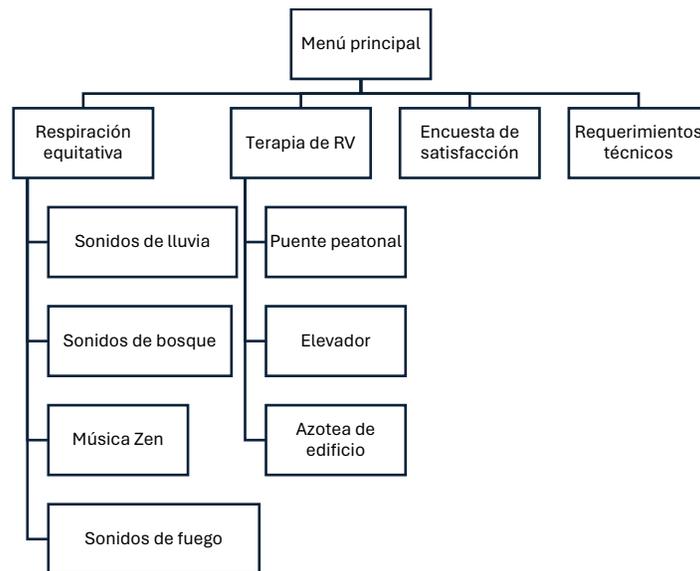


Figura 4. Mapa de navegación de la aplicación.

### 3.6 Realización de pruebas

En esta última etapa se propone la realización de pruebas funcionales con una muestra de usuarios con o sin acrofobia que se exponen a las alturas al asistir a laborar. Para ello, se evalúa que todas las características de la aplicación funcionen según los requerimientos, se buscan y se corrigen posibles errores de navegación y de funcionamiento de los recorridos. Se evalúa la facilidad de uso de la aplicación, desde la navegación hasta los recorridos virtuales. Además, se realizan pruebas de inmersión para evaluar la experiencia de los usuarios al interactuar con los entornos con exposición virtual a las alturas. Para las pruebas de los escenarios virtuales se recomienda usar gafas de RV donde montar un teléfono inteligente con aplicación móvil.

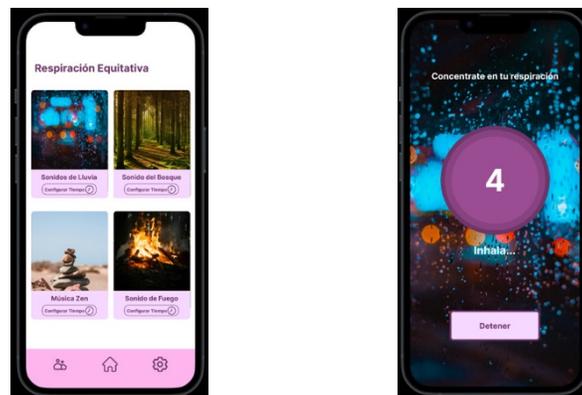
Adicionalmente, se aplicó una encuesta a los usuarios para valorar el nivel de satisfacción de la experiencia de uso de la aplicación y la experiencia con la realidad aumentada. Las preguntas que incluye se listan a continuación:

- a) ¿Cuál es su grado de satisfacción con la experiencia del uso de la aplicación?
- b) ¿Cuál es su grado de satisfacción con la navegación de la aplicación?
- c) ¿Cómo calificaría el aspecto y la sensación de la aplicación móvil?
- d) ¿Cómo calificaría la intuición de los iconos?
- e) ¿En qué grado evaluaría la experiencia inmersiva generada por la realidad virtual?

Es importante mencionar que este análisis es una percepción inicial del uso de la aplicación, sin embargo, a partir de los resultados descubiertos se diseñaran y aplicaran nuevos instrumentos para medir que tan útil puede ser en el tratamiento de la acrofobia.

#### 4 Resultados

Se construyó una aplicación móvil con realidad virtual para la terapia de exposición a la acrofobia en tres escenarios distintos. En la Figura 5.a se puede observar los cuatro ambientes de relajación para realizar ejercicios de respiración profunda equitativa de inhalación y exhalación. Al seleccionar un ambiente, se inicia una relajación con series de 10 respiraciones para liberar la tensión de los usuarios (ver Figura 5.b). Estas escenas incluyen sonidos ambientales alusivos a cada temática y están guiados por voz para lograr resultados más significativos.

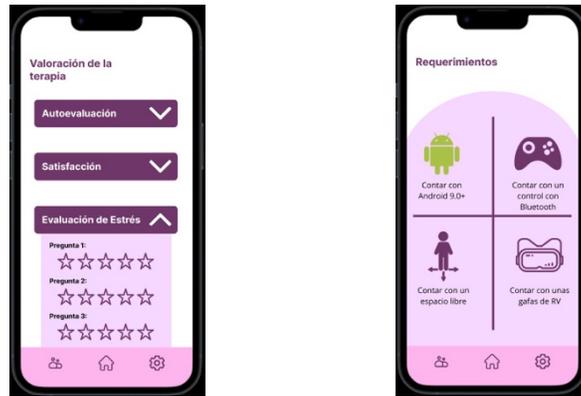


a) Selección de opciones para relajación

b) Repeticiones de respiración

Figura 5. Ejercicios de respiración guiada.

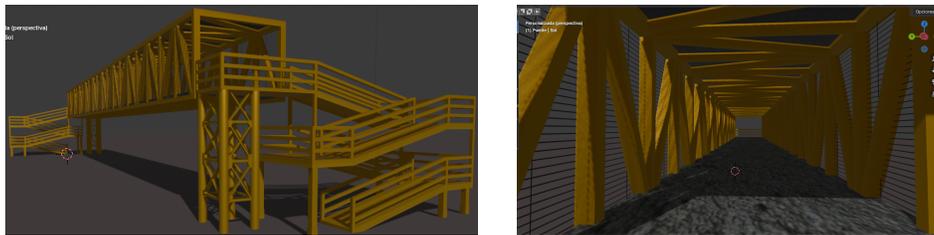
La Figura 6.a muestra las interfaces de la aplicación móvil que incluyen una valoración final de usuario como retroalimentación del uso de la aplicación. La Figura 6.b lista los requerimientos técnicos, de sistema operativo, almacenamiento y recursos, necesarios para que la aplicación pueda funcionar adecuadamente.



a) Encuesta de valoración      b) Requerimientos técnicos

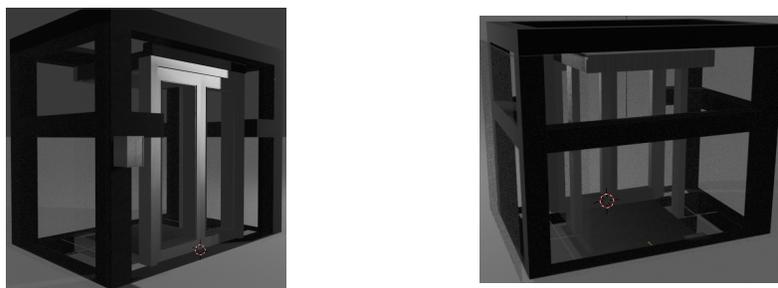
Figura 6. Valoración del usuario sobre el uso de la aplicación.

Las Figuras 7, 8 y 9 muestran los escenarios virtuales seleccionados en este trabajo: puente peatonal, elevador y azotea de edificio, respectivamente. Los escenarios fueron creados aplicando un detallado y efectos realistas que dan al usuario una experiencia inmersiva.



a) Vista desde el inferior      b) Vista desde la parte superior

Figura 7. Recorrido de puente peatonal con realidad virtual.

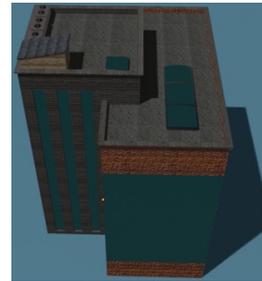


a) Vista lateral      b) Vista frontal

Figura 8. Recorrido de un elevador con realidad virtual.



a) Vista desde la parte superior



b) Vista lateral

Figura 9. Recorrido de la azotea de un edificio con realidad virtual.

La aplicación fue probada con 20 personas, 14 hombres y 6 mujeres, con edades de entre 18 y 22 años, estudiantes del Complejo Regional Centro sede San José Chiapa de la Benemérita Universidad Autónoma de Puebla, 10 de estas personas manifestaron tener algún tipo de temor a las alturas. Al finalizar con la prueba, los usuarios contestaron la encuesta de satisfacción del uso de la aplicación con la siguiente escala: 0- completamente insatisfecha/o, 1- muy insatisfecha/o, 2- algo insatisfecha/o, 3- algo satisfecha/o, 4- bastante satisfecha/o, 5- completamente satisfecha/o. Las preguntas de la encuesta son listadas en la sección 3.6 y el conteo de las respuestas son graficadas en la Figura 10.

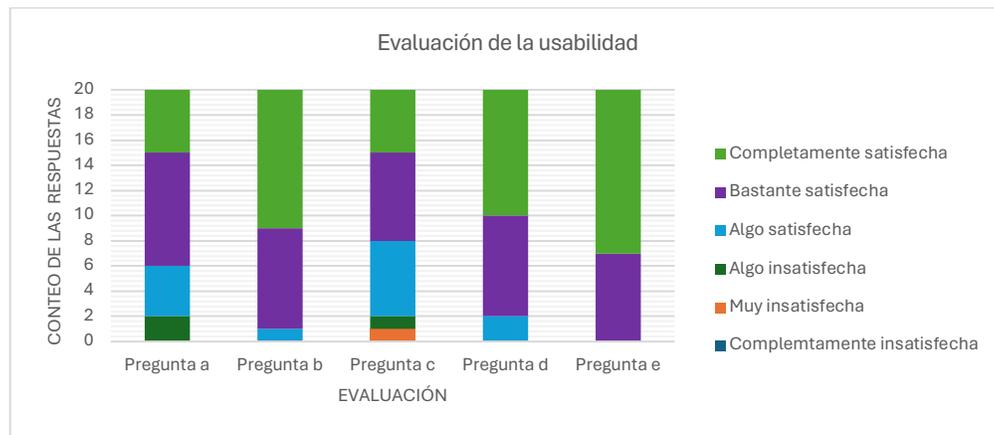


Figura 10. Evaluación de la usabilidad.

Como puede observarse en la Figura 10, ningún usuario contestó que se encuentra completamente insatisfecho con ninguna de las preguntas de la encuesta. Las respuestas muy insatisfecha/o y algo insatisfecha son las de menor frecuencia, mientras que en la mayoría de las respuestas los usuarios contestaron que están bastante satisfechos o completamente satisfechos con el uso de la aplicación de realidad virtual para terapias de exposición a la acrofobia.

## **5 Conclusiones y trabajo a futuro**

En el presente trabajo se propone una metodología para la construcción de una aplicación móvil con realidad virtual que sirva de apoyo en las terapias de exposición a la acrofobia. Se logró construir una aplicación que funciona para teléfonos móviles con sistema operativo Android. La aplicación montada en gafas de realidad virtual se probó con una muestra de 20 personas.

Las pruebas realizadas muestran una alta aceptación de la aplicación en cuanto a diseño, facilidad de uso, técnicas de relajación, experiencia inmersiva de la realidad virtual. Así, la aplicación puede usarse como una herramienta que sirve de apoyo a psicoterapeutas en estos trastornos, lo que reduce costos y evita los riesgos reales a los que pueden enfrentarse las personas.

Como trabajo a futuro, se propone realizar pruebas de medición de estrés para evaluar el estrés de los usuarios antes, durante y después de la sesión de terapia de exposición aplicando cuestionarios de autoevaluación o dispositivos de monitoreo fisiológico, principalmente para personas con acrofobia y con una muestra más grande. Además, podría construirse un juego serio donde el usuario aumente de nivel después de un cierto número de repeticiones a la exposición del escenario según el nivel de avance para mostrar el nivel de mejoría a este trastorno.

## **Agradecimientos**

Los autores agradecen el apoyo otorgado a través del proyecto VIEP con número ID 00322 del ejercicio 2024 de la Benemérita Universidad de Puebla para la realización de este trabajo.

## **Referencias**

- [1] J. R. S. Freitas, V. H. S. Velosa, L. T. N. Abreu, R. L. Jardim, J. A. V. Santos, Peres, B., y P. F. Campos, "Virtual reality exposure treatment in phobias: a systematic review", *Psychiatric Quarterly*, vol. 92, no. 4, pp. 1685-1710, 2021, doi: 10.1007/s11126-021-09935-6.
- [2] M. Krzystanek, S. Surma, M. Stokrocka, M. Romańczyk, J. Przybyło, N. Krzystanek, and M. Borkowski, "Tips for effective implementation of virtual reality exposure therapy in phobias—A systematic review", *Frontiers in psychiatry*, vol. 12, pp. 737351, 2021, doi: 10.3389/fpsyt.2021.737351.
- [3] D. Huppert, M. Wueh, and T. Brandt, "Acrophobia and visual height intolerance: advances in epidemiology and mechanisms", *Journal of neurology*, vol. 267, pp. 231-240, 2020, doi: 10.1007/s00415-020-09805-4.
- [4] A. F. Jiménez Pardo, "Enfoque conductista en el manejo de fobias", Bachelor's thesis, Quito: Universidad Tecnológica Indoamérica, 2022.

- [5] D. Freeman, P. Haselton, J. Freeman, B. Spanlang, S. Kishore, E. Albery, and A. Nickless, "Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial", *The Lancet Psychiatry*, vol. 5, no. 8, pp. 625-632, 2018, doi: 10.1016/S2215-0366(18)30226-8.
- [6] A. Ruiz-García and L. Valero-Aguayo, "Progressive Multimedia Exposure for Specific Phobias: A Clinical Case of Claustrophobia and Acrophobia," *Clin. Case Stud.*, vol. 20, no. 5, pp. 402–416, 2021, doi: 10.1177/15346501211007208.
- [7] Z. Celik, F. B. Alptekin, and K. F. Yavuz, "Acrophobia treatment with virtual reality assisted acceptance and commitment therapy: Case reports," *Dusunen Adam - J. Psychiatry Neurol. Sci.*, vol. 33, no. 3, pp. 317–324, 2020, doi: 10.14744/DAJPNS.2020.00097.
- [8] D. Janecky, E. Kucera, and O. Haffner, "HoloLens 2 and Virtual Reality as Methods for Managing Phobias," *Int. Conf. Cybern. Informatics (K&I)*, pp. 1-7, 2022, doi: 10.1109/KI55792.2022.9925954.
- [9] T. Donker, I. Cornelisz, C. Van Klaveren, A. Van Straten, P. Carlbring, P. Cuijpers, and J. L. Van Gelder, "Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial", *JAMA psychiatry*, vol. 1, no. 76, pp. 682–690, 2019, doi: 10.1001/jamapsychiatry.2019.0219.
- [10] T. Donker, and M. Heinrichs. "Acrophobia and consumer-based automated virtual reality cognitive behavior therapy. *Handbook of Cognitive Behavioral Therapy by Disorder*", pp. 53-64, 2023, doi: 10.1016/B978-0-323-85726-0.00035-1.
- [11] F. Isabel y Á. Urueña, "La realidad virtual para la acrofobia: Evaluación de la usabilidad de una VRBox comercial," vol. 14, no. 1, pp. 62–73, 2023, doi: 10.20318/recs.2023.7329.
- [12] D. S. Muñoz Moncayo, J. A. Santa Muñoz and D. Mayorca-Torres, "Prototipo de sistema de apoyo a la terapia de exposición controlada enfocada en acrofobia mediante realidad virtual y con captura de signos vitales," *Universidad Mariana Boletín Informativo*, vol. 7, no. 3, 2020, pp. 142–146.

**Recuperación de la información efectiva y selectiva a través de  
inteligencia artificial: una revisión sistemática**

**Selective and effective information retrieval powered with artificial  
intelligence: a systematic review**

Daniel Arellano García, Diana Margarita Córdova Esparza

Facultad de Informática, Universidad Autónoma de Querétaro  
Av. de las Ciencias S/N, Juriquilla, Querétaro 76230, México.

daniel.arellano@uaq.mx, diana.cordova@uaq.mx

**Abstract**

The present work aims to empirically substantiate that Information Retrieval Systems with Artificial Intelligence can adapt and improve over time as they learn from new interactions and data. By conducting extensive and up-to-date searches in recent studies, along with appropriate keywords and proposing specific research questions, we will attempt to meet the objective of this research. This search was conducted in search engines such as Google Scholar, ProQuest, and ScienceDirect, yielding a total of 10 research papers, allowing us to discover that the most appropriately used methods suggested for adequately developing the integration of Artificial Intelligence in Information Retrieval are Natural Language Processing techniques, combined with Deep Learning training techniques, integrated with Bayesian Networks and Cognitive Diagnostics.

**Resumen**

El presente trabajo tiene por objetivo fundamentar de manera empírica cómo los Sistemas de Recuperación de la Información con Inteligencia Artificial pueden adaptarse y mejorar con el tiempo a medida que aprenden de nuevas interacciones y datos. Para llevar a cabo este estudio, se hizo una exploración de la literatura a través de los motores de búsqueda: Google Scholar, ProQuest y ScienceDirect, seleccionando un total de 10 artículos de investigación. De acuerdo con el análisis realizado, se puede concluir que los métodos más usados para desarrollar adecuadamente la integración de la Inteligencia Artificial en la Recuperación de la Información son las técnicas de Procesamiento del Lenguaje Natural, aunado a métodos de entrenamiento de Aprendizaje Profundo, el uso de Redes Bayesianas y Diagnósticos Cognitivos.

---

Keywords and phrases: Sistema de Recuperación de Información, Inteligencia Artificial, Aprendizaje Automático, Aprendizaje Profundo, Procesamiento del Lenguaje Natural, Redes Bayesianas.

---

## **1 Introducción**

La recuperación de la información (RI) es una disciplina que se ocupa de la representación, almacenamiento, organización y acceso a elementos de información. Su objetivo principal es obtener información que sea útil o relevante para el usuario [1]. La RI considera aspectos claves como: a) Relevancia, la cual se evalúa en relación con las necesidades de información del usuario, y no solo a la información coincidente con patrones léxicos en una consulta. b) Métricas de evaluación, los criterios de evaluación para los sistemas de RI comprenden la precisión (documentos relevantes recuperados) y la precisión (la proporción de documentos relevantes recuperados) y el recall (la proporción de documentos relevantes que fueron recuperados de todos los documentos relevantes disponibles). c) Aplicaciones, las técnicas de IR se utilizan ampliamente en motores de búsqueda, sistemas de filtrado de información, resumen de documentos y más, lo que la convierte en un campo crucial para gestionar y acceder a la información en diversos contextos [2].

La RI desempeña un papel significativo en la vida diaria de las personas y organizaciones en general debido a su integración en varias funciones útiles, como: la navegación por Internet, sistemas de preguntas y respuestas, asistentes personales, autómatas de interacción y bibliotecas digitales, entre otras.

La importancia que tienen los datos y la información proveniente de diversas fuentes (desde nivel personal a nivel empresarial), su pérdida o destrucción puede ocasionar diferentes problemáticas y en consecuencia tener percances irreparables que incluso pueden impactar de manera financiera a una organización. El Aprendizaje Automático (ML, por sus siglas en inglés) ha sido determinante en diversos aspectos de los sistemas modernos de RI como, por ejemplo: la personalización de resultados de búsqueda [3], mejorar la comprensión y generación de lenguaje natural en interfaces conversacionales [4] y escalabilidad y eficiencia en Big Data [5]. Adicionalmente, se suman las técnicas de entrenamiento y preentrenamiento de la IA para propósitos más específicos, como los algoritmos de clasificación y las redes neuronales para reconocer patrones y características de datos que han sido verificados previamente [6] y los modelos de lenguaje preentrenados, como BERT y GPT, que se utilizan en RI para entender el contexto y la semántica de las consultas y los documentos [7][8].

El objetivo principal de la Inteligencia Artificial (IA) implementada en los Sistemas de Recuperación de Información (SRI) es reconocer y recuperar información que esté asociada con la solicitud del usuario. Dado que pueden ser relevantes múltiples registros, los resultados se clasifican con frecuencia según su puntuación de similitud con la consulta del usuario. De acuerdo con investigaciones previas en el campo de RI, los sistemas tradicionales de recuperación de texto se basaban predominantemente en la coincidencia de términos entre documentos y consultas. Sin embargo, estos SRI basados en términos tienen inconvenientes, como: polisemia, sinonimia y brechas léxicas, que pueden limitar su efectividad. Recientemente, el campo del Procesamiento del Lenguaje Natural (PLN, Natural Language Processing) ha experimentado avances significativos

debido a la mayor disponibilidad de grandes conjuntos de datos etiquetados y al aumento de la capacidad de procesamiento [9], lo que ha permitido a los investigadores emplear métodos de aprendizaje profundo para diversos propósitos, como por ejemplo: el desarrollo de modelos neuronales para la RI [10][11], la mejora de la capacidad de los SRI para comprender y procesar lenguaje natural [12], la aplicación de redes convolucionales para detectar información previamente verificada [9], o el procesamiento y captura de dependencias a largo plazo en datos textuales y generar de respuestas en IA conversacional [6].

Este aprendizaje continuo permite que el sistema sea más efectivo y preciso a lo largo del tiempo, respondiendo mejor a las necesidades cambiantes de los usuarios. La IA puede utilizarse para predecir las necesidades de información de los usuarios antes de que realicen una búsqueda. A través del análisis de datos y patrones de comportamiento, los SRI pueden anticipar y sugerir información relevante de manera proactiva. Las ventajas que se infieren son relevantes debido a que permiten mejorar la personalización de las búsquedas, la comprensión del lenguaje natural, la automatización, la adaptación continua y el análisis predictivo. Estos desarrollos [6][13] no sólo optimizan la experiencia del usuario, también potencian la capacidad de los sistemas para soportar decisiones críticas y mejorar la eficiencia en la recuperación de información.

Las nuevas tecnologías basadas en IA han brindado soporte para tratar los problemas de RI. Por consiguiente, en este estudio sistemático se examinan los métodos neuronales actuales y probados en este campo de estudio, además abordar sus principales beneficios y el análisis de propuestas innovadoras que tienen como finalidad recuperar la información de una manera selectiva a través del uso de técnicas de IA.

## **2 Marco teórico**

En esta sección se describen los fundamentos teóricos que sustentan la comprensión de este artículo:

### **2.1 Aprendizaje automático (*machine learning*)**

Es la rama de la IA que se enfoca en el desarrollo de algoritmos y modelos que permiten a las computadoras aprender y tomar decisiones basadas en datos. En lugar de ser programadas de manera explícita para realizar una tarea específica, las máquinas utilizan datos de entrada para identificar patrones, hacer predicciones y mejorar su rendimiento con el tiempo a medida que reciben más datos [14][15].

### **2.2 Aprendizaje profundo (*deep learning*)**

Es un subconjunto del ML que se basa en el uso de redes neuronales artificiales con múltiples capas (de ahí el término "profundo") para modelar y resolver problemas complejos. Estas redes neuronales profundas están inspiradas en la estructura y el funcionamiento del cerebro humano, y están diseñadas para reconocer patrones y representar datos de manera jerárquica [7].

### 2.3 Procesamiento del lenguaje natural (PLN)

Es la rama de la IA que se enfoca en la interacción entre las computadoras y el lenguaje humano. Su objetivo es permitir a las computadoras entender, interpretar, y generar lenguaje humano de manera que sea útil [16].

Con el uso de algoritmos avanzados de aprendizaje profundo, el rendimiento de los sistemas de RI ha mejorado considerablemente, lo que lleva a una recuperación de información más precisa y eficiente para los usuarios finales [13]. En relación con las técnicas de aprendizaje profundo empleadas en RI, se han empleado arquitecturas de redes neuronales como:

### 2.4 Redes neuronales convolucionales (CNN)

Utilizadas principalmente para el procesamiento de imágenes, estas redes aplican operaciones de convolución para detectar características locales en las imágenes, como bordes y texturas [14].

### 2.5 Redes neuronales recurrentes (RNN)

Las RNN se utilizan para procesar datos secuenciales, como texto o series temporales. Estas redes tienen conexiones que permiten el uso de información pasada para influir en la salida actual [17].

### 2.6 Transformers

Son arquitecturas modernas que ha revolucionado el PLN y otras áreas, utilizando mecanismos de atención para manejar relaciones a largo plazo en los datos [18].

### 2.7 Aprendizaje por transferencia personalizado

Es una técnica en el campo del ML y el DL que aprovecha el conocimiento adquirido en una tarea previa para mejorar el rendimiento o acelerar el aprendizaje en una tarea nueva pero relacionada. En lugar de entrenar un modelo desde cero, se reutilizan los modelos pre-entrenados en grandes conjuntos de datos y se ajustan para tareas específicas con datos limitados [19]. Adicionalmente, Los sistemas pueden ajustarse para ofrecer resultados que se alineen con las necesidades y niveles de conocimiento específicos del usuario, lo que mejora la selectividad en la recuperación [20].

### 2.8 Técnicas semánticas

El análisis semántico en el PLN es el proceso de comprender y extraer el significado de palabras, frases y textos completos. El desarrollo de técnicas semánticas es fundamental para interpretar el contexto y las relaciones entre palabras para derivar un significado coherente y preciso [21].

### 2.9 Algoritmos de recomendación

La IA permite la personalización en los sistemas de recuperación de información, ajustando los resultados de búsqueda según las necesidades específicas del usuario o contexto, lo que contribuye a una recuperación más efectiva y selectiva [2].

En la literatura, la predicción para la personalización de información se ha relacionado con el problema de realizar consultas difíciles. Una consulta difícil es aquella que tiene tantas respuestas posibles que es complejo recuperar las más apropiadas. Esto normalmente ocurre cuando la consulta es muy corta, ambigua o su tema es muy general. En tales situaciones, la personalización ayuda a proporcionar resultados más certeros sobre las preferencias usuario, quien a su vez estará más satisfecho con el SRI [19].

### 3 Metodología

En esta sección se describe la metodología utilizada para llevar a cabo la revisión sistemática.

El primer paso consistió en el planteamiento de las siguientes preguntas de investigación que guían el estudio:

- ¿Cómo se puede mejorar la precisión y relevancia de los resultados de búsqueda en los sistemas de recuperación de información?
- ¿Qué técnicas de inteligencia artificial son más efectivas para procesar y recuperar información de datos no estructurados como texto libre, imágenes y videos?
- ¿Qué estrategias pueden implementarse para asegurar la escalabilidad de los sistemas de recuperación de información al manejar grandes volúmenes de datos en tiempo real?

Como segundo paso y para dar respuesta a cada una las preguntas enlistadas, se realizó una investigación sobre los tópicos que se relacionan con la integración de la Inteligencia Artificial en la optimización de los Sistemas de Recuperación de la Información, considerando: artículos de investigación concerniente a la tecnología y la informática. Se consideró un periodo de búsqueda de 2019 a 2024, la cadena booleana utilizada en las bases de datos ProQuest y Google Scholar fue la siguiente: "Inteligencia Artificial" and "Sistema de recuperación de información" and "SRI" or "recuperación de información". Mientras que para la base de datos ScienceDirect se realizó la búsqueda utilizando las palabras clave en el idioma inglés: "Artificial Intelligence" and "Information Retrieval System" and "IRS" or "Information Retrieval".

Los resultados de la consulta en cada una de las bases de datos, arrojaron diversos documentos para su análisis como se observa en la Tabla 1.

Tabla 1. Total de artículos en cada base de datos consultada (Elaboración propia).

Base de datos	Número de artículos
Google Scholar	24
ProQuest	127
ScienceDirect	15
TOTAL	166

Como tercer paso se utilizó el diagrama de flujo PRISMA 2020 [22] (ver Figura 1) en el que se ilustra el total de fuentes consultadas y los criterios de elegibilidad para determinar la cantidad de trabajos de investigación que se incluyen en el análisis.

Para determinar el número de artículos analizados en el estudio se establecieron los siguientes criterios de inclusión y exclusión.

#### Criterios de inclusión

- Artículos publicados entre 2019 a 2024.
- Artículos científicos escritos en español e inglés.
- Artículos relacionados directamente con las preguntas de investigación que guían este estudio.

#### Criterios de exclusión

- Artículos publicados antes del 2019.
- Artículos científicos escritos en un idioma diferente al español e inglés.
- Artículos que no estuvieran directamente relacionados con las preguntas de investigación.
- Artículos duplicados.

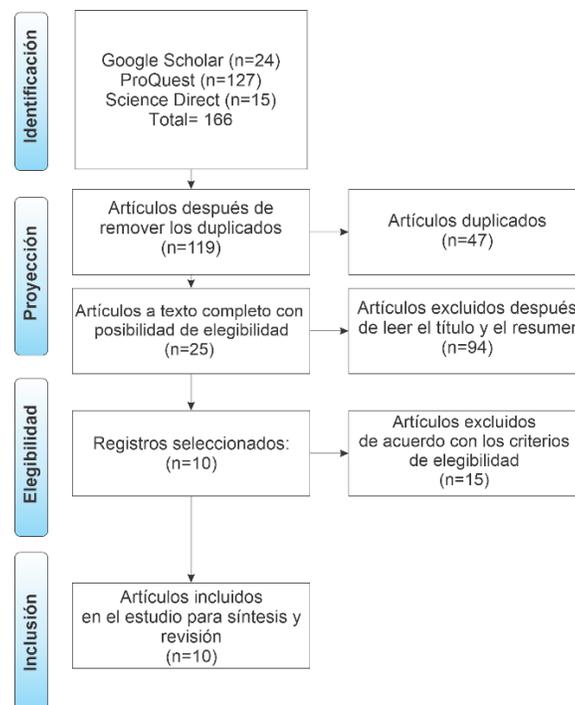


Figura 1. Diagrama de flujo PRISMA para la selección de los artículos incluidos en la revisión sistemática.

En la Tabla 2 se observa la cantidad de artículos incluidos por cada base de datos consultada.

Tabla 2. Número artículos incluidos en el estudio por base de datos consultada (Elaboración propia).

Base de datos	Número de artículos
Google Scholar	8
ProQuest	0
ScienceDirect	2
TOTAL	10

La muestra de artículos analizada en este estudio se presenta en la Tabla 3.

Tabla 3. Artículos seleccionados para la revisión sistemática (Elaboración propia).

Autor y referencia	Año	Título de la investigación
Celi-Parraga et al. [4]	2021	Técnicas de procesamiento de lenguaje natural en la inteligencia artificial conversacional textual
Chakraborty et al. [6]	2023	Information retrieval algorithms and neural ranking models to detect previously fact-checked information
Chouni et al. [5]	2019	Information retrieval system based semantique and big data
Fernández et al. [3]	2011	Modelos de Recuperación de Información basados en Redes de Creencia
Hambarde et al. [9]	2023	Information Retrieval: Recent Advances and beyond
Mitra et al. [11]	2017	Neural Models for Information Retrieval
Vicente-López et al. [13]	2018	Predicting IR personalization performance using pre-retrieval query predictors
Vitres Sala et al. [23]	2018	Semantic processing of information in Information Retrieval Systems
Yaguachi et al. [24]	2020	Uso del procesamiento del lenguaje natural (PLN) en la implementación de agentes conversacionales (bots) para perfiles de Facebook
Zamani et al. [25]	2022	Retrieval-Enhanced Machine Learning

## 4 Resultados

En la Tabla 4, se muestran las principales categorías de análisis encontradas en cada uno de los trabajos de investigación.

Tabla 4. Características de los estudios seleccionados (Elaboración propia).

Autor, Referencia y Año	Tópico central	Origen de los datos	Técnicas empleadas	Rendimiento de los resultados
Celi-Parraga et al. [4] (2021)	IA integrada a SRI	Texto conversacional	Procesamiento del Lenguaje Natural (PNL)	Mejor capacidad de comprensión y respuesta en IA conversacional
Chakraborty et al. [6] (2023)	IA integrada a SRI	Información previamente verificada	Algoritmos de ranking neuronal	Alta precisión en la detección de información verificada
Chouni et al. [5] (2019)	IA integrada a SRI	Big Data semántico	Algoritmos semánticos y tecnologías de Big Data	Escalabilidad y precisión mejorada en grandes volúmenes de datos
Fernández et al. [3] (2011)	Redes Bayesianas	Datos variados	Modelos de recuperación basados en redes de creencia	Mejora en la precisión y manejo de incertidumbre
Hambarde et al. [9] (2023)	IA integrada a SRI	Datos variados	Avances recientes en algoritmos de recuperación	Incremento en la precisión y eficiencia de los SRI modernos
Mitra et al. [11] (2017)	IA integrada a SRI	Datos no estructurados	Modelos neuronales para la recuperación de la información	Alta precisión en la recuperación de la información.
Vicente-López et al. [13] (2018)	Recuperación de información	Consultas personalizadas	Predictores de rendimiento de personalización	Mejora en la precisión de los resultados personalizados
Vitres Sala et al. [23] (2018)	Recuperación de información	Datos semánticos	Procesamiento semántico de información	Mejora en la relevancia y organización de los datos
Yaguachi et al. [24] (2020)	IA integrada a SRI	Perfiles de Facebook	Agentes conversacionales (bots)	Mejora en la interacción y respuesta de los bots
Zamani et al. [25] (2022)	IA integrada a SRI	Datos variados	Machine Learning mejorado por recuperación	Incremento en la precisión y eficiencia de los modelos de ML

De acuerdo con el análisis realizado una línea de investigación en tendencia para la recuperación de la información es el uso de redes bayesianas [3]. La lógica del formalismo basado en redes

bayesianas junto a otras herramientas experimentales como el SRI SMART se ha aplicado con éxito a los problemas existentes en esta disciplina.

Los autores Vicente-López et al. [13] realizaron un estudio exhaustivo utilizando predictores para consultas difíciles (previos a la recuperación). Estos predictores se basan en diferentes suposiciones sobre cómo los usuarios asignan relevancia a la lista de resultados recuperados. Dado que tanto la consulta como el perfil afectan los resultados de personalización. También hicieron la propuesta de generar nuevos predictores para mejorar el rendimiento de personalización; como línea de trabajo futuro podría ser no incluir toda la información del perfil del usuario, sino solo aquella que sea más relevante para la consulta dada, y esto es particularmente importante si los perfiles son heterogéneos y representan varias áreas de interés.

El artículo más reciente de Xing en 2024 sobre la Gestión Segura de Documentos y los SRI Inteligentes [26] nos destaca que el uso de modelos de lenguaje basados en aprendizaje automático para identificar patrones en las consultas y los documentos asegura que los resultados estén alineados con las expectativas y necesidades del usuario. Para el análisis y preparación de los datos, los autores Celi-Parraga et al. [4] y Viltres Sala et al. [23] sugieren diversas técnicas de preprocesamiento, como la tokenización, la lematización, la clasificación y el clustering con el objetivo de manejar grandes volúmenes de información no estructurada.

La integración de técnicas avanzadas de PLN como el procesamiento semántico, en modelos preentrenados ha sido fundamental para mejorar los sistemas conversacionales [4]. Estas técnicas permiten una comprensión más profunda del lenguaje natural. Esto mejora la recuperación de información en datos no estructurados, facilitando aplicaciones prácticas en chatbots y asistentes virtuales. Los modelos que consideran el contexto del usuario y el documento son más efectivos para proporcionar resultados relevantes. La incorporación de ontologías y grafos de conocimiento permite una mejor estructuración y relación de la información [23]. La evolución continua de estas técnicas es esencial para abordar los desafíos del procesamiento de datos no estructurados, los autores Moreira et al. [27] y Yaguachi et al. [24] indican que esto puede lograrse con el desarrollo de agentes conversacionales efectivos que puedan manejar datos textuales no estructurados, pues el análisis de sentimientos y la personalización de las respuestas son cruciales para mejorar la interacción y satisfacción del usuario.

La combinación de técnicas semánticas con tecnologías de Big Data asegura la escalabilidad del sistema, permitiendo manejar grandes volúmenes de datos de manera eficiente. Los índices distribuidos y las plataformas de procesamiento en paralelo pueden mantener la velocidad y eficiencia del sistema a medida que el volumen de datos crece. Otra técnica desarrollada para mejorar la precisión y relevancia de los modelos de aprendizaje automático es el aprendizaje mejorado por recuperación [25], que combina recuperación de información y aprendizaje automático. Esta técnica se integra en sistemas de recuperación en tiempo real, mejorando la capacidad del sistema para manejar grandes volúmenes de datos y consultas instantáneamente. Se

han desarrollado métodos para optimizar su rendimiento, asegurando que los sistemas puedan escalar eficazmente con el aumento de datos.

La precisión se mejora personalizando las recomendaciones en función de los criterios definidos por los usuarios, lo que sugiere que la incorporación de sistemas expertos con bases de conocimiento y reglas definidas puede mejorar la precisión en los resultados de búsqueda en ciertos dominios, de acuerdo con el sistema experto desarrollado por Guevara Orellana [28]. La mejora en la relevancia se logra mediante una adecuada construcción de perfiles de usuarios y necesidades, que permite un filtrado más preciso de la información. Adicionalmente, Martínez Musiño [29] concluye en su trabajo que los sistemas de indexación efectivos, combinados con algoritmos basados en métricas bibliométricas, mejoran la relevancia al destacar las publicaciones más citadas y útiles en un área específica.

Estos métodos han mejorado la representación de datos textuales y han aumentado la comprensión de los sistemas de RI sobre las consultas en lenguaje natural. Además, se han implementado mecanismos basados en atención como la arquitectura Transformer para mejorar la capacidad de los sistemas de RI para centrarse en partes críticas de la consulta y los documentos con fines de coincidencia [9].

La solución que se busca con respecto a la Recuperación de Información implica un estudio complejo: desde la incertidumbre involucrada en varios atributos del problema que pueden estar relacionados con la representación de documentos y consulta hasta el tamaño del problema a resolver, en cuanto al número de documentos y términos que contiene. Todo lo anterior supone obstáculos a la hora de procurar la organización y representación en sus mecanismos y una inferencia sólida. Para poder dar continuidad de forma adecuada a la solución de cualquier problema, se requiere poner en práctica los métodos correctos en condiciones aceptables en cuanto al tiempo de respuesta.

## **5 Discusión y conclusiones**

Los artículos revisados en el presente trabajo indican que la integración de IA en los SRI y el uso de redes Bayesianas se han enfocado en mejorar significativamente la precisión, eficiencia y capacidad de adaptación de estos sistemas, a través de la combinación de técnicas semánticas y tecnologías de Big Data [5], así como la utilización de redes neuronales y algoritmos de recomendación. Adicionalmente, la implementación de arquitecturas distribuidas y el uso de plataformas de computación en la nube son esenciales para asegurar la escalabilidad de los sistemas, permitiendo manejar grandes volúmenes de datos en tiempo real [25][30].

La capacidad de los SRI para entender y procesar grandes volúmenes de datos no estructurados, como texto, imágenes y videos, ha mejorado significativamente mediante el uso de redes neuronales y técnicas de ML, resultando en una recuperación más relevante y precisa de

información. Adicionalmente, las tecnologías de Big Data y los algoritmos avanzados permiten que los SRI procesen y analicen datos a gran escala de manera eficiente.

De manera paralela, la IA ha permitido la personalización de resultados de búsqueda basados en el perfil del usuario y el contexto de la consulta, brindando una mayor satisfacción del usuario, resultados de búsqueda más relevantes de acuerdo con el contexto deseado y adaptados a sus necesidades específicas.

La IA sigue avanzando en su capacidad para procesar y entender datos no estructurados como texto, imágenes y videos, lo que abre nuevas posibilidades para la recuperación de información. Estos modelos siguen siendo relevantes para la recuperación de información, ofreciendo una manera robusta de manejar la incertidumbre y las probabilidades en la toma de decisiones [3]. Con el uso de las técnicas basadas en redes Bayesianas, la toma de decisiones en condiciones de incertidumbre en los SRI ha mejorado considerablemente, gracias a una mejor modelización de relaciones complejas entre datos cuando se manejan evidencias probabilísticas de diversas hipótesis.

El futuro del desarrollo en estas áreas deberá centrarse en crear modelos más avanzados, mejorar la interpretabilidad y transparencia, y explorar nuevas aplicaciones y dominios para maximizar el impacto y la utilidad de estas tecnologías.

Es crucial mejorar la interpretabilidad y transparencia de los modelos de IA utilizados en SRI. Para mejorar la confianza en los sistemas de IA se debe primero comprender cómo y por qué se generan ciertos resultados por parte de los usuarios y para esto se necesita continuar desarrollando modelos de IA más avanzados y complejos que puedan comprender y procesar mejor la semántica y el contexto de las consultas de búsqueda a través de técnicas de DL y redes neuronales más profundas.

Las redes bayesianas pueden aplicarse a nuevos dominios y áreas donde el manejo de la incertidumbre y la probabilidad sea crítico. El nuevo enfoque añadido en dichas aplicaciones permitiría optimizar algoritmos y métodos de redes Bayesianas para mejorar su eficiencia computacional en grandes volúmenes de datos y realizar cálculos complejos en tiempo real de manera más eficiente. Adicionalmente, la combinación de redes Bayesianas con otras técnicas de IA, como redes neuronales y algoritmos de ML, puede ser usada para crear modelos híbridos que aprovechen las fortalezas de ambas aproximaciones.

Aunque las conclusiones de este estudio son un antecedente valioso para futuras investigaciones, es esencial tener en cuenta tanto las limitaciones como las posibles líneas de investigación que puedan surgir a partir de esta revisión sistemática.

Una de las principales limitaciones de este trabajo es el tamaño reducido de la muestra debido a las palabras clave utilizadas, el periodo de tiempo seleccionado y el número de bases de datos consultadas a las cuales se tuvo acceso. A continuación, se enlistan de manera breve algunas propuestas de trabajos a futuro que tienen el potencial de avanzar el campo de la recuperación de

información para mejorar significativamente la experiencia del usuario y la precisión de los sistemas de búsqueda en una variedad de aplicaciones:

- Desarrollar y probar nuevos algoritmos de recomendación que integren mejor la personalización y la seguridad, especialmente en sistemas que manejan datos sensibles.
- Explorar nuevas técnicas semánticas y su integración con tecnologías emergentes como la computación cuántica para mejorar la recuperación de información en grandes volúmenes de datos.
- Investigar formas de optimizar aún más las arquitecturas distribuidas para mejorar la eficiencia y reducir los costos de procesamiento en tiempo real.
- Investigar la integración de nuevas y diversas fuentes de datos, como datos de redes sociales, datos de IoT y datos en tiempo real, para enriquecer los resultados de búsqueda y proporcionar información más contextual y relevante.
- Profundizar en la integración de IA y redes neuronales en sistemas de recuperación de información para mejorar la precisión y relevancia de los resultados, especialmente en dominios específicos como la salud y el derecho.
- Evaluar la eficacia y eficiencia de los Sistemas de Recuperación de Información en tiempo real, identificando áreas de mejora y potenciales innovaciones tecnológicas.

## **Agradecimientos**

Los autores de este artículo agradecen al Consejo Nacional de Ciencia y Tecnología (CONACYT) y al Tecnológico Nacional de México Campus Orizaba por el apoyo otorgado en la realización de este trabajo.

## **Referencias**

- [1] Ceri, S., Bozzon, A., Brambilla, M., Della Valle, E., Fraternali, P., Quarteroni, S. (2013). An Introduction to Information Retrieval. In: Web Information Retrieval. Data-Centric Systems and Applications. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-39314-3\\_1](https://doi.org/10.1007/978-3-642-39314-3_1)
- [2] Veliz, E. (2020). Inteligencia Artificial-IN2-1. <https://doi.org/10.13140/RG.2.2.14627.48167>
- [3] Fernández, J. M., Granada, L., & De, M. (2011). Universidad de Granada Escuela Técnica Superior de Ingeniería Informática. Modelos de Recuperación de Información basados en Redes de Creencia, Tesis Doctoral.
- [4] Celi Parraga, R. J., Varela Tapia, E. A., Acosta Guzmán, I. L., & Montaña Pulzara, N. R. (2021). Técnicas de procesamiento de lenguaje natural en la inteligencia artificial conversacional textual. AlfaPublicaciones, 3(4.1), 40–52. <https://doi.org/10.33262/ap.v3i4.1.123>

- [5] Chouni, Y., Erritali, M., Madani, Y., & Ezzikouri, H. (2019). Information retrieval system based semantique and big data. *Procedia Computer Science*, 151, 1108–1113. <https://doi.org/10.1016/j.procs.2019.04.157>
- [6] Chakraborty, T., la Gatta, V., Moscato, V., & Sperli, G. (2023). Information retrieval algorithms and neural ranking models to detect previously fact-checked information. *Neurocomputing*, 557. <https://doi.org/10.1016/j.neucom.2023.126680>
- [7] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.
- [8] Mitra, B., & Craswell, N. (2018). An introduction to neural information retrieval. In *Foundations and Trends in Information Retrieval* (Vol. 13, Issue 1, pp. 1–129). Now Publishers Inc. <https://doi.org/10.1561/15000000061>
- [9] Hambarde, K. A., & Proenca, H. (2023). Information Retrieval: Recent Advances and beyond. *IEEE Access*, 11, 76581–76604. <https://doi.org/10.1109/ACCESS.2023.3295776>
- [10] Kuna, H., Rey, M., Martini, E., Solonezen, L., & Podkowa, L. (2014). Desarrollo de un Sistema de Recuperación de Información para Publicaciones Científicas del Área de Ciencias de la Computación. *Revista Latinoamericana de Ingeniería de Software*, 107–114.
- [11] Mitra, B., & Craswell, N. (2017). Neural Models for Information Retrieval. <http://arxiv.org/abs/1705.01509>
- [12] Kenter, T., Borisov, A., van Gysel, C., Dehghani, M., de Rijke, M., & Mitra, B. (2017). Neural networks for information retrieval. *SIGIR 2017 - Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1403–1406. <https://doi.org/10.1145/3077136.3082062>
- [13] Vicente López, E., de Campos, L. M., Fernández Luna, J. M., & Huete, J. F. (2018). Predicting IR personalization performance using pre-retrieval query predictors. *Journal of Intelligent Information Systems*, 51(3), 597–620. <https://doi.org/10.1007/s10844-018-0498-3>
- [14] LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., & Jackel, L. (1989). Handwritten digit recognition with a back-propagation network. *Advances in neural information processing systems*, 2.
- [15] Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

- [16] Eisenstein, J. (2019). Introduction to natural language processing. MIT press.
- [17] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
- [18] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [19] de Campos, L. M., Fernández-Luna, J. M., Huete, J. F., Ribadas-Pena, F. J., & Bolaños, N. (2024). Information Retrieval and Machine Learning Methods for Academic Expert Finding. *Algorithms*, 17(2). <https://doi.org/10.3390/a17020051>
- [20] Chen, Y., & Liang, S. (2024). BNMI-DINA: A Bayesian Cognitive Diagnosis Model for Enhanced Personalized Learning. *Big Data and Cognitive Computing*, 8(1). <https://doi.org/10.3390/bdcc8010004>
- [21] Goddard, C., & Schalley, A. C. (2010). *Semantic Analysis*.
- [22] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Alonso-Fernández, S. (2021). Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Revista española de cardiología*, 74(9), 790-799.
- [23] Viltres Sala, H., Leyva, R., Febles, J. P., & Sentí, V. E. (2018). Tipo de artículo: Artículo original  
Temática: Tecnologías de la información y las telecomunicaciones, Semantic processing of information in Information Retrieval Systems. *Revista Cubana de Ciencias Informáticas*, 12(1). <http://rcci.uci.cu>
- [24] Yaguachi, R., & Belén, V. (2020). Uso del procesamiento del lenguaje natural (PLN) en la implementación de agentes conversacionales (bots) para perfiles de Facebook.
- [25] Zamani, H., Diaz, F., Dehghani, M., Metzler, D., & Bendersky, M. (2022). Retrieval-Enhanced Machine Learning. *SIGIR 2022 - Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2875–2886. <https://doi.org/10.1145/3477495.3531722>
- [26] Xing, L. (2024). Secure Official Document Management and intelligent Information Retrieval System based on recommendation algorithm. *International Journal of Intelligent Networks*, 5, 110–119. <https://doi.org/10.1016/j.ijin.2024.02.003>

[27] Moreira, D., Cruz, I., Gonzalez, K., Quirumbay, A., Magallan, C., Guarda, T., Andrade, A., & Castillo, C. (2020). Análisis del Estado Actual de Procesamiento de Lenguaje Natural Analysis of the Current State of Natural Language Processing.

[28] Guevara Orellana, W. S. (2020). Prototipo de un Sistema Experto de Orientación Universitaria utilizando el concepto de Recuperación de Información de la Inteligencia Artificial.

[29] Martínez Musiño, C. (2020). La producción documental de la Inteligencia Artificial y la Ciencia de la Información en Scopus– Análisis informétrico. E-Ciencias de La Información. <https://doi.org/10.15517/eci.v11i1.42252>

[30] Suárez López, D., & Álvarez Rodríguez, J. M. (2023). Recuperación de información en IoT basada en ontologías de dominio. *Revista Ibérica de Sistemas Y Tecnologías de Información*, 247–255.

## **Tecnología y derechos humanos: Cómo las apps transforman la educación infantil**

### **technology and human rights: How apps are transforming early childhood education**

Juan Carlos Cruz Armas<sup>1</sup>, María del Carmen Santiago Díaz<sup>1</sup>, Judith Pérez Marcial<sup>1</sup>, Raúl Antonio Aguilar Vera<sup>2</sup>, Julio César Díaz Mendoza<sup>2</sup>, Ana Claudia Zenteno Vázquez<sup>1</sup>, Gustavo Trinidad Rubín Linares<sup>1</sup>.

<sup>1</sup>Facultad de Ciencias de la Computación, Benemérita  
Universidad Autónoma de Puebla. Av. San Claudio y 14 Sur,  
C. U., Col. San Manuel. 72592 Puebla, Puebla.

<sup>2</sup>Facultad de Matemáticas, Universidad Autónoma de Yucatán

juan.cruza@alumno.buap.mx, {marycarmen.santiago, judith.perez,, ana.zenteno, gustavo.rubin}@correo.buap.mx, avera@correo.uady.mx, julio.diaz@correo.uady.mx

### **Resumen**

La infancia es una etapa crucial en el desarrollo humano, ya que es un periodo en el que se debe garantizar la protección y el respeto de los derechos de los niños. Sin embargo, datos actuales por UNICEF revelan una realidad preocupante donde los derechos de los niños son violentados y ponen en riesgo su desarrollo por lo que se propone realizar una revisión de aplicaciones que permita su conocimiento y difusión de los derechos humanos y de aquellas que presenten maneras innovadoras de aprendizaje e interactivas. Este reporte concentra un análisis e identifica elementos que servirán para un trabajo futuro.

### **Abstract**

Childhood is a crucial stage in human development, as it is a period in which the protection and respect of children's rights must be guaranteed. However, current data from UNICEF reveal a worrying reality where children's rights are violated and their development is put at risk. Therefore, a review of applications that allow their knowledge and dissemination of human rights and those that present innovative and interactive ways of learning is proposed. This report concentrates an analysis and identifies elements that will serve for future.

---

Keywords and phrases: Derechos de los Niños, Teoría de Juegos, Educación,

---

## **1 Introducción**

Los derechos de los niños están reconocidos como derechos humanos fundamentales, los cuales buscan garantizar un buen desarrollo integral, tomando en consideración la fragilidad y sus necesidades específicas. Este marco legal tiene raíces históricas, como el de Rousseau, el cual nos menciona sobre el derecho de los niños al juego y la exploración, y el otro movimiento el de Hegel y Von Humboldt, donde promovía la educación orientada a la participación adulta. [1] Estos conceptos, en el siglo XVIII, se plasmaron en la convención sobre derechos de los niños y en leyes nacionales como lo es la Ley general de los derechos de Niñas, Niños y Adolescentes en México, donde nos garantizan los derechos esenciales como la vida, la educación, la salud, la identidad y la no violencia. También, establecen principios rectores como el interés superior de la infancia, no discriminación y igualdad, buscando la corresponsabilidad de las familias, la sociedad y el estado para poder garantizar el cumplimiento de estos derechos. [2] [3] [4]

La infancia es una etapa crucial en el desarrollo humano, ya que es un periodo en el que se debe garantizar la protección y el respeto de los derechos de los niños. Sin embargo, datos actuales revelan una realidad preocupante. Según UNICEF, cerca de 400 millones de niños menores de cinco años en el mundo, es decir, 6 de cada 10 niños de esa edad, sufren maltrato psicológico o agresiones físicas en sus hogares. De ellos, aproximadamente 330 millones experimentan castigos físicos. Esto refleja no solo una crisis del riesgo de la seguridad de los niños sino también en su desarrollo emocional y cognitivo.[5]

Desde 2021, el número de niños que no estudian a nivel mundial ha aumentado en 6 millones, alcanzado un total de 250 millones, según datos de la UNESCO. Este incremento se debe principalmente a la exclusión masiva de niñas y mujeres en Afganistán, donde 7,8 millones de niños no tienen acceso a la educación, a nivel global, 16% de los niños y jóvenes (desde primaria hasta el segundo ciclo de secundaria) permanecen fuera del sistema educativo y 122 millones de ellos son niñas. En África Subsahariana, 19,7 % de los niños no asisten a la escuela, convirtiéndose en la región más afectada. Este panorama evidencia una violación sistemática de los derechos de los niños a la educación, afectando su desarrollo y limitando sus oportunidades de futuro. [6]

## **2 Métodos de aprendizaje**

En la educación infantil se utilizan varios métodos que están diseñados para el desarrollo integral de los niños, así fomentando sus habilidades cognitivas emocionales y sociales.

El método Montessori fue creado por María Montessori, tiene como base la autonomía y el aprendizaje. Tiene varias etapas de desarrollo (0-6, 6-12, 12-18 años) y fomenta un entorno preparado, de autoevaluación y libertad con límites. Pueden tener actividades que incluyen simulaciones de asambleas, teatro y proyectos comunitarios para enseñar sus derechos como lo es la igualdad y el respeto. También promueve la independencia, requiere una formación especializada y materiales específicos. [7]

En la metodología de aprendizaje basado en proyectos (ABP) los estudiantes se involucran en la resolución de problemas reales mediante la investigación, creación y reflexión., pueden fomentar un aprendizaje significativo, como la colaboración y la aplicación práctica. Los estudiantes investigan, diseñan proyectos y los presentan, así desarrollan sus habilidades críticas y sociales.[8]

La gamificación integra elementos de juego en el aprendizaje, toma en cuenta los puntos y recompensas, para motivar y mejorar el rendimiento. Esta estrategia fomenta habilidades sociales, personalización del aprendizaje y la retroalimentación inmediata. Tiene un diseño adecuado, puede ayudar a reducir el estrés y generar experiencias significativas. Apps como búsquedas del tesoro, trivias o escape rooms que estimulan el aprendizaje dando recompensas en cada nivel. Los videojuegos también son herramientas para la educación y el desarrollo cognitivo. Tienen grandes beneficios como la mejora de la coordinación y el trabajo en equipo, sin embargo, deben manejarlos adecuadamente y con responsabilidad para evitar efectos negativos como sedentarismo y aislarse socialmente. Tienen que ser juegos enfocados donde se combinan el entretenimiento con enseñanza, así fortaleciendo el aprendizaje como derechos y valores. [9] [10]

### **3 Metodología**

La metodología empleada consiste en un proceso sistemático diseñado para identificar, recopilar, analizar y sintetizar la información actualizada y relevante sobre aplicaciones lúdicas y didácticas que enseñan a los niños temas que tienen relación con sus derechos humanos. Este enfoque permite garantizar la calidad y pertinencia de los datos recopilados, así como proporcionar un análisis, riguroso y estructurado de las herramientas disponibles que sirven para fomentar el aprendizaje de manera más interactiva y significativa.

Los métodos de aprendizaje han experimentado una notable transformación en los últimos años, gracias a los avances tecnológicos y cómo ha crecido el enfoque en la personalización de la educación. Estas innovaciones buscan no solo optimizar el aprendizaje, también es preparar a los estudiantes para los retos de la actualidad. A continuación, se presenta una revisión del estado del arte que involucra herramientas tecnológicas que abordan la enseñanza de los derechos de los niños, el impacto educativo y las estrategias empleadas para promover el respeto y conciencia hacia estos derechos. [11]

#### **3.1 Cuaderno de ejercicios para la enseñanza de los derechos humanos**

El Cuaderno de Ejercicios para la Enseñanza de los Derechos Humanos se presenta como una herramienta pedagógica diseñada para sensibilizar a niños, docentes y familias acerca de los derechos fundamentales de los más pequeños. Este material busca fomentar una reflexión profunda sobre la importancia de proteger y garantizar el bienestar de la infancia, mediante actividades prácticas, creativas y participativas.

El cuaderno comienza con una invitación a reflexionar sobre las vulneraciones que enfrentan muchos niños alrededor del mundo, tales como la falta de acceso a la educación, el juego, una

alimentación adecuada o el amor y cuidado familiar. Señala que la infancia, una etapa crucial para el desarrollo humano, a menudo es afectada por factores como la violencia, la negligencia y la pobreza. Este apartado insta a los adultos a asumir el compromiso de proteger a los niños, promoviendo su desarrollo pleno y garantizando sus derechos fundamentales. El cuaderno aborda temas esenciales relacionados con los derechos de los niños: derecho a vivir en un entorno libre de violencia y derecho a un medio ambiente sano. Incluye una amplia gama de actividades diseñadas para involucrar a los niños, sus familias y la comunidad educativa en el aprendizaje y vivencia de los derechos humanos: celebración de días internacionales, diario escolar, "comida de los países del mundo", creaciones artísticas colaborativas, veladas culturales, exposiciones y huerto escolar. [12]

### 3.2 La teoría videojuegos

La Teoría combina distintos enfoques interdisciplinarios para analizar diferentes medios con diferentes perspectivas. Parte de la teoría de juegos, que nació por 1944 con Von Neumann y Morgenstern, se enfoca en las decisiones interactivas de los jugadores, clasificando juegos en cooperativos, no cooperativos, estáticos, dinámicos y por el nivel de información. Usando conceptos clave de equilibrio de Nash, estrategias y pagos, las cuales permiten entender las dinámicas de interacción y los resultados. En los videojuegos, se exploran las mecánicas de juego, la narrativa y dinámicas lúdicas que interactúan con la experiencia del jugador. Los juegos con narrativa analizan historias y personajes, mientras la ludología estudia las estructuras y las reglas. Otros aspectos que incluyen el impacto social y la cultura de los videojuegos, la psicología del jugador y el diseño estético del juego que contribuye a la experiencia del jugador. Además, abarca modelos de negocio y economía, analizando cómo los videojuegos son desarrollados, financiados y distribuidos. La teoría integra elementos narrativos, artísticos y tecnológicos, destacando el valor cultural y comercial de los videojuegos. [13] Existen juegos educativos como los siguientes:

#### a) Khan Academy Kids

Una aplicación educativa gratuita diseñada para niños de 2 a 8 años. La aplicación ofrece una gran variedad de actividades, lecciones, libros y juegos que cubren temas como la lectura, escritura, matemáticas, habilidades socioemocionales y muchos más. Aquí los niños aprenden a contar, sumar y restar con juegos de matemáticas de forma divertida, mejoran la lectura y su escritura con actividades interactivas. Su objetivo es proporcionar una experiencia de aprendizaje divertida y atractiva que inspire a los niños a aprender y descubrir. Los personajes animados, como Cody el oso y sus amigos, guían y motivan a los niños a lo largo de su viaje de aprendizaje. La aplicación también permite a los padres y educadores seguir el progreso de los niños y personalizar su experiencia de aprendizaje. Khan Academy Kids es popular entre los niños porque combina el aprendizaje con el juego, haciendo que el proceso educativo sea entretenido y accesible. La aplicación está diseñada para ser intuitiva y fácil de usar, lo que permite a los niños explorar y aprender a su propio ritmo. Además, al ser completamente gratuita y disponible en múltiples plataformas, es una herramienta accesible para familias y escuelas. [14] Ver figura 1.



Figura 1. Khan Academy Kids

## b) Right Runner

Este juego invita a los jugadores a sumergirse en una aventura épica donde luchan por los derechos de todos los niños, presentando una temática claramente alineada con el propósito de tu estudio. A través de su narrativa, el juego permite a los jóvenes expresarse, inspirar a otros y reclamar sus derechos, mientras exploran distintos entornos y superan obstáculos que simbolizan las barreras que enfrentan en la vida real. El hecho de que esté ambientado en América Latina y el Caribe añade una dimensión cultural valiosa, mostrando cómo los derechos del niño pueden abordarse en diferentes contextos regionales. Los jugadores no solo deben superar obstáculos, sino que también tienen la posibilidad de interactuar con amigos virtuales que se unen a la lucha, recolectar megáfonos y difundir su mensaje, lo que refuerza la importancia de la colaboración y la participación para lograr un cambio. En cuanto a las mecánicas del juego, los controles sencillos permiten a los niños moverse rápidamente a través de diferentes escenarios, como deslizarse, saltar y realizar trucos, mientras aprenden de manera lúdica sobre los derechos que les corresponden. Esta interactividad no solo hace que el juego sea entretenido, sino que también favorece el aprendizaje práctico, al enseñar a los jugadores a identificar y superar obstáculos para defender sus derechos. Además, el capítulo final del juego, donde los niños deben "gritar" sus derechos a través del micrófono de su celular, simboliza la importancia de que sus voces sean escuchadas. También se destaca por su gratuidad y accesibilidad, al no incluir anuncios ni compras dentro de la aplicación. Esto lo convierte en una herramienta educativa abierta a un público más amplio, permitiendo que más niños accedan al conocimiento de sus derechos sin barreras económicas. Esta característica es especialmente relevante en el contexto de tu investigación, donde la accesibilidad a herramientas de aprendizaje es crucial para garantizar que todos los niños puedan beneficiarse de ellas. Finalmente, el mensaje central de *Right Runner* refuerza la importancia de los derechos de los niños, subrayando que todos tienen derechos sin importar dónde vivan o quiénes sean. Este aspecto resuena con la Convención sobre los Derechos del Niño, un punto clave en tu investigación, y muestra cómo un videojuego puede ser utilizado para promover la conciencia y el conocimiento sobre estos derechos de manera atractiva y eficaz.

Efectividad y Resultados: “En poco más de dos meses desde su lanzamiento, la App ha sido ya descargada más de 60.000 veces y tiene 4,4 estrellas de promedio en las valoraciones de los jugadores, tanto en la tienda de Apple como en la de Google”, dice Moreno González, que añade que, debido al éxito de las descargas en tan corto plazo, Apple ha notificado su interés de resaltar el juego en el mes de marzo de 2020. [15] Vea figura 2.



Figura 2. Right Runner

### c) My Rights and Duties

El juego My Right and Duties lanzado el 11 de noviembre de 2018 y desarrollado por Dubai Police es uno de los juegos que representa la importancia de educar a los futuros jóvenes de la mejor manera y tiene como objetivo concienciar a los niños sobre los derechos y deberes personales hacia su familia y la sociedad. Juega a este juego para elevar tu nivel de conciencia y responsabilidades hacia la comunidad. Conoce tus derechos de una manera divertida y sencilla. El juego contiene cuatro tipos de juegos interesantes y tres mundos con más de 100 etapas divertidas. Habla sobre los derechos y deberes familiares, abordando temas como el respeto, comunicación y colaboración. Derechos y deberes sociales, informa sobre los derechos y responsabilidades en la sociedad. Aspectos como la ciudadanía activa, la participación cívica y respeto a las leyes.

Efectividad y Resultados: tiene una valoración de 4.5 estrellas en Google Play. Los usuarios la encuentran útil para aprender sobre sus derechos y responsabilidades. [16] Ver figura 3.



Figura 3. My Rights and Duties

#### d) Rights Arcade

Amnistía Internacional lanzó el 24 de enero de 2022 Rights Arcade, una innovadora aplicación de videojuegos diseñada para educar a las nuevas generaciones sobre derechos humanos. Presentada en el Día Internacional de la Educación, esta aplicación tiene como objetivo sensibilizar a los jóvenes sobre derechos fundamentales como la libertad de expresión y la libertad de reunión pacífica, entre otros. Está diseñado para fortalecer el movimiento de derechos humanos a través de una educación que impulsa a la acción. El juego permite a los jugadores aprender sobre los derechos humanos de manera interactiva y reflexiva, fomentando un mayor entendimiento y compromiso con estos temas. Agnès Callamard, secretaria general de Amnistía Internacional, destaca que el juego está dirigido especialmente a los jóvenes, quienes son cruciales para definir y proteger los derechos humanos en el presente y en el futuro. El enfoque del juego es flexible, permitiendo a los jugadores avanzar a su propio ritmo mientras exploran historias inspiradas en hechos reales. Los jugadores siguen las experiencias de tres personas reales: Ahmed Kabir Kishore, un caricaturista injustamente acusado bajo la Ley de Seguridad Digital en Bangladesh; Zhang Zhan, una periodista ciudadana condenada a prisión por informar sobre la COVID-19 en China; y Panusaya Sithijirawattanakul, una activista estudiantil en Tailandia enfrentando múltiples cargos por protestar. Estas historias, aunque ficticias en su presentación, están basadas en casos reales y permiten a los jugadores tomar decisiones que reflejan su comprensión de los derechos humanos y su aplicación en la vida cotidiana. El juego está disponible en inglés, chino simplificado, tailandés y coreano, y puede ser descargado en dispositivos iOS y Android, con la garantía de accesibilidad en áreas con conexión a Internet deficiente. Rights Arcade se actualizará periódicamente con nuevos juegos y ampliará su disponibilidad a más idiomas en el futuro.

Efectividad y resultados: Esta app, tiene poca información, pero por lo revisado desde la appstore, tiene una calificación de 5 estrellas, pero no significa ser muy efectiva ya que no son muchos los usuarios en calificar esta app. [17] Ver figura 4.



Figura 4. Rights Arcade

e) Among Us

Es un juego multijugador en línea desarrollado por InnerSloth. En este juego, los jugadores asumen uno de dos roles: Tripulante o Impostor. El juego se desarrolla en una nave espacial, una base aérea o una estación en un planeta alienígena. Los Tripulantes tienen la tarea de completar varias misiones alrededor del mapa, mientras que los Impostores intentan sabotear estas misiones y eliminar a los Tripulantes sin ser descubiertos. El objetivo para los Tripulantes es completar todas las tareas asignadas o descubrir y expulsar a todos los Impostores. Por otro lado, los Impostores deben eliminar a todos los Tripulantes o sabotear misiones clave para ganar el juego. Los Tripulantes realizan tareas simples como conectar cables y descargar datos, mientras que los Impostores pueden sabotear sistemas críticos para crear caos y eliminar a los Tripulantes. Cuando se descubre un cuerpo o alguien presiona el botón de emergencia, los jugadores discuten y votan para expulsar a quien creen que es un Impostor. Es muy popular entre los niños por varias razones. Fomenta la comunicación y el trabajo en equipo, lo que lo hace muy atractivo para los niños que disfrutan jugando con amigos. Las reglas son fáciles de entender y el juego está disponible en múltiples plataformas, incluyendo móviles y PC. La mezcla de tareas simples con la emoción de descubrir o ser descubierto como Impostor mantiene a los jugadores enganchados. Además, los jugadores pueden personalizar sus personajes con diferentes colores y accesorios, lo que añade un elemento divertido y personal al juego. [18] Ver figura 5.

f) Minecraft

Uno de los videojuegos más populares jugados por niños en 2024 es Minecraft, el cual ha mantenido su relevancia por su versatilidad. Este juego ofrece dos modos principales: *Creativo*, que permite a los niños explorar su creatividad construyendo mundos sin restricciones, y *Supervivencia*, donde deben reunir recursos y protegerse de enemigos. Minecraft es valorado no solo por su entretenimiento, sino también por fomentar habilidades como la colaboración y la resolución de problemas, lo que lo hace ideal tanto para niños como para adultos. [19] Ver figura 6.

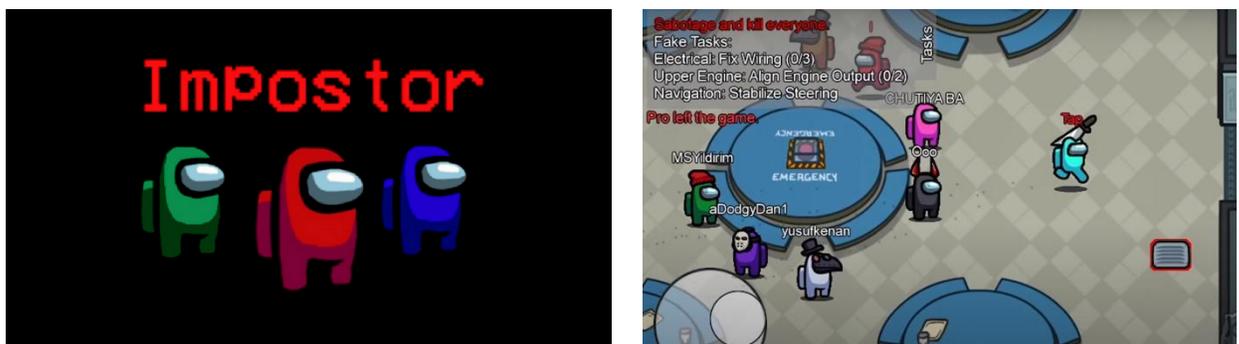


Figura 5. Among Us



Figura 6. Minecraft

#### 4 Resultados

A continuación, presentamos una tabla donde compara los videojuegos enfocados en su mayoría, en la educación y aprendizaje, analizando aspectos como la edad recomendada, la plataforma en que están disponibles, sus objetivos educativos, el entretenimiento, su duración, historia, ventajas y desventajas. Cada juego tiene diferentes temáticas, fomentan la creatividad y la enseñanza de los derechos de los niños. La tabla 1 comparativa tiene la finalidad de proporcionar un panorama sobre cómo los juegos combinan distintos elementos lúdicos con objetivos educativos, destacando sus fortalezas y posibles limitaciones y según sus características.

#### 5 Conclusiones

Las aplicaciones tecnológicas y principalmente los videojuegos vienen siendo una magnífica opción para que los niños comprendan sus derechos y las App de juegos como Right Runner, Among Us, My Rights and Duties y Rights Arcade comparten las ventajas de ser gratuitos, lo que ayuda a tener más accesibilidad. Sin embargo, tienen diferentes enfoques, mientras que Right Runner tiene una experiencia corta y divertida, Rights Arcade ofrece un contenido más profundo de los derechos humanos. Mientras títulos como Minecraft tiene un alto potencial creativo y multijugador, ofrece una experiencia de creatividad y habilidades sociales, aunque su costo inicial puede limitar su alcance. Among Us, que no tiene un enfoque educativo, pero sí fomenta habilidades importantes como la deducción y el trabajo en equipo a través del modo multijugador, con opciones más extensas como Khan Academy Kids y Minecraft, frente a experiencias más breves como My Rights and Duties. En general, estos resultados reflejan que cada juego satisface necesidades específicas y puede ser aprovechado según los intereses y objetivos educativos del usuario.

Tabla 1. Comparativo de App educativas

Juego	Edad Recomendada	Plataforma	Aspectos Educativos	Divertido	Duración	Historia	Ventajas	Desventajas
<a href="#">Right Runner</a>	8-14 años	iOS	Enseñar los derechos de los niños	Divertido	Corto	Tiene	Gratis, educativo	Solo es para iOS
<a href="#">Khan Academy Kids</a>	3-7 años	iOS, Android	Fomentar el aprendizaje temprano en diversas áreas (lectura, matemáticas, etc.)	Moderado	Largo	No	Gratis, contenido extenso y variado, educativo	Requiere conexión a internet, puede ser abrumador por la cantidad de contenido
<a href="#">Minecraft</a>	7+	PC, Consolas, Móvil	Creatividad, resolución de problemas, colaboración	Divertido	Largo	Si	Altamente creativo, educativo, multijugador	Costo inicial, compras dentro del juego, puede ser adictivo
<a href="#">Among Us</a>	10+ años	PC, Consolas, Móvil	Deducción, Trabajo en equipo	Divertido	Medio	No	Gratis en móvil, multijugador	frustrante para algunos, requiere conexión a internet
<a href="#">Rights Arcade</a>	10-16 años	iOS, Android	Enseñar sobre derechos humanos y su protección	No	Medio	Sí	Gratis, educativo, promueve derechos	Gráficos simples
<a href="#">My Rights and Duties</a>	6-12 años	iOS, Android	Enseñar a los niños sobre sus derechos y deberes	Moderado	Corto	No	Gratis, educativo	Gráficos simples, puede no atraer a todos los niños

## Referencias

[1] Butler, C. (Ed.). (2012). *Child rights: The movement, international law, and opposition* (pp. 13-37). Purdue University Press. <https://doi.org/10.2307/j.ctt6wq6tt>

[2] UNICEF Comité Español. (2024). *Convención sobre los Derechos del Niño*. UNICEF. Recuperado de <https://www.unicef.es/publicacion/convencion-sobre-los-derechos-del-nino-0n>.

[3] Cámara de Diputados del H. Congreso de la Unión. (2014). *Ley para la Protección de los Derechos de Niñas, Niños y Adolescentes*. Última reforma publicada en el DOF el 2 de abril de 2014. Secretaría General, Secretaría de Servicios Parlamentarios.

[4] Gobierno de México. (2000). *Ley para la Protección de los Derechos de Niñas, Niños y Adolescentes*. Diario Oficial de la Federación. <https://www.dof.gob.mx/> (Fecha de la última reforma: 2 de abril de 2014).

[5] Unicef. (2024, junio 11). *Cerca de 400 millones de niños y niñas de todo el mundo sufren disciplina violenta*. Recuperado el 11 de enero de 2025, de <https://www.unicef.org/lac/comunicados-prensa/cerca-de-400-millones-de-ninos-y-ninas-de-todo-el-mundo-sufren-disciplina-violenta>.

[6] Unesco. (2023, septiembre 23). *250 millones de niños sin escolarizar: Lo que debemos saber acerca de los datos recientes de la educación global*. Recuperado el 11 de enero de 2025, de <https://www.unesco.org/es/articles/250-millones-de-ninos-sin-escolarizar-lo-que-debemos-saber-acerca-de-los-datos-recientes-de-la#:~:text=Una%20encuesta%20reciente%20realizada%20en,asistir%20a%20la%20escuela%20secundaria>.

[7] Britton, L. (2000). *Montessori Play and Learn* (P. Paterna Molina, Trad., 1.ª ed. en esta presentación, 2017). Vermilion, The Random House Group. (pp. 14-26).

[8] Secretaría de Educación Pública (SEP). (2022). *Metodología del aprendizaje basado en proyectos (ABP)*. SEP.

[9] Contreras Espinosa, R. S., & Eguía, J. L. (Eds.). (2017). *Experiencias de gamificación en aulas* (pp. 7-99). InCom-UAB Publicacions, 15. Institut de la Comunicació, Universitat Autònoma de Barcelona. ISBN 978-84-944171-6-0.

[10] Videojuegos y los niños. Consejos y cuidados. Clínica Universidad de Navarra. (2023). <https://www.cun.es>. Recuperado 11 de septiembre de 2024, de <https://www.cun.es/chequeos-salud/infancia/videojuegos-ninos>

[11] ¿Qué metodologías para el aprendizaje se utilizarán en el nuevo modelo pedagógico? (2024, 21 febrero). Red de Colegios Semper Altius. Recuperado 22 de septiembre de 2024, de <https://www.semperaltius.edu.mx/blog-post/metodologias-para-el-aprendizaje>

[12] Naciones Unidas. (2003). *Cuaderno de ejercicios para la enseñanza de los derechos humanos*. Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos, publicado por la Oficina Regional de Educación de la UNESCO para América Latina y el Caribe (OREALC/UNESCO Santiago). (pp. 8-100).

[13] Cerda Tena, E., Pérez Navarro, J., & Jimeno Pastor, J. L. (2004). *Teoría de juegos* (528 págs.). Pearson Educación, S.A. Recuperado el 23 de septiembre de 2024.

[14] Khan Academy. (2008). *Un recurso de aprendizaje personalizado para todas las edades*. Recuperado el 23 de septiembre de 2024, de <https://es.khanacademy.org/about#:~:text=Un%20recurso%20de%20aprendizaje%20personalizado,dentro%20como%20fuera%20del%20aula>

[15] UNICEF. (20 de marzo de 2020). *Right Runner*. UNICEF. Recuperado el 11 de septiembre de 2024, de <https://www.unicef.org/lac/right-runner>

[16] *My Rights and Duties - Aplicaciones en Google Play*. (s. f.). <https://play.google.com/store/apps/details?id=com.dubaipolice.rightsandduties&hl=es&pli=1>

[17] Amnistía Internacional. (2022, 24 enero). *Amnistía presenta una aplicación de videojuegos para facilitar el aprendizaje sobre derechos humanos*. Recuperado 11 de septiembre de 2024, de <https://www.amnesty.org/es/latest/news/2022/01/amnesty-launches-game-app-to-make-human-rights-learning-a>

[18] Fernández, Y. (2021, enero 26). *Among Us: Qué es, cómo se juega y cómo descargarlo en Android, iOS o Windows*. Xataka. Recuperado el 23 de septiembre de 2024, de <https://www.xataka.com/basics/among-us-que-como-se-juega-como-descargarlo-android-ios-windows>

[19] Espeso, P. (2016, febrero 2). *Descubre qué es Minecraft y por qué todo el mundo habla de él*. EDUCACIÓN 3.0. Recuperado el 23 de septiembre de 2024, de <https://www.educaciontrespuntocero.com/noticias/que-es-minecraft-educacion/>

## **Identificación automática de género a partir de habla en español**

### **Automatic gender identification from spanish speech**

Brayan Blancas Monsalvo, Maya Carrillo Ruiz

Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla.

brayan.blancas@alumno.buap.mx, maya.carrillo@buap.mx.

#### **Abstract**

Gender recognition in speech refers to the process of identifying the gender of a speaker based on their voice. Each speaker has determined acoustic characteristics by the structure of their vocal apparatus. Therefore, by analyzing specific parameters from speech, it is feasible to identify an individual's gender. In this work, two speech representation schemes are used: Mel-Frequency Cepstral Coefficients and Mel Spectrograms. The classification algorithms: KNN, SVC, SGDC, Perceptron and Random Forest are applied for gender identification, with SGDC achieving 96% accuracy on a dataset with 2,000 examples.

#### **Resumen**

El reconocimiento del género del habla se refiere al proceso de identificar el género de un hablante en función de su voz. Cada hablante tiene características acústicas particulares determinadas por la estructura de su aparato bucal, por lo tanto, analizando parámetros específicos a partir del habla es factible identificar el género de un individuo. En este trabajo se utilizan dos esquemas de representación del habla: Coeficientes Cepstrales de Frecuencias de Mel y Espectrogramas de Mel. Los algoritmos de clasificación: KNN, SVC, SGDC, Perceptrón y Random Forest son aplicados para la identificación del género del hablante, de los cuales SGDC alcanza una exactitud del 96% con un conjunto de datos de 2,000 ejemplos.

---

Keywords and phrases: Automatic Gender Identification, Machine Learning, MFCC, Mel Spectrogram.

---

### **1 Introducción**

Los seres humanos somos capaces de recuperar información asociativa, cuando escuchamos hablar a una persona conocida, aunque no la veamos, la identificamos. En otras ocasiones, aunque no conozcamos a una persona, a través de su voz intuimos su estado de ánimo, hacemos conjeturas de su género y edad. Cada hablante tiene características acústicas particulares determinadas por la

estructura de su aparato vocal. De manera general la voz de las mujeres tiene patrones de frecuencias o timbre más altos que la de los hombres. Por lo tanto, analizando parámetros específicos, a partir del habla es factible identificar el género de un individuo [1].

El reconocimiento de género del habla se refiere al proceso de identificar el género de un hablante en función de las características de su voz. Para este propósito se utilizan algoritmos de aprendizaje automático que analizan diversas características acústicas del habla, como el tono, las frecuencias y la calidad de la voz [2].

En este trabajo se comparan diferentes algoritmos de aprendizaje automático para identificar el género del hablante a partir de audios en español. Se utilizan conjuntos de datos de diferentes tamaños para entrenar los modelos y un conjunto de datos diferente para realizar predicciones. La principal aportación de este proyecto es la comprobación de que los espectrogramas de Mel permiten identificar con mayor exactitud el género del hablante. El trabajo está organizado de la siguiente forma: En la sección 2 se presentan algunos conceptos y trabajos relacionados, en la 3 se describe la arquitectura del sistema, en la 4 están los experimentos y resultados obtenidos y en la 5 las conclusiones y trabajo futuro.

## **2 Marco teórico y estado del arte**

A continuación, se definen conceptos importantes para la comprensión de este artículo y algunos trabajos relacionados.

### **2.1 Representaciones utilizadas**

Para caracterizar el habla existen diferentes representaciones, en este trabajo se utilizan los Coeficientes Cepstrales en las Frecuencias de Mel (MFCC) que son coeficientes para la representación del audio basada en la percepción de las señales acústicas por parte del oído humano. Un problema fundamental en el procesamiento del habla es obtener una codificación compacta de las características del audio. Una de las técnicas más usadas para la extracción de estas características son los MFCC [3]. Otra representación utilizada son los Espectrogramas de Mel, un espectrograma mel es una variante del espectrograma que se utiliza comúnmente en el procesamiento del habla y en tareas de aprendizaje automático, es una imagen del audio.

### **2.2 Algoritmos de clasificación**

K vecinos más cercanos o KNN (K-Nearest Neighbors) es un algoritmo de aprendizaje automático supervisado que puede usarse tanto para clasificación como regresión. Es uno de los más sencillos debido a la forma de clasificar, y es que únicamente se realiza una votación de los k vecinos más cercanos a un punto de dato, y gana la clase que tenga la mayor cantidad de representantes. Por lo tanto, KNN no crea un modelo, sino que almacena las instancias de entrenamiento [9].

Máquina de soporte vectorial o SVM (Support Vector Machine) es un algoritmo de aprendizaje automático que tiene como principal propósito encontrar un hiperplano que sea capaz de separar

puntos de datos en clases. Los puntos de datos más cercanos al hiperplano se conocen como vectores de soporte, y se maximiza la distancia o margen entre dicho hiperplano y los vectores para reducir el riesgo de clasificar erróneamente nuevos puntos de datos. Un kernel es una función usada para aumentar la dimensión en un espacio de características para poder resolver problemas que no son linealmente separables [8]. La implementación LinearSVC proporcionada por scikit-learn es una optimización de SVM para problemas de clasificación lineal, donde la diferencia principal es el uso del kernel lineal y la función de pérdida para la maximización del margen [10].

La clase SGDC (Stochastic Gradient Descent Classifier) proporcionada por scikit-learn es un modelo de clasificación que usa por defecto SVM, y tiene como función de optimización el algoritmo de gradiente descendiente estocástico. Dicho algoritmo calcula el gradiente de la función objetivo y actualiza los parámetros del modelo en dirección opuesta al gradiente, para minimizar o maximizar la función objetivo. El término estocástico se refiere a que se toman muestras aleatorias en lotes y no todo el conjunto de datos en cada iteración [17].

Un perceptrón es el modelo de redes neuronales más simple que existe. Está conformado por un conjunto de entradas, pesos asignados a cada entrada, una función de activación, un bias o sesgo y una salida. La forma en la que un perceptrón consigue hacer clasificaciones es que realiza una suma del producto de cada entrada con su correspondiente peso y el sesgo, el resultado de la suma pasa por una función de activación la cual determina la salida final de la neurona. Al crear capas de perceptrones y asignarles un rol, siendo las más comunes capa de entrada, capa oculta y capa de salida, se pueden resolver problemas de clasificación al optimizar el peso de las entradas y salidas [11]. La implementación de scikit-learn llamada MLPC (Multi-Layer Perceptron Classifier) tiene por defecto una capa oculta de 100 neuronas y la función de activación ReLU [12].

Un árbol de decisión es un modelo de aprendizaje automático usado para problemas de clasificación y regresión. A partir de las características de los datos, un árbol puede aprender decisiones las cuales tomará para clasificar los datos [13]. Un nodo de división consiste en una condición sobre una característica, y los nodos hoja contienen la clasificación final. Un bosque aleatorio o Random Forest combina el resultado de múltiples árboles de decisión para mejorar la precisión, siendo cada árbol creado con una muestra aleatoria de datos [14].

### 2.3 Trabajos relacionados.

Levitan S.I. et al. [4] presentan identificación de género en hablantes de inglés, y alemán. Para la identificación en inglés utilizan un subconjunto del corpus HMIHY ("How May I Help You") de 4520 enunciados con etiquetas de género. La mayor exactitud que obtienen utilizando como representación del audio frecuencia fundamental es de 93.5 % y con MFCC 92.8 % utilizando regresión logística para la clasificación. Para el alemán utilizan el corpus aGender que tiene audios de mujeres, hombre y niños, considerando estas tres clases obtienen una exactitud de 83.3 y 85% para frecuencia fundamenta y MFCC respectivamente, utilizando Random Forest como clasificador.

Finalmente combinan audios en inglés y alemán eliminando los audios de niños y la mayor exactitud que obtienen es de 92.1% con regresión logística.

Abella M. G. J. , et al. [5] presentan un estudio que utiliza una combinación de análisis de componentes principales (Principal Component Analysis) y funciones kernel (Kernel PCA) para reducir datos, así como MFCC que se combinan con varias características de corto y largo plazo para detectar el género del habla. Recopilan 250 voces filipinas en el Metro de Manila, las limpian y extraen 25 características como MFCC, tono, tasa de cruce por cero, energía, centroide espectral, flujo espectral y edad. Los clasificadores utilizados son Random Forest, KNN y Redes Neuronales o ANN (Artificial Neural Network). Crean 4 modelos en total en los que utilizan Kernel PCA y PCA. De los 4 modelos, aquellos sin Kernel PCA tuvieron mayor exactitud. El clasificador que obtuvo la mayor exactitud fue el de ANN con un 92.06 %.

K. Nugroho, et al. [6] identifican el género de personas javanesas mediante su voz para ello utilizan MFCC. La clasificación de voz la realiza utilizando algoritmos de aprendizaje profundo, combinados con el método de descomposición en valores singulares (SVD) para reducir las dimensiones de los datos producidos. Utilizan una relación de división del 70% para los datos de entrenamiento y del 30% para los datos de prueba, obtienen la mayor exactitud, con el clasificador de aprendizaje profundo, que es del 97.78 % con regresión logística del 95,56 % y con SVM del 93.33 %.

### 3 Arquitectura de la aplicación

Para realizar los experimentos de identificación de género se diseñó una aplicación cuyas etapas principales se encuentra en la siguiente figura.

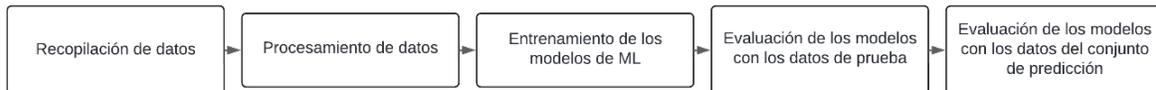


Figura 1. Etapas para entrenar modelos de aprendizaje automático para identificación de género.

#### 3.1 Recopilación de datos

En la primera etapa para construir la aplicación de identificación de género se recopilan datos para entrenar y probar los modelos. En este caso, el principal desafío fue encontrar audios en español correctamente etiquetados para un enfoque de aprendizaje supervisado. Entre los diversos conjuntos de datos (data sets) de audios disponibles, se encontró uno de acceso libre, TEDx Spanish Corpus [15]. Este conjunto de datos contiene discursos de varios expositores de TEDx, clasificados según el género del ponente, e incluye el audio, las transcripciones y un archivo de Excel con un resumen de la información contenida. Este conjunto de datos se utilizó para entrenar los modelos de aprendizaje automático.

Se conformó un segundo conjunto de datos uniendo audios de voluntarios de 4 países de habla hispana: Chile, Colombia, Perú y Venezuela [7]. Este segundo conjunto de datos se utilizó para realizar predicciones.

Los archivos de audio de TEDx poseen las siguientes características:

1. Son de 3 a 10 segundos, con una duración de 24 horas y 29 minutos
2. Los 11,243 archivos cuentan con 142 oradores distintos, 102 hombres y 40 mujeres
3. Están clasificados según el género del hablante

Los archivos de audio del conjunto de predicción son 4 corpus de un total de 6 disponibles, hecho mediante crowdsourcing o, en otras palabras, con voluntarios que grabaron su voz en alta calidad. Junto a los archivos de audio, se encuentra un archivo .csv que contiene la transcripción de cada uno de los archivos mencionados. La Tabla 1 resume características del conjunto de predicción.

Tabla 1. Información detallada sobre los 4 corpus consultados (conjunto de predicción)

Corpus	Género	Número de archivos	Duración (horas)	Número de oradores
Chile	Hombre	2,636	4.31	18
	Mujer	1,738	2.84	13
Colombia	Hombre	2,534	3.84	17
	Mujer	2,369	3.74	16
Perú	Hombre	2,918	4.87	20
	Mujer	617	1	5
Venezuela	Hombre	1,754	2.40	12
	Mujer	1,603	2.41	11

### 3.2 Procesamiento de datos

Con los conjuntos de datos recopilados, se realiza un procesamiento para que los datos sean de utilidad. Primero, es crucial extraer características que los modelos de aprendizaje automático (ML) puedan usar para identificar patrones y hacer clasificaciones a partir de ellos. Se eligieron dos técnicas para la extracción de características a partir de archivos de audio: MFCC y espectrogramas de Mel. Ambas representaciones se utilizaron con el objetivo de compararlas y determinar cuál ofrece los mejores resultados. Se estableció un número de 13 coeficientes para MFCC y 132 bandas para el espectrograma de Mel.

La librería librosa de Python usada para el procesamiento de los archivos de audio tiene implementado MFCC y los espectrogramas de Mel. El procesamiento de los audios procesados con MFCC [16] se hizo con 13 coeficientes de Mel, una tasa de muestreo de 22,050 Hz (número de muestras por segundo) y normalización de tipo ortogonal, transformada discreta del coseno de tipo II y una normalización del filtro Mel de tipo slaney. Para los audios procesados con espectrograma de Mel [18], se realizó con 132 bandas de Mel, una tasa de muestreo de 22,050 Hz, un tamaño de salto entre ventanas adyacentes de 512, un tamaño de ventana de la transformada rápida de Fourier (FFT) de 2048, función de ventana de tipo hann y con la señal centrada antes de aplicar la FFT.

Al indicar una tasa de muestreo de 22,050 Hz, se crean 22,050 registros por cada segundo analizado, lo que aumenta la dimensionalidad del espacio de características en cada representación, MFCC o espectrograma de Mel, por lo que se realizó un promedio para así reducir el número de características y contar con un único registro por cada audio procesado.

### 3.3 Entrenamiento de modelos de ML

Dado que se usaron dos representaciones para obtener características de los audios, se crearon modelos de ML para cada representación, es decir, entrenados con datos procesados mediante MFCC y espectrogramas de Mel. Se definieron diferentes tamaños de corpus equilibrados (mismo número de hombres y mujeres), iniciando con 200 ejemplos hasta 2,000 con un incremento de 200. Los corpus son independientes y generados aleatoriamente en cada incremento.

### 3.4 Evaluación de los modelos con datos de prueba

Se definieron modelos ML para cada corpus definido considerando el 80% de los datos para entrenar y 20% para probar. La exactitud presentada es la correspondiente al conjunto de pruebas.

### 3.5 Evaluar los modelos con nuevos datos

Los modelos se probaron con un conjunto de datos de predicción, diferente del conjunto de datos de entrenamiento, lo que significa que son datos que los modelos no han visto previamente. Se extrajeron 25 muestras aleatorias de hombres y mujeres de cada país (conjunto de predicción), y se experimentó con los modelos para evaluar su capacidad de reconocer el género del locutor, a pesar de las diferencias entre los audios del dataset de TEDx y el conjunto de predicción.

La principal diferencia entre el dataset de TEDx y el conjunto de predicción es que los audios de TEDx son extracciones de audio de exposiciones, mientras que los archivos del conjunto de predicción son grabaciones de alta calidad realizadas por voluntarios. Además, los archivos del conjunto de predicción presentan un acento marcado, correspondiente al lugar de origen del voluntario que realiza la grabación.

## **4 Experimentos y resultados**

A continuación, se presentan los experimentos y resultados obtenidos.

#### 4.1 Experimentos

El experimento principal consistió en conocer cuál representación es la mejor, espectrogramas de Mel o MFCC para entrenar modelos de ML que sean capaces de reconocer el género del locutor a partir de un audio.

Se eligieron 5 algoritmos de aprendizaje automático para crear modelos de predicción: KNN, SVC, SGDC, MLPC y Random Forest. Por cada algoritmo, se crearon dos modelos, uno para cada representación de los datos, MFCC y espectrogramas de Mel, todos creados usando el módulo scikit-learn de Python y usando los parámetros predeterminados.

Todos los algoritmos tienen diferencias significativas entre sí, por lo cual se eligieron algoritmos distintos para descubrir cuál forma de resolver el problema de clasificación que nos ocupa, sería la adecuada para este trabajo. Se consideraron algoritmos basados en árboles de decisión, distancia, funciones, y redes neuronales.

Se estableció una serie de conjuntos de datos, partiendo desde las 200 hasta las 2,000 con un incremento de 200 siendo 50% hombres y 50% mujeres en cada iteración, como se mencionó previamente, esto con el propósito de saber cómo cambia la exactitud de los modelos en función del número de ejemplos disponibles para entrenar.

Con los modelos debidamente entrenados y evaluados con datos de pruebas, estos se verificaron con datos ajenos a los vistos durante el entrenamiento, para saber que modelo conseguiría la mejor exactitud.

#### 4.2 Resultados

Los resultados obtenidos de los datos de prueba (pertenecientes al conjunto de TEDx) utilizando espectrograma de Mel se pueden observar en la Figura 2. En ella se observa que el algoritmo que mejor comportamiento presenta fue SVC, con una exactitud máxima del 100% y una mínima del 95%.

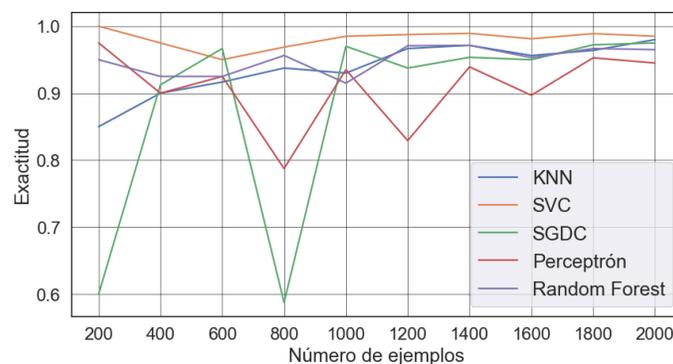


Figura 2. Exactitud del conjunto de pruebas variando el número de ejemplos, usando espectrograma de Mel.

Los resultados obtenidos con el conjunto de pruebas usando MFCC se pueden observar en la Figura 3. Todos los modelos parecen seguir la tendencia esperada, entre mayor el número de ejemplos disponibles para entrenar, mejor la exactitud. El modelo creado con Random Forest es el que alcanzó la mejor exactitud, siendo 96.11%. En contraste con los modelos que usan espectrograma de Mel, hay diferencias significativas en la exactitud que alcanzan SVC y SGDC con el máximo número de ejemplos, pues esta decrece.

Los resultados obtenidos de los datos del conjunto de predicción (Chile, Colombia, Perú y Venezuela) usando espectrograma de Mel se puede observar en la Figura 4. Se aprecia una gran diferencia entre los resultados de pruebas realizadas con los mismos datos de entrenamiento, y con nuevos datos nunca vistos durante el entrenamiento. KNN es quién tiene la exactitud más baja, 77.5%. El caso de SVC y SGDC es interesante, SVC empieza con la exactitud más alta con el menor número de ejemplos y SGDC lo contrario, con 95% y 58% respectivamente, pero al llegar a 2,000 ejemplos, se consigue lo contrario, SVC finaliza con una exactitud más baja que la lograda con SGDC, 80.5% y 96% respectivamente.

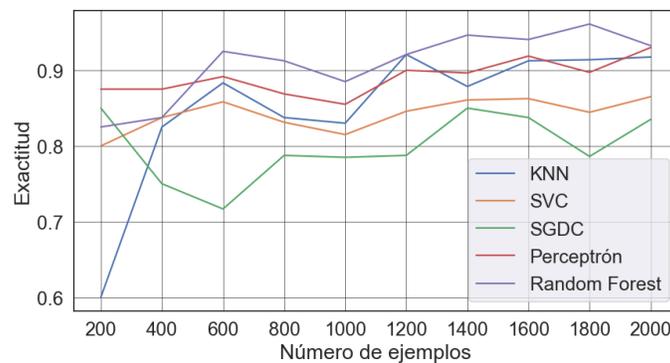


Figura 3. Exactitud del conjunto de pruebas variando el número de ejemplos, usando MFCC.

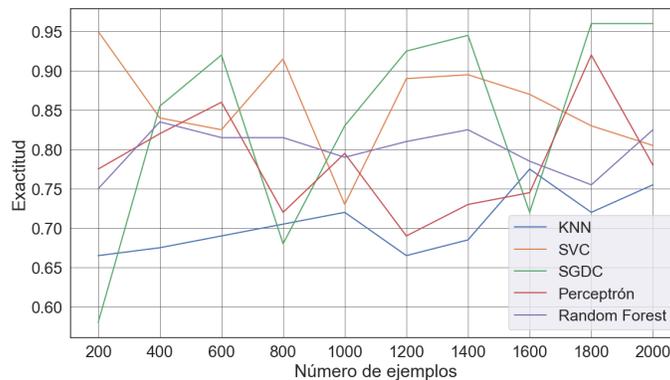


Figura 4. Exactitud del conjunto de predicción variando el número de ejemplos, usando espectrograma de Mel.

Los resultados obtenidos de los datos de predicción usando MFCC se observan en la Figura 5. Comparándolos con los resultados con espectrograma de Mel, no consiguen en su punto máximo acercarse al modelo con mayor exactitud usando espectrograma de Mel, 76% con perceptrón y 400 ejemplos vs 96% con SGDC y 2,000 ejemplos.

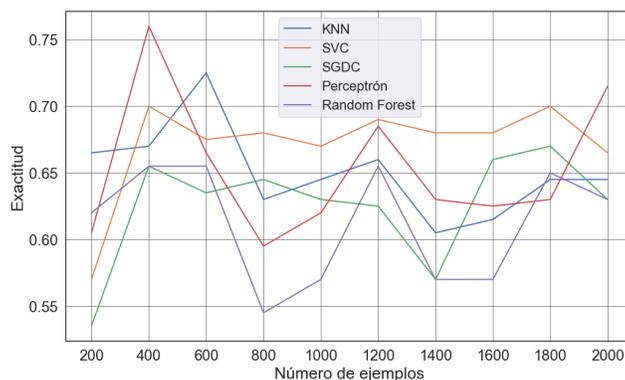


Figura 5. Exactitud del conjunto de predicción variando el número de ejemplos, usando MFCC.

## 5 Conclusiones y trabajo a futuro

Los modelos entrenados con el dataset de TEDx y espectrograma de Mel obtuvieron una mayor exactitud de manera general a comparación de los creados con MFCC, siendo SVC quién alcanzó la mayor exactitud con el mínimo número de ejemplos. Mientras que los modelos entrenados con el dataset de TEDx y MFCC mostraron una tendencia de aumentar la exactitud a medida que aumenta el número de ejemplos durante el entrenamiento. Random Forest alcanzó la mayor exactitud y se observó una diferencia entre la exactitud alcanzada por SVC y SGDC y los modelos restantes.

Durante la predicción con nuevos datos (conjunto de predicción) se observó un contraste con los resultados de las pruebas. Siendo lo más común que el rendimiento de los modelos alcanzaran menor exactitud. Con los modelos con espectrograma de Mel, el modelo que mostró la mejor exactitud fue SGDC con 96% mientras que los demás no consiguen una exactitud cercana. Los modelos MFCC obtuvieron menor exactitud que los creados empleando espectrograma. Siendo el modelo perceptrón con MFCC el que mejor se comporta con 76% de exactitud. Debe considerarse que para llevar a cabo las predicciones se utilizaron audios de hablantes de diferentes países, lo que implica acentos diferentes y a pesar de ello las predicciones de todos los modelos estuvieron por encima del azar.

A pesar de que los modelos entrenados con MFCC mostraron buenos resultados durante las pruebas con el conjunto de TEDx, no pudieron igualar el rendimiento de los modelos creados con espectrograma de Mel. Sugiriendo que los espectrogramas de Mel captan de mejor manera patrones importantes al momento de clasificar hombres y mujeres. Y no menos importante es la elección del algoritmo de clasificación adecuado.

Aunque los resultados obtenidos no pueden compararse con los presentados en el estado del arte, porque utilizan diferentes corpus, características, clasificadores e idiomas, podemos decir que la exactitud obtenida está dentro del rango obtenido en estos trabajos.

Como trabajo futuro se planea utilizar otros algoritmos de clasificación como redes neuronales recurrentes o convolucionales y experimentar con algoritmos metaheurísticos para optimizar los parámetros de los algoritmos de aprendizaje utilizados.

## Referencias

[1] Belin, P., Bestelmeyer, P. E. G., Latinus, M., y Watson, R., "Understanding voice perception", in *British Journal of Psychology*, 102, 2011, pp. 711-725.

[2] Rani, Poonam, and Bhupender Yadav, "A Survey on Gender and Emotion Recognition Using Voice", in *International Journal of Recent Research Aspects* 5.4 2018.

[3] Wei Han, Cheong-Fat Chan, Chiu-Sing Choy and Kong-Pang Pun, "An efficient MFCC extraction method in speech recognition", in *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2006, doi: 10.1109/ISCAS.2006.1692543. <https://ieeexplore.ieee.org/document/1692543>

[4] Levitan, Sarah Ita, Taniya Mishra, and Srinivas Bangalore, "Automatic identification of gender from speech", in *Proceeding of speech prosody*. Semantic Scholar, 2016.

[5] M. G. J. Abella, J. M. L. Borja, A. C. F. Zuñiga and J. C. De Goma, "Analyzing Gender Detection in Speech in Adult Filipino Citizens," 2022 12th International Conference on Software Technology and Engineering (ICSTE), Osaka, Japan, 2022, pp. 73-81, doi: 10.1109/ICSTE57415.2022.00018.

[6] K. Nugroho, E. Noersasongko, Purwanto, Muljono and H. A. Santoso, "Javanese Gender Speech Recognition Using Deep Learning And Singular Value Decomposition," in 2019 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, 2019, pp. 251-254, doi: 10.1109/ISEMANTIC.2019.8884267.

[7] Guevara-Rukoz, A., Demirsahin, I., He, F., Chu, S. H. C., Sarin, S., Pipatsrisawat, K., ... & Kjartansson, O. (2020, May). Crowdsourcing Latin American Spanish for low-resource text-to-speech. In *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 6504-6513).

[8] A. Patle and D. S. Chouhan, "SVM kernel functions for classification," *IEEE Xplore*, Jan. 01, 2013. <https://ieeexplore.ieee.org/abstract/document/6524743>

[9] sklearn.neighbors — Nearest Neighbors, Scikit-learn. [Online]. Available: <https://scikit-learn.org/stable/modules/neighbors.html#classification>

[10] sklearn.svm.LinearSVC, Scikit-learn. [Online]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.svm.LinearSVC.html#sklearn.svm.LinearSVC>

[11] Popescu, M. C., Balas, V. E., Perescu-Popescu, L., & Mastorakis, N. (2009). Multilayer perceptron and neural networks. *WSEAS Transactions on Circuits and Systems*, 8(7), 579-588.

[12] sklearn.neural\_network.MLPClassifier, Scikit-learn. [Online]. Available: [https://scikit-learn.org/stable/modules/generated/sklearn.neural\\_network.MLPClassifier.html#sklearn.neural\\_network.MLPClassifier](https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPClassifier.html#sklearn.neural_network.MLPClassifier)

[13] sklearn.tree — Decision Trees, Scikit-learn. [Online]. Available: <https://scikit-learn.org/stable/modules/tree.html#tree>

[14] sklearn.ensemble.RandomForestClassifier, Scikit-learn. [Online]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>

[15] Hernandez-Mena, C. D. (2019). TEDx Spanish Corpus. Audio and transcripts in Spanish taken from the TEDx Talks; shared under the CC BY-NC-ND 4.0 license. Mexico City: Universidad Nacional Autonoma de Mexico.

[16] librosa.feature.mfcc, Librosa. [Online]. Available: <https://librosa.org/doc/main/generated/librosa.feature.mfcc.html>

[17] Stochastic Gradient Descent, Scikit-learn. [Online]. Available: <https://scikit-learn.org/stable/modules/sgd.html#sgd>

[18] librosa.feature.melspectrogram, Librosa. [Online]. Available: <https://librosa.org/doc/main/generated/librosa.feature.melspectrogram.html>

## **Prototipo para la detección de vehículos y motocicletas en puntos ciegos utilizando el modelo YOLO**

### **Prototype for vehicle and motorcycle detection in blind spots using the YOLO model**

Abraham Isaí Cantú Sánchez, José Guadalupe Martínez Herrera, Aldo de Jesus Peña  
Cuellar, Jesús David Treviño Gandarilla, Marco Aurelio Nuño Maganda

Universidad Politécnica de Victoria, Av. Nuevas Tecnologías No. 5902 Carretera Victoria-  
Soto La Marina Km. 55, Parque Científico y Te, 87138 Ciudad Victoria, Tamaulipas

2130328@upv.edu.mx

#### **Abstract**

This project develops and evaluates a prototype for detecting vehicles and motorcycles, designed to identify objects in the blind spot of a car using videos captured from within the vehicle under real traffic conditions. The system employs YOLO (You Only Look Once) for real-time object detection. Although initial results show high accuracy (98%) with the COCO dataset, performance is more limited (45%) with our own dataset, which contains fewer images. Furthermore, under dense traffic and adverse weather conditions, accuracy significantly drops to 25%. The development involves capturing training and testing videos, as well as implementing a user-friendly interface using PyQt6. As the project progresses, improvements are anticipated to enhance performance in more complex scenarios. The ultimate goal is to improve road safety by providing a tool that, once optimized, can detect vehicles in the driver's blind spot.

#### **Resumen**

Este proyecto desarrolla y evalúa un prototipo de detector de vehículos y motocicletas, diseñado para identificar objetos en el punto ciego de un automóvil utilizando videos capturados desde el interior del vehículo en condiciones de tráfico real. El sistema utiliza YOLO (You Only Look Once) para la detección de objetos en tiempo real. A pesar de que los resultados iniciales muestran una alta precisión (98%) al utilizar el dataset COCO, el desempeño es más limitado (45%) con nuestro propio dataset, que contiene menos imágenes. Además, en condiciones de tráfico denso y meteorología adversa, la precisión disminuyó significativamente (25%). El desarrollo del sistema incluye la captura de videos de entrenamiento y prueba, así como la implementación de una interfaz en PyQt6 para una experiencia de usuario amigable. A medida que se avanza en el proyecto, se esperan mejoras que optimicen su rendimiento en escenarios más complejos. El objetivo final es

mejorar la seguridad vial proporcionando una herramienta que, una vez optimizada, podrá detectar vehículos en el punto ciego del conductor.

---

Keywords: YOLO, PyQt6 , Detección de Objetos, Software, Segmentación de Video, Modelos Preentrenados.

---

## **1 Introducción**

Hoy en día la seguridad vial sigue siendo un desafío global, especialmente en lo que respecta a la prevención de accidentes causados por la falta de visibilidad en los puntos ciegos de los vehículos. A pesar de los avances en tecnologías automotrices, este problema persiste como una de las principales causas de colisiones en carreteras. Actualmente, existen soluciones a estos problemas como los espejos convexos, sensores de proximidad y cámaras, pero a menudo resultan insuficientes o imprecisas, especialmente en situaciones de tráfico denso o condiciones adversas.

El proyecto aborda este vacío mediante el desarrollo de un sistema más preciso y eficiente de detección de vehículos y motocicletas en los puntos ciegos, utilizando algoritmos avanzados de visión por computadora y aprendizaje automático. Al emplear el algoritmo YOLO, conocido por su capacidad de detección de objetos en tiempo real, se busca superar las limitaciones de las tecnologías actuales, brindando una solución más robusta y adaptada a las necesidades del conductor moderno.

La justificación de este desarrollo radica en la necesidad de mejorar las tasas de detección, reduciendo el tiempo de respuesta y la precisión en la identificación de vehículos y motocicletas, lo que contribuirá a la prevención de accidentes y aumentará la seguridad en las carreteras.

## **2 Prueba de software**

La prueba del software se divide en varias fases clave que incluyen la recopilación de datos, el procesamiento, el entrenamiento del modelo, la implementación y la evaluación del sistema.

### **2.1 Recopilación de datos**

La primera fase implicó la recopilación de vídeos de tráfico tomados desde el interior de un automóvil en diversas condiciones de tráfico. Estos videos reflejaban situaciones comunes en donde los vehículos y las motocicletas podían estar en el punto ciego. Para complementar la calidad de los datos capturados, también se utilizaron recursos de internet, especialmente cuando hubo problemas de enfoque o calidad de imagen. Se prioriza la recolección de videos que incluyeran tanto vehículos como motocicletas, capturando un total de 139 imágenes.

### **2.2 Procesamiento de datos**

Para optimizar el rendimiento del modelo, se realizó un etiquetado manual de los cuadros extraídos de los videos capturados. Esto incluyó la identificación de la presencia de vehículos, motocicletas, y en algunos casos, sus marcas y modelos. El proceso de etiquetado se realizó utilizando la herramienta LabelMe, que permitió generar anotaciones precisas para el entrenamiento del modelo.

Una vez capturados los videos, se realizó una segmentación en cuadros individuales para facilitar el análisis. Las imágenes resultantes fueron sometidas a un proceso de normalización, ajustando la iluminación y la resolución para asegurar la consistencia entre los datos de entrada. Este preprocesamiento fue fundamental para mejorar la precisión en la detección.

### 2.3 Selección y ajuste de modelos

Para la detección de objetos, se eligió el algoritmo YOLO por su velocidad y eficacia en la detección en tiempo real. Se exploraron dos versiones:

- YOLO v3: Ofreció un buen balance entre precisión y velocidad.
- YOLO v8: Una versión más reciente que mejoró la precisión y la eficiencia computacional.

### 2.4 Entrenamiento del modelo

El núcleo del sistema es un modelo YOLO entrenado usando tanto modelos pre-entrenados como datos específicos recopilados para motocicletas. El proceso de entrenamiento incluyó:

- Carga de modelos pre-entrenados: Se utilizaron pesos pre-entrenados para las versiones seleccionadas de YOLO, lo que redujo el tiempo de entrenamiento.
- Entrenamiento personalizado: Se realizó un entrenamiento adicional específico para la detección de motocicletas usando el dataset etiquetado manualmente de 139 imágenes. Sin embargo, este tamaño de dataset mostró ser insuficiente, por lo que se recurrió al uso del dataset COCO, que contiene más de 330,000 imágenes y alrededor de 200,000 archivos etiquetados, lo que mejoró considerablemente la precisión del modelo.
- Optimización y validación: Se aplicaron técnicas de optimización de hiper-parámetros y validación cruzada para asegurar la robustez del modelo.

### 2.5 Implementación del sistema

Una vez entrenado, el modelo fue integrado en un prototipo que procesa videos en tiempo real. La interfaz de usuario fue desarrollada utilizando PyQt6, proporcionando una experiencia amigable para el conductor. El sistema, diseñado para integrarse en el tablero del automóvil, emitía alertas visuales y auditivas cuando se detectaban vehículos o motocicletas en el punto ciego.

El sistema fue implementado en un MacBook Pro M2, lo que permitió un procesamiento eficiente de los cuadros de video y la generación de alertas en tiempo real.

### 2.6 Evaluación del sistema

La evaluación del sistema se realizó bajo diversas condiciones de tráfico y climatológicas. Los criterios de evaluación incluyeron:

- Precisión de detección: El sistema alcanzó una precisión del 98% en la detección de vehículos y motocicletas utilizando el dataset COCO. Al utilizar únicamente el dataset personalizado de 139 imágenes, la precisión disminuyó a 45%.
- Velocidad de procesamiento: El tiempo de procesamiento fue adecuado para el uso en tiempo real, aunque se planea refinar la medición de los tiempos exactos en futuros experimentos. Actualmente, se estima que el procesamiento de cada cuadro ocurre dentro de un rango aceptable de milisegundos.
- Robustez: Aunque el sistema fue robusto en la mayoría de las condiciones, la precisión cayó a 25% en situaciones de tráfico denso y baja iluminación.

A pesar de estos desafíos, los resultados generales fueron prometedores, demostrando el potencial de este sistema para mejorar la seguridad vial en situaciones de tráfico cotidiano.

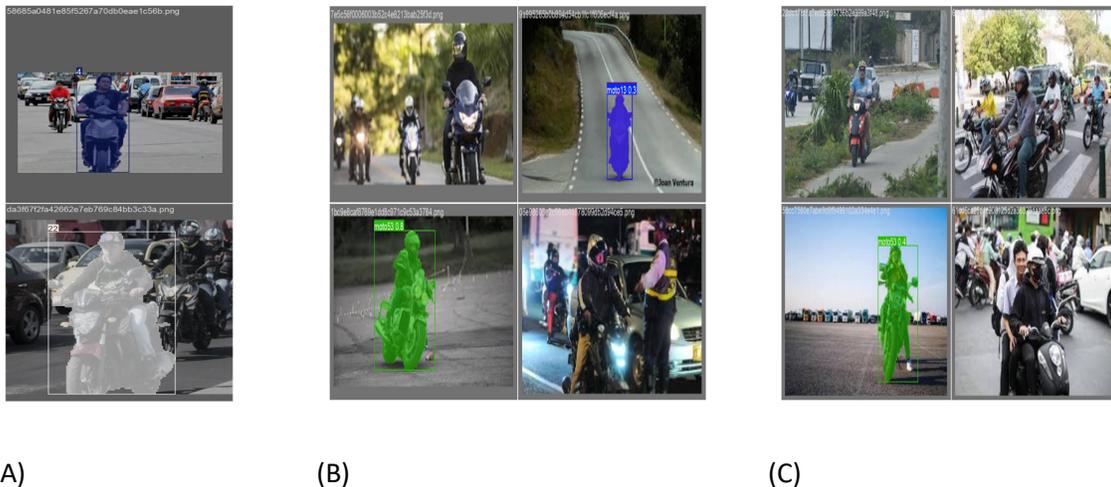


Figura 1. Resultado en la detección de motocicletas (dataset hecho por nosotros).

Figura 1(A). En esta imagen se presenta un ejemplo de la detección de motocicletas. Sin embargo, la detección no es completamente efectiva, ya que se puede observar que algunas motocicletas no fueron detectadas correctamente.

Figura 1(B). Aquí se muestra otro ejemplo de la detección de motocicletas. En dos de los casos se logró identificar correctamente el modelo de la motocicleta, pero en otros dos casos la detección falló.

Figura 1(C). Esta imagen ilustra un caso donde la detección no fue particularmente efectiva. De cuatro intentos, sólo en uno se detectó correctamente la motocicleta, mientras que en los otros tres casos la detección falló.

El entrenamiento del modelo mejoró eficazmente logrando detectar cada modelo en diferentes ambientes como autopista, carretera, entre otros.

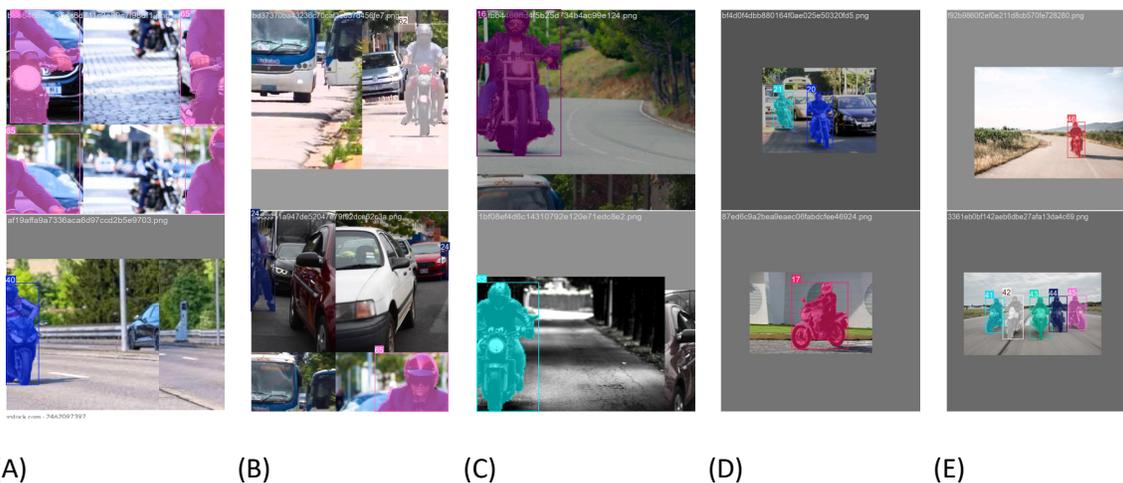


Figura 2. Detección de motocicletas en distintos entornos.

Figura 2(A). Esta subfigura muestra la detección exitosa de todas las motocicletas cercanas al vehículo.

Figura 2(B). Aquí se observa cómo la aplicación logra detectar cada una de las motocicletas de manera individual.

Figura 2(C). Se presenta la detección de motocicletas en una carretera, mostrando variaciones en comparación con un entorno de autopista.

Figura 2(D). En esta subfigura se destaca la eficacia de la aplicación al detectar varias motocicletas simultáneamente.

Figura 2(E). Finalmente, se muestra que la aplicación detecta múltiples motocicletas al mismo tiempo con éxito, comprobando su rendimiento.

Fuente del dataset: [COCO - Common Objects in Context \(cocodataset.org\)](https://cocodataset.org)

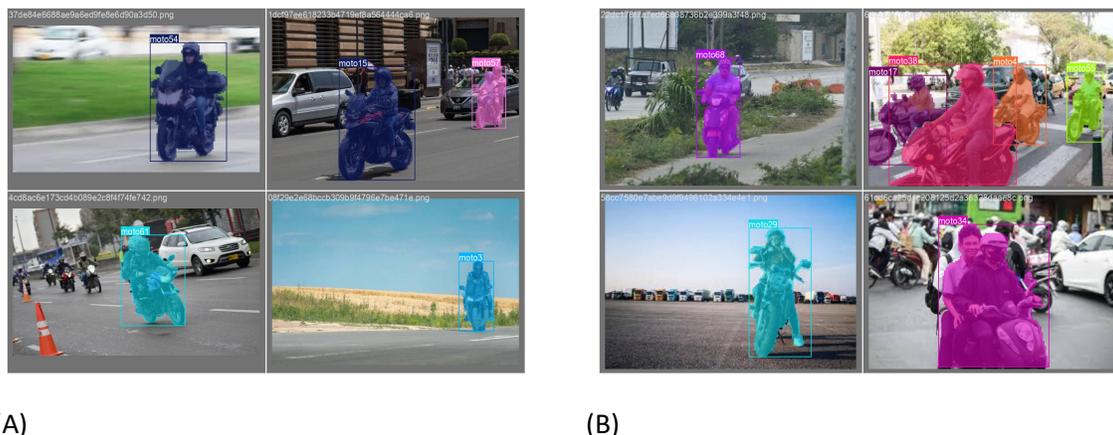


Figura 3. Detección eficaz de motocicletas cercanas.

Figura 3(A). En esta subfigura se observa cómo la aplicación detecta eficazmente todas las motocicletas cercanas sin ningún inconveniente, utilizando el dataset COCO como base para el entrenamiento del modelo de detección.

Figura 3(B). De manera similar, esta subfigura muestra la detección precisa de todas las motocicletas en el área cercana, reforzando la efectividad de la aplicación. El dataset COCO fue empleado para mejorar la precisión de las detecciones.

Fuente del dataset: [COCO - Common Objects in Context \(cocodataset.org\)](https://cocodataset.org)

### 3. Metodología

En las fases iniciales, se probó el clasificador Cascade obtenido de fuentes en línea, pero mostró baja efectividad en la clasificación de objetos. Posteriormente, se realizaron pruebas con otros datasets proporcionados por colaboradores, aunque los resultados fueron inconsistentes. Esto motivó el desarrollo de modelos más complejos utilizando YOLO y la herramienta LabelMe para la anotación de imágenes.

Dado que el conjunto inicial de 139 imágenes resultó insuficiente para un entrenamiento efectivo, se optó por utilizar el conjunto de datos COCO, con más de 330,000 imágenes y aproximadamente 200,000 archivos. Este recurso fue invaluable para el desarrollo de un modelo más preciso y eficiente.

A pesar de la eficiencia del modelo YOLO, el tiempo de entrenamiento fue largo, llegando a dos horas para procesar únicamente las 139 imágenes iniciales de motocicletas. Esto resaltó la necesidad de un dataset más robusto, lo que justificó la adopción del dataset COCO para mejorar los tiempos de entrenamiento y la precisión del modelo.

#### 4. Resultados

El sistema demostró una alta precisión en la detección tanto de vehículos como de motocicletas bajo diversas condiciones de tráfico y luminosidad. Durante las pruebas, se observó que el modelo logró identificar correctamente la mayoría de los vehículos y motocicletas en el punto ciego del automóvil, proporcionando alertas visuales y auditivas de manera oportuna. La implementación de YOLO v3 y YOLO v8 permitió mantener un balance óptimo entre velocidad y precisión, adaptándose eficazmente a las variaciones en el entorno de prueba [10].

En el caso de las motocicletas, a pesar de las variaciones en tamaño y forma, el modelo logró identificarlas con una precisión considerable. Sin embargo, se notaron algunos fallos en situaciones de tráfico más complejo o en condiciones de baja iluminación. Estos resultados sugieren que, aunque el sistema es efectivo, aún hay espacio para mejoras en la detección de motocicletas en escenarios más desafiantes.



(A)

(B)

Figura 4. Sistema de detección y alerta de colisión.

Figura 4(A). El sistema detecta tanto vehículos como motocicletas, y muestra un mensaje de alerta de "ALERTA DE COLISIÓN" cuando alguno de ellos se encuentra demasiado cerca.

Figura 4(B). En esta subfigura se presenta un caso en el que los vehículos y motocicletas están a una distancia segura, por lo que la aplicación no muestra ningún mensaje de alerta de colisión.

Fuente: [Los mejores 10.000+ videos de Trafico Automoviles · Banco de Imágenes Gratis · Videos de archivo de Pexels](#)

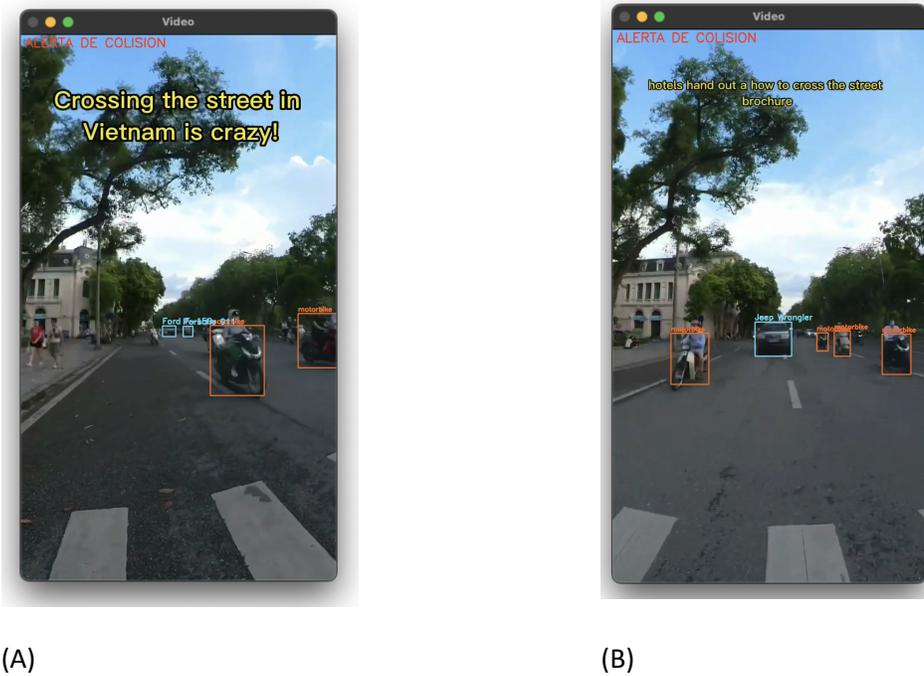


Figura 5. Alerta de colisión en casos de proximidad.

**Figura 5(A) y Figura 5(B).** En ambas subfiguras se muestran pruebas donde los vehículos o motocicletas se encuentran muy cerca del vehículo, lo que provoca que la aplicación emita un mensaje de advertencia de "ALERTA DE COLISION" debido a la proximidad y el riesgo de colisión.

Fuente: [Crossing the street in Vietnam is crazy! 🤪 \(youtube.com\)](https://www.youtube.com/watch?v=...)

## Diagrama de clases



A continuación se explican las principales clases de la aplicación:

Clase Classifier en classifier.py

### *Descripción General*

La clase Classifier se encarga de cargar un modelo de Tensor Flow y usarlo para predecir la marca y modelo de autos en imágenes. Utiliza las configuraciones especificadas en el archivo config.py.

### Atributos

- self.model: El modelo de TensorFlow cargado desde el archivo especificado en config.model\_file.
- self.labels: Lista de etiquetas de clases cargadas desde el archivo especificado en config.label\_file.
- self.input\_layer: Nombre de la capa de entrada del modelo.
- self.output\_layer: Nombre de la capa de salida del modelo.

### Métodos

- \_\_init\_\_(self): Constructor que inicializa el modelo y las etiquetas.
  - Carga el modelo de TensorFlow.
  - Carga las etiquetas de clase.
  - Configura los nombres de las capas de entrada y salida.
- preprocess(self, image): Método que preprocesa la imagen para que sea compatible con el modelo.
  - Redimensiona la imagen al tamaño especificado en config.classifier\_input\_size.
  - Convierte la imagen a un tensor de TensorFlow.
- predict(self, image): Método que realiza la predicción sobre la imagen preprocesada.
  - Preprocesa la imagen.
  - Ejecuta el modelo en la imagen.
  - Interpreta las salidas del modelo para obtener la predicción de marca y modelo del auto.

### Configuración en config.py

#### Descripción General

Este archivo contiene las configuraciones necesarias para la clase `Classifier`, como rutas a archivos de modelo y etiquetas, y nombres de capas del modelo.

#### Atributos

- model\_file: Ruta al archivo del modelo de TensorFlow (model/trained\_model.pb).
- label\_file: Ruta al archivo de etiquetas (model/labels.txt).
- input\_layer: Nombre de la capa de entrada del modelo (input\_1).
- output\_layer: Nombre de la capa de salida del modelo (softmax/Softmax).
- classifier\_input\_size: Tamaño de entrada esperado por el clasificador ((128, 128)).

### Uso del Clasificador en main.py

#### Descripción General

El archivo main.py demuestra cómo se utiliza la clase `Classifier` junto con un modelo de detección de objetos YOLO para procesar un video, detectar objetos, y clasificar autos por marca y modelo.

### Proceso General

1. Inicialización:
  - Importa los módulos necesarios (numpy, time, cv2, os, y classifier).
  - Carga el clasificador (car\_color\_classifier = classifier.Classifier()).
  - Carga etiquetas de COCO para el modelo YOLO.
  - Inicializa colores para las etiquetas.
  - Carga los archivos de configuración y pesos de YOLO.
  - Configura el umbral de confianza y supresión de no máximos.
2. Procesamiento de Video:
  - Abre un archivo de video.
  - Itera sobre los fotogramas del video.
  - Obtiene las capas de salida del modelo YOLO.
  - Crea un blob de imagen y lo pasa por la red YOLO.
  - Recoge las coordenadas, confianzas, e IDs de clase de las detecciones.
  - Aplica supresión de no máximos para eliminar detecciones redundantes.
  - Anota los fotogramas con cajas delimitadoras y etiquetas.
3. Clasificación y Detección de Colisiones:
  - Si se detecta un auto, utiliza el clasificador para predecir la marca y modelo del auto.
  - Dibuja las etiquetas y cajas delimitadoras en el fotograma.
  - Detecta posibles colisiones basadas en el área de la caja delimitadora.
  - Muestra el fotograma anotado en una ventana.

## 5 Conclusión

El desarrollo del detector de vehículos y motos ha demostrado ser un avance significativo en la mejora de la seguridad vial. Al identificar de manera precisa y oportuna los vehículos en el punto ciego, el sistema proporciona una herramienta valiosa para los conductores, reduciendo así el riesgo de accidentes. La capacidad del sistema para reconocer vehículos específicos y motocicletas en tiempo real subraya la efectividad de las tecnologías de visión por computadora y aprendizaje automático aplicadas en este contexto.

A través de la implementación de este proyecto, se ha logrado demostrar el potencial de los sistemas de detección automatizados para contribuir a un entorno de conducción más seguro. Los resultados obtenidos indican que, con una continua optimización y refinamiento, este tipo de tecnologías podría convertirse en un componente estándar en los vehículos modernos, ofreciendo una capa adicional de protección para los conductores y otros usuarios de la vía .

## Referencias

- [1] Pandas development team, "pandas: Powerful Python data analysis toolkit," pandas documentation, [Online]. Disponible en: [\url{https://pandas.pydata.org/}](https://pandas.pydata.org/). [Consultado: 09-Jul-2024].
- [2] NumPy developers, "NumPy: the fundamental package for scientific computing with Python," NumPy documentation, [Online]. Disponible en: [\url{https://numpy.org/}](https://numpy.org/). [Consultado: 09-Jul-2024].
- [3] Matplotlib development team, "Matplotlib: Visualization with Python," Matplotlib documentation, [Online]. Disponible en: [\url{https://matplotlib.org/}](https://matplotlib.org/). [Consultado: 09-Jul-2024].
- [4] Seaborn development team, "Seaborn: Statistical data visualization," Seaborn documentation, [Online]. Disponible en: [\url{https://seaborn.pydata.org/}](https://seaborn.pydata.org/). [Consultado: 09-Jul-2024].
- [5] W. McKinney, *\textit{Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython}*, O'Reilly Media, 2012.
- [6] scikit-learn development team, "scikit-learn: Machine Learning in Python," scikit-learn documentation, [Online]. Disponible en: [\url{https://scikit-learn.org/}](https://scikit-learn.org/). [Consultado: 09-Jul-2024].
- [7] Kaggle Inc, "Kaggle: Your Home for Data Science," Kaggle documentation, [Online]. Disponible en: [\url{https://www.kaggle.com/}](https://www.kaggle.com/), junio de 2023. [Consultado: 09-Jul-2024].
- [8] J. VanderPlas, *\textit{Python Data Science Handbook: Essential Tools for Working with Data}*, O'Reilly Media, 2016.
- [9] M. Zhao, D. Lee, S. Kim, y otros, "Using Python for Data Analysis in Research: A Case Study," *Data Science Journal*, vol. 19, no. 5, pp. 45-56, 2020.
- [10] A. Geron, *\textit{Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems}*, O'Reilly Media, 2019.

## **Uso de diccionarios en Python e interpolación para conocer las tendencias en X sobre el aumento de feminicidios en México**

## **Use of dictionaries in Python and interpolation to understand trends on X regarding the increase in femicides in Mexico**

Maria Beatriz Bernábe Loranca<sup>1</sup>, Melissa Mendoza Bernábe<sup>2</sup>, Rogelio González Velázquez<sup>1</sup>

<sup>1</sup>Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación,  
Puebla, México,

<sup>2</sup>Universidad Iberoamericana Puebla, México.

beatriz.bernabe@gmail.com, melrosa88@gmail.com, rogelio.gzzvzz@gmail.com

### **Abstract**

The news circulating on social media has led to increased concerns about violence against women, particularly in relation to femicides in Mexico. This has prompted the need to review people's perceptions on social media to analyze how opinions spread and go viral. In this context, the use of dictionary creation techniques applied to content analysis on X (formerly Twitter) allows for the study of the relationships between the words used to refer to femicides and the thoughts of users.

This article proposes an initiative focused on creating dictionaries to structure comments on X and, consequently, organize keywords related to femicides in order to quantify the problem. Additionally, linear interpolation based on data from INEGI has been used to make forecasts.

### **Resumen**

Las noticias que circulan en redes sociales han provocado un aumento de preocupaciones sobre la violencia contra las mujeres, esencialmente en relación con los feminicidios en México. Esto ha motivado la necesidad de revisar las percepciones de las personas de los usuarios en redes sociales con el fin de analizar la manera en que las opiniones se propagan y viralizan. En este contexto, el uso de técnicas de creación de diccionarios aplicados al análisis de contenido en X (anteriormente Twitter) permite estudiar las relaciones entre las palabras utilizadas para referirse a los feminicidios con los pensamientos de los usuarios.

Este artículo propone una iniciativa centrada en la creación de diccionarios para estructurar comentarios en X y consecuentemente, organizar palabras clave relacionadas con los feminicidios para dimensionar el problema. Asimismo, se ha utilizado una interpolación lineal basada en datos del INEGI para realizar pronósticos.

---

Keywords and phrases: Diccionarios, Femicidios, Interpolación, Redes Sociales, Python

---

## 1 Introducción

El término "femicidio" fue introducido por la académica estadounidense Diana E.H. Russell, quien lo definió como "el asesinato de mujeres por el simple hecho de ser mujeres" [1]. En México, los homicidios de mujeres en Ciudad Juárez, Chihuahua, se denuncian a partir de en 1993, lo cual adquirió una importancia tanto a nivel nacional como internacional. Veinte años después, en México se reconoció el femicidio como delito dentro del código penal [2, 3].

A partir de este contexto, se plantea un ejercicio computacional que utiliza la creación de diccionarios para analizar las opiniones de los usuarios en X. Este enfoque tiene como objetivo generar datos e información que permitan entender la percepción pública sobre los femicidios en México en los últimos años. Al enmarcar el problema dentro de una red social, es posible aplicar métodos estadísticos, como la interpolación lineal que ayuda a identificar tendencias y patrones en las discusiones.

El trabajo se encuentra estructurado de la siguiente manera: La sección 1 se ocupa de la presente introducción. En el apartado 2, se exponen los aspectos preliminares. La sección 3 se ocupa de presentar la metodología de extracción de información en X. Con el fin de analizar los datos del INEGI y de señalar los diccionarios creados, la sección 4 se encarga de describir el conteo de las palabras implícitas en los diccionarios y además se expone un cálculo estadístico estimativo para los años 2019-2024 sobre femicidios.

## 2 Preliminares

Los femicidios son definidos como el asesinato de mujeres por razones de género, enraizados en dinámicas de poder propias de un sistema patriarcal [4]. El término "femicidio" se adapta a diferentes contextos culturales y sociales, lo que implica variaciones en su interpretación, expresión y legislación. Aunque el concepto mantiene una idea central, su aplicación cambia considerablemente según el entorno.

El término "violencia de género" ha conseguido importancia sustancial en México, especialmente ante el supuesto incremento de casos reportados en los últimos años. En 1993, el decreto sobre la eliminación de la violencia contra la mujer de la Organización de las Naciones Unidas (ONU), definió la violencia contra la mujer como "todo acto de violencia basado en el género que tenga como consecuencia posible o real un daño físico, sexual o psicológico, incluidas las amenazas, la coerción o la privación arbitraria de la libertad sin importar que ocurra en la esfera pública o privada" [5]. Por otra parte, en el convenio del Consejo de Europa sobre la prevención y combate contra la violencia

hacia las mujeres, las palabras de "violencia de género" significan en términos amplios "violencia ejercida contra una mujer por el hecho de serlo" [6, 7].

La erradicación de la violencia de género ha impactado y generado reflexión de distintas maneras en las últimas décadas. A pesar de que las iniciativas para combatir la violencia hacia la mujer se asentaron en los años 90, el concepto de "dignidad" emerge como un componente central en los derechos de las mujeres y, particularmente, en la lucha contra los feminicidios [8].

### 2.1 Sobre el procesamiento en lenguaje natural, redes sociales y Python

El lenguaje natural (LN) es el medio utilizado por los seres humanos para comunicarse cotidianamente [9]. Simbólicamente parece sencillo y es lo opuesto ya que el proceso de comunicación involucra herramientas tanto intuitivas como metodológicas y automatizadas. El Procesamiento del Lenguaje Natural (PLN) deriva de las ciencias computacionales cuya función principal es el diseño e implementación de software y hardware destinados al análisis computacional del LN, de este modo, se logra que las máquinas comprendan este tipo de lenguaje [10]. En los últimos años, el PLN ha avanzado gracias al crecimiento exponencial del volumen de datos y al desarrollo de algoritmos más eficientes para su procesamiento.

Las diversas plataformas de relación entre usuarios, quienes comparten, difunden y transforman información, son servicios basados en web [11]. La gran cantidad de información y usuarios involucrados en las redes sociales impacta en el pensamiento de diversas situaciones en distintos contextos que son compartidos por la red. Específicamente, sobre la interacción en X, es posible que cualquier usuario web que pertenezca a esta red social comparta cualquier opinión en público, así como comunicarse de forma privada. Esta plataforma ofrece características que facilitan el análisis y la extracción de información, lo que la convierte en un recurso valioso para aplicar técnicas de PLN, utilizando la programación en Python [12].

## 3 Descubrimiento de información y acercamiento a la solución del problema

La estructura del problema que se presenta en este documento propone una secuencia de pasos para obtener una implementación razonable. Estos pasos incluyen desde la identificación de la información relevante hasta la obtención de conclusiones, una secuencia aplicable a diversos problemas (ver figura 1).

### 3.1 Palabras clave y hashtags

El primer paso en la extracción de datos es la caracterización de palabras clave que representen el problema, de este modo, se agiliza las descargas de las opiniones de los usuarios y además se conserva una coherencia de la situación. Un enfoque pragmático es rastrear cuentas de personas influyentes asociadas con el problema de los feminicidios. Tras identificar dichas cuentas, se seleccionaron palabras representativas que se buscan a través de hashtags, como #NiU-naMas,

#NiUnaMenos, #Feminicidio, #VivasNosQueremos, Acoso sexual, Movimiento feminista, Asesinato machista, Feminicida.

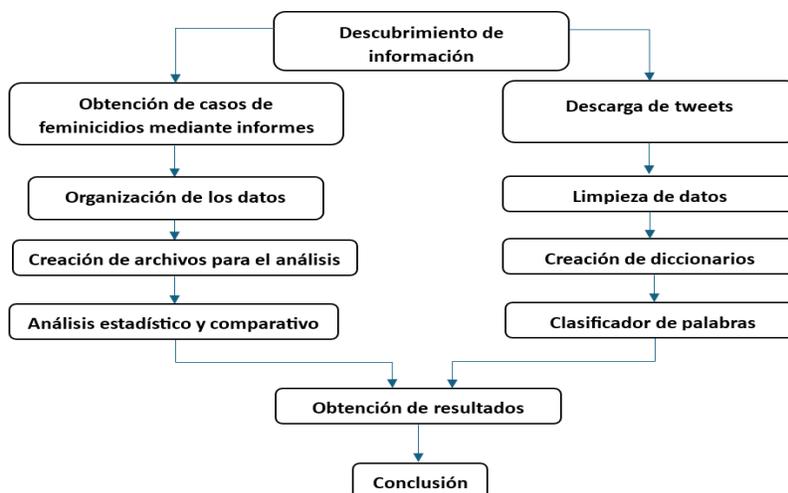


Figura 1. Arquitectura del problema de feminicidios

### 3.2 Extracción de datos con Vicinitas

Las descargas de las opiniones (Tweets), inicialmente se hicieron durante 10 días consecutivos seguido de descargas semanales. Se empleó la herramienta web Vicinitas que facilita la recopilación de datos importantes sobre los tweets. Las palabras clave reconocidas incluyeron #NiUnaMas, #NiUnaMenos, #Feminicidio, #VivasNosQueremos, Acoso sexual, Movimiento feminista, Asesinato machista y Feminicida. Posteriormente, se aplicaron procedimientos sencillos de filtros para limpiar los datos extraídos. Para una revisión no subjetiva, dos personas auxiliares examinaron y seleccionaron los tweets relevantes (excluyendo re-tweets y respuestas) y para concluir con cualquier desacuerdo, se recurrió a la intervención de un tercer revisor independiente. Este proceso es muy importante porque de esta elección de tweets depende la calidad de los diccionarios

Para acceder a la API de X, se utilizó la biblioteca Tweepy de Python junto con las credenciales de acceso obtenidas mediante el registro como desarrollador en la plataforma. El acceso a la API permite descargar datos basados en las palabras clave seleccionadas.

### 3.3 Diseño de diccionarios y excepción de análisis de sentimientos

Para generar los diccionarios, se desarrolló un programa en Python que analiza los archivos CSV exportados desde Vicinitas registrando la aparición de palabras. El software se enfoca en términos que tienen cinco o más letras ya que muchas de las palabras clave relacionadas con los feminicidios cumplen con este criterio.

Se llevó a cabo un análisis de sentimientos (AS), sin embargo, los resultados no permitieron segmentar las palabras en categorías positivas o negativas debido a la predominancia de términos



Las palabras que se muestran en la figura 2 se generaron por las ventajas del programa de diccionarios. En la Tabla 1 se pueden ver las 10 palabras más comunes de la información recogida con su frecuencia asociada. La tabla 2 señala los vocablos que se identificaron en la segunda parte de descargas y creación de los respectivos diccionarios.

Tabla 1. Palabras con más frecuencia dentro de los diccionarios

PALABRA	FRECUENCIA	PALABRA	FRECUENCIA	PALABRA	FRECUENCIA
FEMINICIDIO	10089	DENUNCIA	3874	JUSTICIA	2816
FEMINICIDA	9283	CHICA	3836	REPORTAR	2813
VICTIMAS	9129	ASESINADA	3679	LUCHA	2715
MUJER	7538	PUEBLA	3383	911	2707
DEMASIADAS	5628	MACHISTA	3237	FORZÁNDOLA	2707
CRISIS	5393	CASOS	3114	ASESINO	2620
VIDA	5147	TODAS	3049	DENUNCIADO	2601
HOMBRES	4902	MÉXICO	2972	DENUNCIAS	2585
VICTIMA	4598	FEMINISTA	2881	CUARENTENA	2570
SOFIA	4242	HISTORIA	2828	FEMINICIDIOS	2493

Tabla 2. Diccionario extendido

PALABRA	FRECUENCIA	PALABRA	FRECUENCIA	PALABRA	FRECUENCIA
MARZO	2328	#FEMINICIDIOS	1918	AGRESOR	1711
ASESINATO	2278	VIOLACION	1901	FEMINISTA	1707
CADAVER	2252	#VIOLENCIAMACHISTA	1872	ABUSO	1607
#FEMINICIDIO	2213	IVANA	1776	MATO	1603
#VIOLENCIADEGENERO	2176	FAMILIA	1773	DERECHOS	1598
ASESINADAS	2171	MUCHAS	1772	MIEDO	1589
DENUNCIAR	2139	ESTADO	1765	HERMANA	1547
MENOR	2112	DETENIDO	1762	MADRE	1528
NIÑAS	2099	#FEMINICIDIOINFANTIL	1756	AGRESORES	1512
#TODASDUELEN	2024	NICOLE	1724	ULTIMA	1466

#### 4.2 Casos de feminicidios

Una vez creados los diccionarios, se intentó compararlos con los datos del mapa diseñado por María Salguero; el mapa intenta describir con informes evaluados, los feminicidios registrados por año en México [14]. En la figura 3 se pueden observar diversas categorías sin ser claras y además imprecisas, no existe una manera exacta de saber el número de casos de feminicidios de cada estado, por tanto, no ha sido posible correlacionar las palabras del estudio de Salguero con los diccionarios que conseguimos porque son trabajos excluyentes, aun así, se contabilizaron los casos registrados por

estado entre 2016 y 2019 del INEGI y se procedió a hacer una interpolación de 2016 a 2019 y una extrapolación entre 2019 y 2024 para calcular un pronóstico con Excel.



Figura 3. Mapa de feminicidios de María Salguero.

### 4.3 Casos de feminicidios con Python e interpolación

Se utilizó la biblioteca “pandas” de Python para reestructurar la información del INEGI dando como resultado la Tabla 3. En dicha tabla, se observan las cifras agrupadas con la variabilidad de feminicidios por año y estado. Este proceso es independiente de los diccionarios.

Los feminicidios tienen bajas cifras que se contradicen con la información de los medios y de las manifestaciones de los movimientos feministas. Si bien la respuesta no es nada fácil, se requiere de una investigación profunda y multidisciplinaria para acercarse a establecer conclusiones aproximadas del problema. Los feminicidios realmente reportan pocos casos debido a que, en caso de ser feminicidio, el acto no se etiqueta inmediatamente con ese nombre hasta que es demostrado.

Tabla 3. Feminicidios por estado en los años 2016 y 2019.

ESTADO	2016	2019	ESTADO	2016	2019
AGUASCALIENTES	8	17	MICHOACÁN	102	59
BAJA CALIFORNIA	86	93	MORELOS	82	37
BAJA CALIFORNIA SUR	14	10	NAYARIT	6	8
CAMPECHE	5	14	NUEVO LEÓN	68	68
COAHUILA	45	39	OAXACA	99	52
COLIMA	30	21	PUEBLA	97	67
CHIAPAS	52	19	QUERÉTARO	14	12
CHIHUAHUA	119	169	QUINTANA ROO	37	29
DISTRITO FEDERAL	2	75	SAN LUÍS POTOSÍ	27	39
DURANGO	10	16	SINALOA	85	42
GUANAJUATO	99	238	SONORA	36	66
GUERRERO	166	84	TABASCO	39	32
HIDALGO	36	29	TAMAULIPAS	43	24
JALISCO	68	72	TLAXCALA	5	18
MÉXICO	386	182	VERACRUZ	142	84

La Tabla 3 muestra la cantidad de feminicidios reportados en cada estado entre 2016 y 2019. En 2016 el estado de México tuvo más feminicidios, mientras que Campeche y Tlaxcala registraron 5. Del mismo modo para 2019 (en la tabla aparece el distrito federal por separado, y sumados sus casos al estado de México, se tuvieron 388 casos). Guanajuato en 2019 reporto 238 casos siendo así el estado con más feminicidios y Nayarit fue el estado con menos casos.

La tabla 3 ejemplifica un promedio de 65 feminicidios por estado en 2016 y en 2019 disminuyó a 55. En 2016 la cifra menor fue 2 y el máximo de 386. En 2019 los casos se encuentran en el intervalo [4,238]. Con esta información, es claro afirmar que en 2019 hubo menos feminicidios denunciados que se autorizaron con ese nombre. Estos resultados suponen que la baja de casos responde a que se divulgó en muchos lugares tanto televisivos como en las redes la denuncia formal de las diversas manifestaciones de violencia de género, entonces, se establece la conjetura de que la diversificación de las protestas provocó una alerta en las autoridades para atender casos de violencia, promover campañas de alerta y sensibilización para evitar en la medida posible un feminicidio. Por otra parte, se considera que tanto en campañas sobre conciencia poblacional, las ordenes de castigo y otras medidas expuestas en los medios, provocó que la información se extendiera, lo que permite asumir que se fomentó la prevención de feminicidios. En este punto, los diccionarios creados facilitarán la búsqueda de casos en las redes sociales en estudios posteriores.

En la tabla 4, se puede observar el cálculo de interpolación entre 2016-2019 y de extrapolación a partir de 2019.

La interpolación se hizo entre los años 2016 y 2019, mientras que la extrapolación se calculó de 2020 a 2024. Se observa una gran variabilidad en ambos casos, lo cual responde a que la extrapolación lineal tiene limitaciones cuando se proyectan muchos años hacia el futuro, por ejemplo, algunos estados mostraron una fuerte reducción entre 2016 y 2019, el modelo lineal tiende a proyectar una continuación de esa tendencia, lo que llevó a resultados negativos; para ello, se ajustó la ecuación porque no tiene sentido que haya feminicidios negativos a pesar de que se interpreta como una disminución. En la tabla 4, se observa una tendencia a la disminución en varios estados. La extrapolación apunta a una reducción en los feminicidios en los próximos años y la conjetura se centra en que podría estar vinculada esta baja a políticas de prevención o cambios sociales. Respecto a los picos de incremento, la extrapolación muestra un aumento (como Guanajuato), indicando que requieren más atención y recursos. En relación con la incertidumbre en el futuro, la extrapolación debe ser interpretada con cautela, especialmente para los años 2024 en adelante, donde el modelo puede no capturar adecuadamente las complejidades del fenómeno.

## **5 Conclusiones**

A partir del conteo de los casos de homicidios femeninos dados por el INEGI, se identificó que la violencia de género era común, al parecer crecía el número de homicidios, pero en años recientes se observa una disminución tanto en los datos que ofrece como en la interpolación y extrapolación desarrollada en este trabajo. Adicionalmente, el PLN, a través de la creación de diccionarios, ha

contribuido a encontrar información y tendencias en opiniones que pueden dar lugar a polarizar la situación de feminicidios generando conciencia sobre este problema. Las palabras contenidas en los diccionarios identifican el descubrimiento de información utilizando descargas de la red social como es el caso de este trabajo, con este procedimiento ha sido posible contar con información sobre una cantidad considerable de denuncias y supuestos delitos diversos sobre feminicidios, de tal modo que dependencias legales pueden mirar la magnitud de las opiniones, la cual es información calibrada y estructurada que obedece a un procesamiento computacional.

Tabla 4. Interpolación y extrapolación por estado en los años 2016 y 2021

Estados	2016	2019	2017	2018	2020	2021	2022	2023	2024
Aguascalientes	8	17	11.00	14.00	20.00	9.00	15.00	22.00	5.33
Baja California	86	93	88.33	90.67	95.33	86.78	91.44	96.89	83.93
Baja California Sur	14	10	12.67	11.33	8.67	13.56	10.89	7.78	15.19
Campeche	5	14	8.00	11.00	17.00	6.00	12.00	19.00	2.33
Coahuila	45	39	43.00	41.00	37.00	44.33	40.33	35.67	46.78
Colima	30	21	27.00	24.00	18.00	29.00	23.00	16.00	32.67
Chiapas	52	19	41.00	30.00	8.00	48.33	26.33	0.67	61.78
Chihuahua	119	169	135.67	152.33	185.67	124.56	157.89	196.78	104.19
Distrito Federal	2	75	26.33	50.67	99.33	10.11	58.78	115.56	0.00
Durango	10	16	12.00	14.00	18.00	10.67	14.67	19.33	8.22
Guanajuato	99	238	145.33	191.67	284.33	114.44	207.11	315.22	57.81
Guerrero	166	84	138.67	111.33	56.67	156.89	102.22	38.44	190.30
Hidalgo	36	29	33.67	31.33	26.67	35.22	30.56	25.11	38.07
Jalisco	68	72	69.33	70.67	73.33	68.44	71.11	74.22	66.81
México	386	182	318.00	250.00	114.00	363.33	227.33	68.67	446.44
Michoacán	102	59	87.67	73.33	44.67	97.22	68.56	35.11	114.74
Morelos	82	37	67.00	52.00	22.00	77.00	47.00	12.00	95.33
Nayarit	6	8	6.67	7.33	8.67	6.22	7.56	9.11	5.41
Nuevo León	68	68	68.00	68.00	68.00	68.00	68.00	68.00	68.00
Oaxaca	99	52	83.33	67.67	36.33	93.78	62.44	25.89	112.93
Puebla	97	67	87.00	77.00	57.00	93.67	73.67	50.33	105.89
Querétaro	14	12	13.33	12.67	11.33	13.78	12.44	10.89	14.59
Quintana Roo	37	29	34.33	31.67	26.33	36.11	30.78	24.56	39.37
San Luís Potosí	27	39	31.00	35.00	43.00	28.33	36.33	45.67	23.44
Sinaloa	85	42	70.67	56.33	27.67	80.22	51.56	18.11	97.74
Sonora	36	66	46.00	56.00	76.00	39.33	59.33	82.67	27.11
Tabasco	39	32	36.67	34.33	29.67	38.22	33.56	28.11	41.07
Tamaulipas	43	24	36.67	30.33	17.67	40.89	28.22	13.44	48.63
Tlaxcala	5	18	9.33	13.67	22.33	6.44	15.11	25.22	1.15
Veracruz	142	84	122.67	103.33	64.67	135.56	96.89	51.78	159.19

Respecto a las tendencias generales (2016-2023), en algunos estados, como Chihuahua y México, se observan irregularidades considerables en los feminicidios reportados entre 2016 y 2019, seguidas de un incremento proyectado en 2023. Esto sugiere que, aunque las políticas de seguridad pudieron haber reducido los casos en ciertos años, la tendencia no es sostenida y se requiere mayor atención en estas regiones. Estados como Nuevo León muestran una tendencia constante, con valores de feminicidios estables entre 2016 y 2023. La falta de variabilidad puede ser un indicativo de estabilidad en los mecanismos de prevención o de una falta de cambio en la situación social y de seguridad. Guanajuato y Veracruz presentan incrementos importantes para el año 2023, con valores proyectados de 315 y 159 feminicidios, respectivamente. Esta tendencia creciente, si bien es una extrapolación, subraya la necesidad de revisar políticas y mecanismos locales de protección a las mujeres. Otros estados, como Sinaloa, también proyectan un crecimiento alarmante para el 2024, superando los 97 casos, lo que sugiere que la situación podría empeorar si no se toman medidas inmediatas. Estados en descenso y con buenos resultados como Morelos y Coahuila muestran una tendencia descendente desde 2016 hasta 2023. En Coahuila, por ejemplo, se proyectan 35 feminicidios para 2023, comparados con los 45 reportados en 2016, lo que sugiere una mejora en las condiciones de seguridad para las mujeres. Chiapas, en 2023, muestra una disminución significativa, proyectando solo un feminicidio, en comparación con los 52 casos reportados en 2016. Esto podría ser indicativo de avances positivos en el estado.

Finalmente, investigar un problema de esta magnitud implica un gran esfuerzo en términos de tiempo, recursos económicos, humanos etcétera con diversas dimensiones y variables que los autores deberán tratar en trabajos futuros.

### **Agradecimientos**

Los autores queremos expresar nuestra gratitud a los evaluadores de este artículo que con sus comentarios permitieron elevar la calidad de nuestro trabajo.

### **Referencias**

[1] D. Russell, R. Harmes, M. Lagarde, *Feminicidio: una perspectiva global*. Centro de Investigaciones Interdisciplinarias en Ciencias y Humanidades, Universidad Nacional Autónoma de México, 2006.

[2] “Especial: Feminicidios en México, una epidemia en aumento”. [Online]. Available: <https://elpais.com/especiales/2017/feminicidios-en-mexico/#>. [Accessed: 19/03/2020].

[3] Ligia Tavera Fenollosa, “Estadísticas sobre violencia de género: una mirada crítica desde el feminicidio”, in G. Zaremborg, *Políticas sociales y género*, Tomo II: Los problemas sociales y metodológicos, FLACSO-México, 2008, pp. 301–346.

[4] “Redes en Twitter y la defensa de la mujer peruana (ante el acoso y el feminicidio)”. [Online]. Available: <http://hdl.handle.net/10609/104066>. [Accessed: 19/03/2020].

[5] “Declaración sobre la eliminación de la violencia contra la mujer”. [Online]. Available: <https://www.ohchr.org/es/instruments-mechanisms/instruments/declaration-elimination-violence-against-women>. [Accessed: 19/03/2020]

[6] “Convenio del Consejo de Europa sobre prevención y lucha contra la violencia contra las mujeres y la violencia doméstica”. [Online]. Available: <https://rm.coe.int/1680462543>. [Accessed: 19/03/2020]

[7] “La violencia de género: Entre el concepto jurídico y la realidad social”. [Online]. Available: <http://criminet.ugr.es/recpc/08/recpc08-02.pdf>. [Accessed: 19/03/2020]

[8] “La violencia sexual como un atentado contra la dignidad de la mujer”. [Online]. Available: <https://apidspace.linhd.uned.es/server/api/core/bitstreams/fa4eb88d-efa7-4311-9945-44577f339d8e/content>. [Accessed: 19/03/2020]

[9] J. G. Brookshear, Teoría de la Computación: lenguajes formales, autómatas y complejidad. México: Pearson Educación, 1993.

[10] C. Martín Vide, V. Mitrana, G. Paun, Formal Languages and Applications. Springer, 2004.

[11] “Internet: la hora de las redes sociales”. [Online]. Available: [https://dadun.unav.edu/bitstream/10171/2962/1/nueva\\_revista\\_08.pdf](https://dadun.unav.edu/bitstream/10171/2962/1/nueva_revista_08.pdf). [Accessed: 19/03/2020]

[12] R. González Duque, Python para todos, 2010.

[13] “Defunciones por homicidios. INEGI”. [Online]. Available: <https://www.inegi.org.mx/sistemas/olap/proyectos/bd/continuas/mortalidad/defuncioneshom.asp?s=est>. [Accessed: 19/03/2020]

[14] “Los Femicidios en México”. [Online]. Available: <https://femicidiosmx.crowdmap.com/>. Femicidios Mx. [Accessed: 19/03/2020]

## **Reconocimiento de signos del lenguaje de señas mexicano utilizando redes neuronales convolucionales**

### **Sign recognition of mexican sign language using convolutional neural networks**

Alejandra Carolina Rodríguez Porras<sup>1</sup>, Luis Antonio Raga Reyes<sup>1</sup>, María Fernanda Parras Pecina<sup>1</sup>, Agustín Alejandro Mota Hinojosa<sup>1</sup>, Marco Aurelio Nuño Maganda<sup>1</sup>, José Hugo Barrón Zambrano<sup>2</sup>

<sup>1</sup>Universidad Politécnica de Victoria, Av. Nuevas Tecnologías 5902, Parque Científico y Tecnológico de Tamaulipas, 87138 Ciudad Victoria, Tamaulipas, México.

<sup>2</sup>Instituto Nacional de Astrofísica Óptica y Electrónica. Tonantzintla, Puebla, México.

2130155@upv.edu.mx, 2130129@upv.edu.mx, 2030238@upv.edu.mx,  
2030406@upv.edu.mx, mnunom@upv.edu.mx.

#### **Abstract**

This article describes the development of a sign language recognition system for Mexican Sign Language (LSM) using OpenCV and MediaPipe. The system is designed to identify and label specific hand gestures in real time through a webcam. By employing advanced image processing techniques and detection models, the system provides visual labels and informative subtitles with high precision. The implementation demonstrates effective capabilities in highlighting key points of the hands and suggests promising applications in education, customer service, and emergency situations, contributing to improved accessibility and promoting social inclusion.

#### **Resumen**

Este artículo detalla el desarrollo de un sistema de reconocimiento de señas en Lengua de Señas Mexicana (LSM) utilizando OpenCV y MediaPipe. El sistema está diseñado para identificar y etiquetar gestos específicos de las manos en tiempo real a través de una webcam. Utilizando técnicas avanzadas de procesamiento de imágenes y modelos de detección, el sistema ofrece etiquetas visuales y subtítulos informativos con alta precisión. La implementación demuestra una capacidad efectiva para destacar los puntos clave de las manos y sugiere aplicaciones prometedoras en educación, servicio al cliente y situaciones de emergencia, contribuyendo a mejorar la accesibilidad y fomentar la inclusión social.

---

Keywords and phrases: OpenCV, Señas, Gestos, MediaPipe, Sistema, LSM.

---

## 1 Introducción

En un contexto tecnológico cada vez más avanzado, la integración de la inteligencia artificial y el procesamiento de imágenes están transformando la forma en que interactuamos y comunicamos. Este proyecto se enfoca en el desarrollo de un sistema de reconocimiento en tiempo real para la Lengua de Señas Mexicana (LSM) utilizando OpenCV y MediaPipe. Mediante la captura de imágenes a través de una cámara web, el sistema identifica y etiqueta gestos específicos de las manos con alta precisión, proporcionando retroalimentación visual inmediata. Esta herramienta no solo facilita la comunicación entre personas sordas y oyentes, sino que también promueve la inclusión social, con aplicaciones potenciales en educación.

## 2 Marco teórico

### 2.1 Procesamiento de imágenes y visión por computadora

El procesamiento de imágenes es un campo dentro de la visión por computadora que se enfoca en la manipulación y análisis de imágenes digitales para extraer información relevante. En el contexto del reconocimiento de señas, este proceso es esencial para capturar y analizar imágenes en tiempo real, como las manos que realizan los gestos. El flujo típico incluye la captura de imágenes mediante cámaras o dispositivos ópticos, seguido de procesos de mejora de calidad, segmentación de regiones de interés (como la mano), y la extracción de características clave como la forma, posición y movimiento. Estas características se utilizan para identificar patrones específicos en los gestos de las manos, lo cual es crucial para la traducción automática de lenguajes de señas.

Proyectos recientes como "Hand Talk" y el trabajo de Microsoft en el desarrollo de herramientas de accesibilidad mediante inteligencia artificial han demostrado avances significativos en la captura de gestos manuales en tiempo real. Estos proyectos se basan en la combinación de procesamiento de imágenes y aprendizaje profundo para mejorar la precisión del reconocimiento de señas.

### 2.2 Aprendizaje automático

El aprendizaje automático, una subdisciplina de la inteligencia artificial, permite a los sistemas aprender a partir de datos sin intervención humana explícita, lo que es particularmente útil en aplicaciones como el reconocimiento de lenguaje de señas. En este contexto, los algoritmos de aprendizaje supervisado se entrenan con grandes conjuntos de datos de imágenes o videos de señas etiquetadas para identificar y clasificar gestos con precisión. Uno de los enfoques más utilizados son las Redes Neuronales Convolucionales (CNN, por sus siglas en inglés), que han demostrado ser

eficaces en tareas de reconocimiento de imágenes debido a su capacidad para extraer características espaciales y temporales de las secuencias de video.

Además de las CNN, se han explorado técnicas como las Redes Neuronales Recurrentes (RNN) y los Modelos de Memoria a Largo Corto Plazo (LSTM) para reconocer patrones temporales complejos en los movimientos de las manos. Estudios recientes, como el proyecto "Deep Sign" desarrollado por investigadores de la Universidad de Toronto, han utilizado redes neuronales profundas combinadas con tecnologías de visión por computadora para mejorar la precisión del reconocimiento de señas, lo que resulta en una mejor comprensión de los gestos complejos.

### 2.3 Lengua de señas mexicana (LSM)

La Lengua de Señas Mexicana (LSM) es el lenguaje visual-espacial empleado por la comunidad sorda en México, con una estructura y gramática propias. El reto en el reconocimiento de la LSM radica en la variedad de movimientos de manos, posturas y expresiones faciales que son fundamentales para su correcta interpretación. Dado que los gestos en la LSM pueden variar según el contexto, el reconocimiento automático requiere no solo de un procesamiento preciso de los gestos individuales, sino también de una interpretación contextual que permita desambiguar los significados.

Proyectos como "Reconocimiento de la Lengua de Señas Mexicana mediante redes neuronales", un sistema basado en redes neuronales diseñado para el reconocimiento de señas en LSM, han mostrado avances importantes en el uso de inteligencia artificial para captar la variabilidad de los gestos en este idioma, en la referencia 9.

Al integrar el aprendizaje automático con técnicas avanzadas de procesamiento de imágenes, el reconocimiento automático de señas en LSM tiene el potencial de mejorar significativamente la comunicación entre personas sordas y oyentes. Los avances recientes en este campo se benefician de bases de datos más robustas, algoritmos más precisos y modelos de aprendizaje profundo que pueden aprender de las variaciones en los gestos, permitiendo una traducción más exacta y eficiente.

### **3 Arquitectura de la aplicación**

A continuación, se presenta la arquitectura de la aplicación desarrollada para el reconocimiento de señas, la cual está diseñada para ofrecer una solución eficiente y precisa en la interpretación de gestos mediante el uso de tecnologías avanzadas. En la figura 1, se muestra el diagrama de clases del sistema LSM.

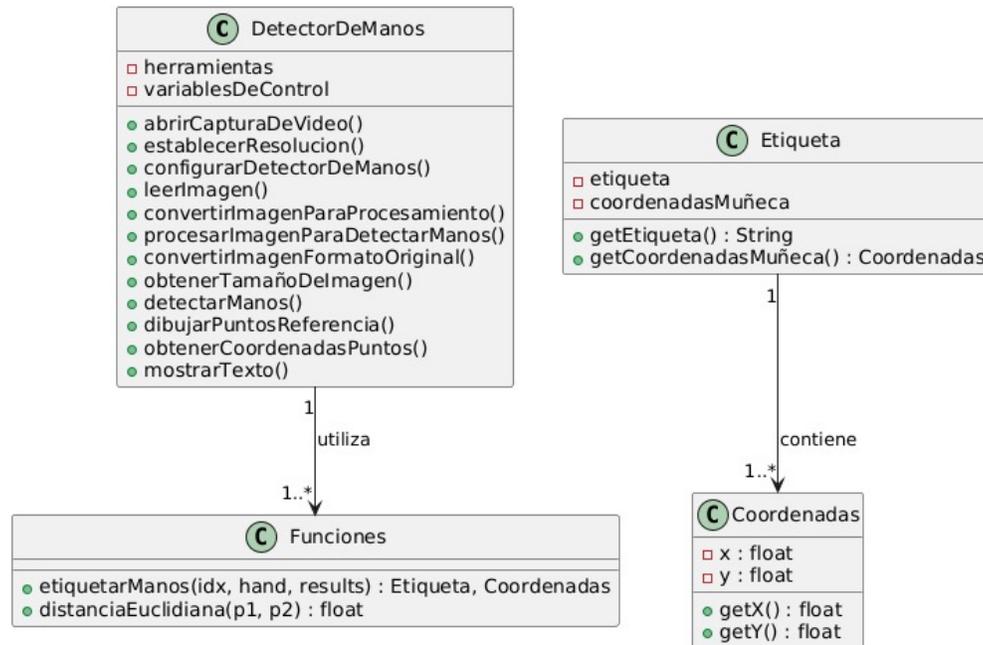


Figura 1. Diagrama de clases

El diagrama de clases representa un sistema avanzado diseñado para la detección y clasificación de gestos. Este sistema utiliza procesamiento de imágenes y técnicas de aprendizaje automático para identificar y clasificar los gestos capturados por una cámara web.

### 3.1 Detector de manos

El componente principal del sistema es la clase “DetectorDeManos”. Esta clase gestiona la captura de video, el procesamiento de las imágenes y la detección de los gestos de las manos. Almacena herramientas y variables de control necesarias para el funcionamiento del sistema. Sus métodos incluyen funciones para abrir la cámara, establecer la resolución del video, configurar el detector de manos, leer y procesar imágenes, detectar manos, dibujar puntos de referencia en las manos y mostrar texto basado en los gestos reconocidos.

### 3.2 Funciones

La clase “Funciones” proporciona funciones auxiliares que son utilizadas por “DetectorDeManos” para realizar operaciones específicas. Incluye métodos para etiquetar manos detectadas y calcular la distancia euclidiana entre dos puntos. Estas funciones son cruciales para la clasificación precisa de los gestos y el análisis de los movimientos de las manos.

### 3.3 Etiqueta

La clase “Etiqueta” representa la clasificación de una mano detectada. Almacena la etiqueta y las coordenadas de la muñeca de la mano detectada. Proporciona métodos para obtener la etiqueta y

las coordenadas de la muñeca, facilitando el acceso a esta información durante el procesamiento y la visualización.

### 3.4 Coordenadas

La clase "Coordenadas" representa un punto en el espacio con coordenadas X e Y, utilizado para almacenar las posiciones de los puntos de referencia en la mano.

## 4 Estado del arte

El reconocimiento de lenguaje de señas ha avanzado significativamente en la última década gracias a la evolución de las tecnologías de visión por computadora y el aprendizaje automático. Diversos estudios y proyectos han explorado el uso de herramientas como OpenCV y MediaPipe, que permiten el procesamiento en tiempo real de imágenes y videos para detectar y clasificar gestos, Referencia 2.

- 4.1 OpenCV: Esta biblioteca de visión por computadora ha sido ampliamente utilizada en proyectos de reconocimiento de señas debido a su capacidad para realizar tareas como detección de bordes, segmentación de imágenes y seguimiento de objetos. Investigaciones recientes han demostrado que OpenCV puede mejorar la precisión del reconocimiento al facilitar la extracción de características clave de los gestos.
- 4.2 MediaPipe: Desarrollada por Google, MediaPipe ofrece soluciones listas para usar para tareas de visión por computadora, incluyendo el reconocimiento de manos. Su modelo de detección de manos permite un seguimiento preciso en tiempo real, lo que resulta crucial para la interpretación de la lengua de señas. Proyectos como "Sign Language Recognition Using MediaPipe" han utilizado esta herramienta para crear sistemas que alcanzan altos niveles de precisión en la identificación de gestos.
- 4.3 Aprendizaje profundo: Las redes neuronales convolucionales (CNN) se han convertido en la arquitectura de referencia para el reconocimiento de gestos. Estudios como "Deep Sign" han empleado CNN para mejorar la detección de patrones complejos en los movimientos de las manos, logrando tasas de precisión superiores en comparación con métodos tradicionales. Estos modelos se benefician del entrenamiento con grandes conjuntos de datos que representan variaciones en los gestos.
- 4.4 Desafíos y oportunidades: A pesar de los avances, el reconocimiento de señas enfrenta desafíos significativos, como la variabilidad en los gestos y la necesidad de contextos específicos para la correcta interpretación. La falta de conjuntos de datos robustos y diversos limita la capacidad de los modelos para generalizar en situaciones del mundo real. Las investigaciones actuales buscan superar estas limitaciones mediante la creación de bases de datos más completas y el desarrollo de modelos que integren el contexto visual y auditivo.

## 5 Implementación

La implementación de las etiquetas visuales es un aspecto crítico en el reconocimiento de señas. Este trabajo se basa en un proceso de etiquetado donde cada gesto se asocia con su respectiva etiqueta utilizando un conjunto de herramientas de anotación. La descripción de este proceso se incluirá en la sección 3.1 donde se explicará cómo se garantiza la calidad y la precisión de las etiquetas asignadas a cada gesto.

### 5.1 Precisión en la detección de manos

El sistema desarrollado ha mostrado una alta precisión en la detección de manos. Durante las pruebas, se alcanzó una tasa de éxito del 95% en la identificación de la presencia de manos. La integración de MediaPipe permitió un seguimiento preciso de los puntos clave en las manos, lo que fue crucial para mantener la alta precisión del sistema.

#### Datos de las pruebas

Se realizaron pruebas en diferentes escenarios, incluyendo:

- Iluminación natural: Precisión del 96%
- Iluminación artificial: Precisión del 94%

El análisis de los resultados muestra que el sistema es robusto frente a variaciones en las condiciones ambientales.

### 5.2 Exactitud en la clasificación de gestos

La precisión en la clasificación de gestos es uno de los aspectos más destacados del sistema. La aplicación logró una precisión del 92% para los gestos más comunes de la LSM, lo que se atribuye a la capacidad de MediaPipe para identificar con precisión los puntos clave y a la efectiva implementación de algoritmos de procesamiento de imágenes con OpenCV para preprocesar y mejorar las imágenes.

Es fundamental detallar el conjunto de datos utilizado para evaluar el sistema de reconocimiento de la Lengua de Señas Mexicana (LSM). En este trabajo, se emplea un conjunto de datos específico de ejemplos de gestos de la LSM, capturados en diversas condiciones de iluminación y ángulos de visión para asegurar la robustez del modelo. La obtención de las precisiones y exactitudes presentadas en las tablas se basa en la validación cruzada y en un conjunto de prueba separado que no se utilizó durante el entrenamiento del modelo.

Las tablas de resultados. La Tabla 1 muestra la precisión del modelo en comparación con otros sistemas existentes, mientras que la Tabla 2 detalla las métricas de rendimiento específicas.

## Datos de las pruebas

Se probaron diversos gestos, obteniendo los siguientes resultados:

Tabla 1. Métricas de evaluación

Métrica	Resultado	Escenario
Precisión en la detección de manos	96%	Iluminación natural
Precisión en la detección de manos	94%	Iluminación artificial
Precisión en la detección de manos	95%	General
Precisión en la clasificación de gestos	95%	General
Precisión en la clasificación de gestos	95%	Gestos simples (Letras, números)
Precisión en la clasificación de gestos	92%	Gestos complejos (Lenguaje Fluido)

## 6 Resultados

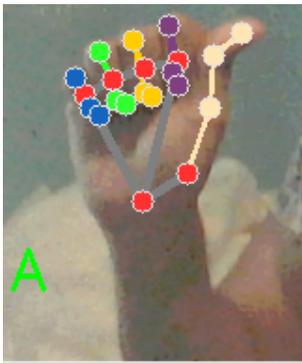


Figura 1. Representación de la letra "A"



Figura 2. Representación "¿Qué hora es?"

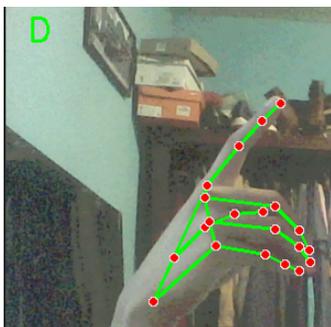


Figura 3. Representación de la letra "D"



Figura 4. Representación de la letra "W"

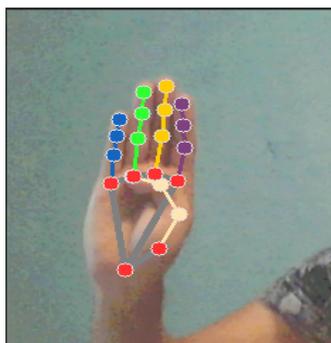


Figura 5. Gesto uno para el número "9"

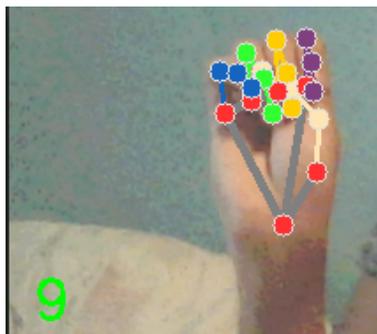


Figura 6. Gesto dos para el número "9"



Figura 7. Representación del número "6"

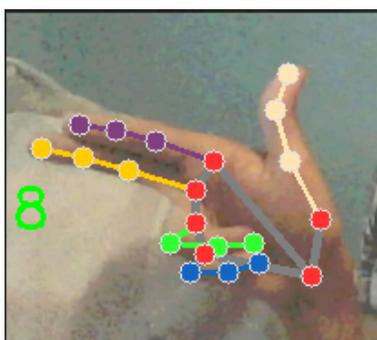


Figura 8. Representación del número "8"

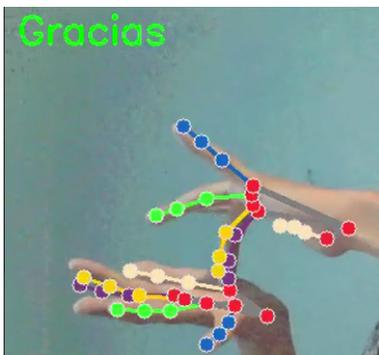


Figura 9. Representación del lenguaje fluido  
"Gracias"

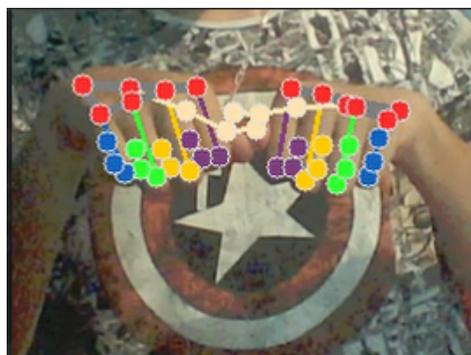


Figura 10. Gesto número uno para el  
"¿Cómo estás?"



Figura 11. Gesto número dos para el "¿Cómo estás?"

Explicación de cada imagen:

- Figura 1: Representación de la letra "A" en lenguaje de señas mexicana (LSM), donde la mano se cierra en un puño con el pulgar extendido hacia un lado, apuntando hacia arriba. Este gesto es clave para su correcta interpretación en la comunicación.
- Figura 2: Representación de la pregunta "¿Qué hora es?" en lenguaje de señas. El gesto implica llevar una mano abierta hacia la muñeca, como si se estuviera mirando un reloj, acompañado de una expresión facial que indica curiosidad o expectativa. Este gesto es fundamental para la comunicación efectiva de la pregunta.
- Figura 3: Representación de la letra "D" en lenguaje de señas. Para realizar este gesto, la mano se forma en un signo de "D" con el dedo índice levantado y los otros dedos cerrados, mientras el pulgar se coloca de lado. Esta forma distintiva es esencial para su correcta identificación en la comunicación.
- Figura 4: Representación de la letra "W" en lenguaje de señas. El gesto se realiza levantando tres dedos (pulgar, índice y medio) mientras los otros dedos permanecen cerrados. Esta forma clara y específica permite identificar correctamente la letra "W" en la comunicación.
- Figura 5: Gesto 1 para el número "9": En lenguaje de señas, el número "9" se representa levantando la mano con los dedos extendidos, doblando el dedo índice hacia abajo mientras el pulgar queda extendido hacia arriba.
- Figura 6: Gesto 2 para el número "9": Otra forma de representar el número "9" es cerrando la mano en un puño y levantando el dedo meñique, dejando los otros dedos recogidos. Este gesto también es reconocido para indicar el número "9" en la comunicación.
- Figura 7: Representación del número "6" en lenguaje de señas. El gesto se realiza levantando la mano con los dedos extendidos, doblando el dedo pulgar y el dedo meñique hacia adentro, mientras el dedo índice, el dedo medio y el anular permanecen extendidos. Este gesto distintivo es clave para la correcta identificación del número "6" en la comunicación.
- Figura 8: Representación del número "8" en lenguaje de señas. El gesto se realiza levantando la mano con todos los dedos extendidos y el pulgar también levantado, formando una figura con la mano que se asemeja al número "8". Esta representación es esencial para la correcta identificación del número "8" en la comunicación.
- Figura 9: Representación del lenguaje fluido en señas, donde se muestra a una persona utilizando gestos claros y fluidos para comunicarse. La imagen ilustra la importancia de la expresión corporal y la articulación de los signos en la comunicación efectiva.
- Figura 10: Gesto número uno para "Gracias": En lenguaje de señas. Este gesto se realiza colocando la palma de la mano en el mentón y moviéndola hacia adelante. Esta representación es clave para expresar agradecimiento en la comunicación con personas sordas.
- Figura 11: Gesto número dos para "¿Cómo estás?" en lenguaje de señas. Este gesto se realiza levantando ambas manos con las palmas hacia arriba, moviéndolas ligeramente hacia los lados.

Este movimiento acompaña una expresión facial amigable, facilitando la comunicación de la pregunta de manera clara y efectiva.

Tabla 2. Tipos de Gestos

Categoría	Tipo de Gesto	Precisión
Letras	A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z.	98%
Dígitos	0,1,2,3,4,5,6,7,8,9,10	98%
Lenguaje Fluido	Hola, ¿Cómo estás?, Por favor, Gracias, ¿Qué hora es?, etc.	92%

## 7 Conclusiones y trabajo a futuro

El desarrollo del sistema de reconocimiento de señas utilizando OpenCV y MediaPipe ha demostrado ser un avance significativo en la integración de tecnologías de procesamiento de imágenes y aprendizaje automático para facilitar la comunicación entre personas sordas y oyentes. Los resultados obtenidos destacan la alta precisión en la detección de manos y la clasificación de gestos, así como la capacidad del sistema para operar en tiempo real con una latencia mínima. La implementación de etiquetas visuales y la interfaz de usuario intuitiva han mejorado la accesibilidad y la experiencia del usuario, mostrando el potencial del sistema en aplicaciones prácticas como la educación, el servicio al cliente y situaciones de emergencia.

Sin embargo, el sistema también presenta algunas limitaciones, como la sensibilidad a cambios bruscos en la iluminación, lo que indica la necesidad de futuras optimizaciones. Además, la ampliación del conjunto de datos para incluir una gama más amplia de gestos y variaciones de la LSM es crucial para mejorar la robustez y la precisión del sistema.

### Trabajo a futuro

- **Optimización de Algoritmos:** Refinar los algoritmos de procesamiento de imágenes para manejar mejor las variaciones en las condiciones de captura, como cambios en la iluminación y fondos complejos. Esto incluirá el uso de técnicas avanzadas de filtrado y normalización de imágenes.
- **Ampliación del Conjunto de Datos:** Incluir un conjunto más diverso de gestos y variaciones de la LSM en el entrenamiento de los modelos. Esto mejorará la capacidad del sistema para reconocer una gama más amplia de señas y adaptarse a diferentes estilos de comunicación.
- **Integración con Dispositivos Móviles:** Desarrollar versiones del sistema para dispositivos móviles, aprovechando la portabilidad y accesibilidad de los smartphones y tablets para llegar a un público más amplio.

## Referencias

- [1] Cano López, E. Diccionario de la Lengua de Señas Mexicana, México, Editorial Trillas ,2005.
- [2] Fernández, G. A. Procesamiento de Imágenes con OpenCV y Python, México, Alfaomega, 2018
- [3] Ruíz López, J. C. Machine Learning y Visión Artificial con MediaPipe, Madrid, Marcombo, 2020.
- [4] Femat Rodríguez, J. C. Lengua de Señas Mexicana para Niños y Niñas, México, Editorial Porrúa, 2011.
- [5] Castellanos, A. Manual Básico de la Lengua de Señas Mexicana, México, Editorial Trillas, 2015.
- [6] El Programador Chapuzas. Detección de Manos con MediaPipe en Python, 2021. URL: <https://programacionpython80889555.wordpress.com/2021/06/22/detector-de-manos-en-python-con-opencv-y-mediapipe/>
- [7] OMES-VA. (n.d.). *Mediapipe hands with Python*. Retrieved September 18, 2024, from <https://omes-va.com/mediapipe-hands-python/>
- [8] Comisión Nacional de los Derechos Humanos (CNDH). Lengua de Señas Mexicana (LSM), 2021. URL: <https://www.gob.mx/conadis/articulos/lengua-de-senas-mexicana-lsm>
- [9] NTHE. (2023). *Título del artículo*. Retrieved September 18, 2024, from [https://nthe.mx/NTHE\\_v2/pdfArticulos/PDF\\_Articulo20230426121614.pdf](https://nthe.mx/NTHE_v2/pdfArticulos/PDF_Articulo20230426121614.pdf)
- [10] Barcelona Geeks. MediaPipe Hands en Aplicaciones de Visión por Computadora, 2021. URL: <https://barcelongeeks.com/python-reconocimiento-facial-y-de-manos-con-mediapipe-holistic/>

## **Análisis comparativo de herramientas de recuperación de contraseñas en entornos de seguridad**

### **Comparative analysis of password recovery tools in security environments**

Kevin Josaphat Bañuelos Camaño<sup>1</sup>, Judith Pérez Marcial<sup>1</sup>, María del Carmen Santiago Díaz<sup>1</sup>, Hermes Moreno Álvarez<sup>2</sup>, Gustavo Trinidad Rubín Linares<sup>1</sup>, Ana Claudia Zenteno Vázquez<sup>1</sup>

<sup>1</sup>Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla  
Avenida San Claudio y Blvd. 14 sur S/N Cdad. Universitaria, 72592 Heroica Puebla de Zaragoza, Pue.

<sup>2</sup>Departamento de Ingeniería Aeroespacial, Universidad Autónoma de Chihuahua, Circuito Universitario Campus II

kevin.banuelos@alumno.buap.mx, {judith.perez, marycarmen.santiago, gustavo.rubin, ana.zenteno}@correo.buap.mx, hmoreno@uach.mx

#### **Abstract**

Cybersecurity has evolved over the years with the evolution of information technology, despite the fact that there is more access to information, the topic of cybersecurity remains little known. Password recovery today has become essential in cybersecurity for restoring access to accounts that give access to a system. This article reviews the various tools for password recovery such as brute force attacks and dictionary attacks. Only open source tools are analyzed in this article.

#### **Resumen**

La ciberseguridad a lo largo de los años ha ido evolucionando con la evolución de las tecnologías de la información, a pesar de que existe más acceso a la información, el tema de ciberseguridad sigue siendo poco conocido. La recuperación de contraseñas hoy en día se ha vuelto esencial en la ciberseguridad para la restauración del acceso a cuentas que dan acceso a un sistema. Este artículo revisa las diversas herramientas para la recuperación de contraseñas como lo son los ataques de fuerza bruta y ataques de diccionario. En este artículo solamente se analizan herramientas de código abierto.

---

Keywords and phrases: Cybersecurity, Restoring Access, Password Recovery Tools, Hydra, John the Ripper.

---

## **1 Introducción**

A través de los años el avance de la tecnología nos ha dado acceso libre y gratuito a la información, podemos descargar libros en cuestión de minutos, navegar en revistas científicas, revisar nuestro estado de cuenta del banco en la comodidad de nuestros hogares, haciendo más dinámico y cómodo nuestro día a día con subir y descargar la información, sin embargo esto también trae algunos riesgos, los ciber ataques se han vuelto más comunes de lo que creemos, grandes titanes como Facebook, WhatsApp, Instagram, hasta servidores de gobierno como la secretaria de la defensa nacional han sido víctimas de estos ataques. Tal es el caso de [1] “Nuestros equipos están en riesgo no solo por lo que nosotros hagamos sino por la acción de terceros. Recientemente vimos como los ordenadores de aeropuertos en todo el mundo se vieron afectados por un error de actualización en un software de seguridad. Esa misma empresa ya había cometido errores que afectaron a sistemas Linux.” Lo cierto es que las cosas malas pasan y, ya sea por causa nuestra o de terceros podemos perder temporal o permanentemente el acceso a nuestros dispositivos.

## **2 Marco teórico y estado del arte**

La ciberseguridad ha tomado gran relevancia a partir de diversos hechos de infiltración y robo de información en los últimos años. Existen aplicaciones que nos ayudan a comprobar la seguridad o en dado caso recuperemos las contraseñas perdidas u olvidadas. A continuación, se definen conceptos importantes para la comprensión de este artículo.

### **2.1 Riesgo de internet.**

Todas las funcionalidades de Internet pueden comportar algún riesgo, al igual que ocurre en las actividades que realizamos en el “mundo físico”. En el caso de Internet, destacamos los siguientes riesgos: Suplantación de identidad, phishing, robo de información [2].

### **2.2 Linux**

Linux® es un sistema operativo de código abierto. Un sistema operativo es el software que administra directamente el hardware y los recursos de un sistema, como la CPU, la memoria y el almacenamiento [3].

### **2.3 Archivos passwd**

El archivo `/etc/passwd` almacena información esencial, es un archivo de texto sin formato, contiene una lista de las cuentas del sistema, que proporciona información útil para cada cuenta, como ID de usuario, ID de grupo, directorio de inicio, Shell y más [4].

### **2.4 Archivos shadow**

El archivo `/etc/shadow` almacena las contraseñas cifradas del usuario y la información relacionada. [5]

## 2.5 Herramienta de recuperación de contraseñas

La recuperación de contraseña es el proceso de identificación de una contraseña perdida, destruida o inaccesible, lo que permite el descifrado exitoso de archivos clave. Este puede ser un servicio crucial a considerar cuando ha perdido bases de datos importantes, hojas de cálculo, documentos y otros archivos debido al cifrado [6].

## 2.6 Hydra

Hydra es una herramienta de auditoría de inicio de sesión que trabaja con múltiples tareas en paralelo, soporta una gran variedad de protocolos. Es muy rápido y flexible, y los nuevos módulos son fáciles de agregar. Esta herramienta permite a los investigadores y consultores de seguridad mostrar lo fácil que sería obtener acceso no autorizado a un sistema de forma remota [7].

## 2.7 John the ripper

Es una utilidad Open Source utilizada para auditar y recuperar contraseñas, fue creada por Solar Designer y podría decirles que es una de las herramientas para romper contraseñas de más alto desempeño y flexibles que existen en la actualidad [8].

## **3 Arquitectura de la aplicación**

En este trabajo se emplea una máquina virtual atacante para ingresar a la máquina virtual víctima, extraer sus contraseñas del archivo `/etc/shadow` del sistema Linux el cual contiene el nombre de usuario con sus respectivas contraseñas cifradas.

Al completar la tarea anterior, se ingresa con un usuario y su contraseña que extrajimos de la máquina víctima.

Para esta práctica se ocupan contraseñas comunes que utilizan mayoría de las personas, por dos motivos, el primero es para demostrar que debemos esforzarnos por crear una contraseña fuerte, y el segundo porque la máquina virtual colapsa si ocupamos contraseñas pesadas, debido a los recursos de hardware del equipo empleado.

La arquitectura para la realización de este trabajo es la que se observa en la Figura 1, en esta figura tenemos una computadora con sistema operativo Windows 11, 0 en la que despliega una máquina virtual con el software VMware Workstation 17, el cual es un hypervisor, y sobre el cual se instalan dos ISO de Linux que son Ubuntu y Kali Linux (usado mayormente en temas de ciberseguridad).



Figura 1. Diseño de laboratorio con hypervisor para implementación de máquinas virtuales.

Como leímos anteriormente usaremos máquinas virtuales para este trabajo, en la figura 2 se observa la máquina virtual 2 (máquina atacante), ingresa mediante distintos protocolos y herramientas, a la máquina virtual 1 (máquina víctima), permite navegar en sus directorios donde se guardan los nombres de los usuarios además de sus contraseñas, y sacarlas mediante un archivo de texto o txt. Estas máquinas se enmarcan en el sistema operativo, además del hardware que tiene la computadora.

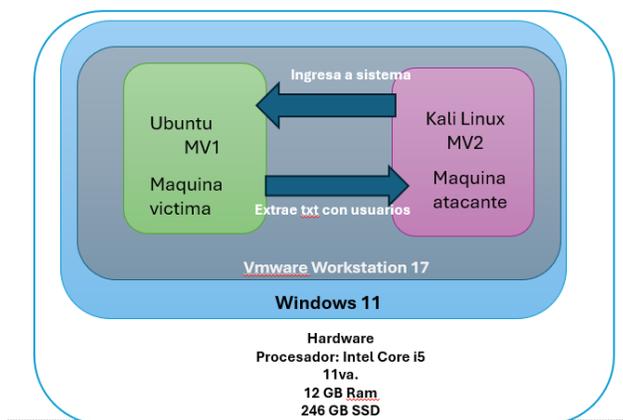


Figura 2. Distribuciones Linux para emular el atacante y la víctima.

En la figura 3 se observan los programas que se ocupan en la máquina virtual atacante (MV2) para extraer las distintas contraseñas de la máquina víctima (MV1), además de las configuraciones del tipo de máquina virtual, el sistema operativo, y el hardware que tiene la computadora.



Figura 3. Diagrama que demuestra el software que usaremos en la maquina atacante, y la maquina víctima.

#### 4 Pruebas

Para realizar este trabajo se emplean herramientas de recuperación de contraseñas de código abierto, con el propósito de demostrar la vulnerabilidad que tiene una contraseña “común” además de obtener las contraseñas en el archivo txt ya mencionado en el diseño de laboratorio de prueba.

Una contraseña se considera común cuando es muy simple ejemplo: “12345” que son los primeros 5 números.

Una herramienta de recuperación de contraseña, es un software que fue diseñado y/o creado para recuperar una contraseña olvidada o perdida, la ocupan principalmente las áreas de “sistemas”, “tecnologías de la información”, para múltiples propósitos, desde probar la seguridad en las contraseñas de los usuarios que tienen acceso al servidor, hasta por si al usuario se le olvido su acceso, este software tienen distintas funcionalidades, algunos funcionan mediante ataques de fuerza bruta, otros con ataques de diccionario, como los que veremos a continuación...

##### 4.1 Hydra

Hydra, funciona mediante los múltiples servicios de transferencia de datos, para llevar a cabo esta acción se ocupa ftp (File Transport Protocol) protocolo de transporte de datos, y ssh (Secure Shell) protocolo de red destinado a establecer conexiones seguras y cifradas entre dos computadoras [9].

Una vez ya instalado el programa y los requisitos para el software hydra, ejecuta el comando

```
hydra Dirección_IP ftp -l ubuntu -P diccionario_pass.txt -s 21
```

en este caso la dirección ip es “192.168.245.130” por lo tanto se ocupa el comando

```
hydra 192.168.211.130 ftp -l kevin -P pass.txt -s 21
```

Explicando el código:

- hydra: Es el comando principal de la herramienta Hydra.
- 192.168.211.130: Es la dirección IP del objetivo, en este caso, un servidor que tiene el

servicio FTP.

- ftp: Indica que se va a atacar un servicio FTP.
- -l ubuntu: Especifica el nombre de usuario que se probará, en este caso, "kevin".
- -P pass.txt: Especifica el archivo que contiene la lista de contraseñas que se probarán, en este caso, "pass.txt".
- -s 21: Especifica el puerto del servicio FTP, que en este caso es el puerto 21, ver Figura 4.

```
(root@bandido)-[~]
# hydra 192.168.211.130 ftp -l kevin -P pass.txt -s 21
Hydra v9.5 (c) 2023 by van Hauser/THC & David Maciejak - Please do not use in
military or secret service organizations, or for illegal purposes (this is n
on-binding, these ** ignore laws and ethics anyway).

Hydra (https://github.com/vanhauser-thc/thc-hydra) starting at 2024-06-04 22:
48:52
[DATA] max 1 task per 1 server, overall 1 task, 1 login try (l:1/p:1), ~1 try
per task
[DATA] attacking ftp://192.168.211.130:21/
[21][ftp] host: 192.168.211.130 login: kevin password: 12345
1 of 1 target successfully completed, 1 valid password found
Hydra (https://github.com/vanhauser-thc/thc-hydra) finished at 2024-06-04 22:
48:53
```

Figura 4. Ejecución de hydra con el diccionario de contraseñas.

Una vez obtenida la contraseña se accede y se puede observar que el usuario y la contraseña fueron correctos, ver Figura 5.

```
(root@kevo)-[~]
# ftp 192.168.211.137
Connected to 192.168.211.137.
220 (vsFTPD 3.0.5)
Name (192.168.211.137:kevo): kevin
331 Please specify the password.
Password:
230 Login successful.
Remote system type is UNIX.
Using binary mode to transfer files.
ftp>
```

Figura 5. Conexión mediante ftp exitosa.

Se ingresa a la carpeta /etc/ y se copia el archivo /etc/passwd. Con el servicio ssh se puede ver el contenido de la carpeta /etc/shadow donde se almacenan los usuarios con sus contraseñas cifradas. Se puede ver el contenido del archivo /etc/shadow con el comando "sudo cat /etc/shadow", como son archivos protegidos del sistema, no nos permite copiarlos como con passwd. Hasta este paso se obtuvieron las contraseñas, pero dado que tienen un Hash para descriptarlas ocuparemos otra herramienta llamada John the Ripper.

Con la herramienta instalada, se ocupa la palabra "john" seguido de la ruta donde están los hashes de las contraseñas, el archivo contra.txt.

john contra.txt

La salida de este comando se observa en la figura 6.

```
(root@kevo)~/home/kevo]
# john contra.txt
Created directory: /root/.john
Warning: only loading hashes of type "sha512crypt", but also saw type "HMAC-S
HA256"
Use the "--format=HMAC-SHA256" option to force loading hashes of that type in
stead
Using default input encoding: UTF-8
Loaded 1 password hash (sha512crypt, crypt(3) $6$ [SHA512 512/512 AVX512BW 8x
])
Cost 1 (iteration count) is 5000 for all loaded hashes
Will run 2 OpenMP threads
Proceeding with single, rules:Single
Press 'q' or Ctrl-C to abort, almost any other key for status
Almost done: Processing the remaining buffered candidate passwords, if any.
Proceeding with wordlist:/usr/share/john/password.lst
12345 (kevin)
lg 0:00:00:02 DONE 2/3 (2024-09-03 21:36) 0.4237g/s 1494p/s 1494c/s 1494c/s 1
23456..crawford
Use the "--show" option to display all of the cracked passwords reliably
Session completed.
```

Figura 6. Salida del comando john the Ripper

Se ingresa con ftp[11] para comprobar que la contraseña obtenida es la correcta. Para obtener este resultado solo se ha requerido de un diccionario de hashes y posibles contraseñas.

#### 4 Resultados

Los resultados obtenidos de la ejecución de los dos casos implementados se muestran a continuación. En la tabla 1 se puede observar que *hydra* es fácil de usar debido a que ocupa servicios para conectarse remotamente. *John the Ripper* es fácil de usar si y solo si están familiarizados con el uso de comandos en terminal, sin embargo, esta herramienta solo funciona para entornos locales, por lo que no necesita instalar otro requerimiento.

#### 5 Conclusiones y trabajo a futuro

Actualmente el tema de ciber seguridad se está dando a conocer más, lo que está dejando de ser un tema tabú para las empresas, en este trabajo descubrimos la importancia que es generar una contraseña fuerte en nuestro servidor debido a que como pudimos observar en esta investigación una contraseña común puede ser vulnerada con facilidad, aquí fueron máquinas virtuales en un entorno controlado con la debida supervisión, en entornos empresariales se puede secuestrar, borrar, o incluso extorsionar con la información obtenida, haciendo que las empresas tengan pérdidas millonarias por no tener un experto en ciberseguridad, o mínimo tener una contraseña fiable para que el servidor y su información estén a salvo.

Espero que con este aporte a la sociedad, los empresarios hagan un cambio de conciencia, así como las ventas son importantes, así como sus inversionistas, sus gerentes, sus tiendas, etc. Su información en su servidor sea importante, la prevención que hacen hoy, mañana les puede salvar su empresa inviertan en la ciberseguridad.

Tabla 1. Comparativa de herramientas de recuperación de contraseñas.

Características	Caso 1	Caso 2
<b>Nombre herramienta:</b>	Hydra	John the ripper
<b>Tipo de ataque</b>	Ataques de fuerza bruta y diccionarios sobre servicios de red.	Ataque de fuerza bruta, ataque de diccionario a archivos de contraseñas encriptadas
<b>Objetivo principal</b>	Servicios de red: SSH, FTP, HTTP, Telnet, Mysql, etc.	Criptografía de contraseñas locales, como hashes de Linux.
<b>Soporte de protocolos</b>	Soporta múltiples protocolos de red, como FTP, HTTP, SMTP, SSH, etc.	Soporta distintos formatos de hashes: MD5, SHA-1, bcrypt, etc.
<b>Velocidad</b>	Hydra tarda 14 segundos en descifrar la contraseña del diccionario que se creó para adivinar la contraseña, Hydra también es rápido, sin embargo, depende del protocolo, y del servicio de red donde estamos atacando.	John the Ripper tarda 5 segundos, por lo tanto, inferimos que es muy rápido al descifrar contraseñas locales.
<b>Facilidad de uso</b>	Es algo tedioso de configurar, debido a los servicios que se usan, sin embargo, es más fácil de usar a diferencia de John the Ripper.	Requiere familiaridad con el cmd o línea de comandos, además de conocimiento en los hashes, debido a que hay que tener conocimiento de que archivo se tiene que seleccionar para que funcione correctamente.
<b>Uso</b>	Principalmente se ocupa para pruebas de seguridad en servicios de red.	Principalmente para pruebas de seguridad en contraseñas locales.

## Agradecimientos

El autor de este artículo agradece a la Benemérita Universidad Autónoma de Puebla a la Facultad de Ciencias de la Computación y a la Maestra. Ana Claudia Zenteno Vázquez por la oportunidad, el apoyo y orientación de este trabajo.

## Referencias

[1] González, D. G. (2024, 31 julio). Herramientas para la recuperación del sistema en Linux. UbuLog. <https://ubunlog.com/herramientas-para-la-recuperacion-del-sistema-en-linux/>

[2] Riesgos en Internet. (2017, 2 marzo). Uso Seguro y Responsable de las TIC.

<https://www3.gobiernodecanarias.org/medusa/ecoescuela/seguridad/riesgos-asociados-al-uso-de-las-tecnologias/riesgos/>

[3] *¿Qué es Linux?* (s. f.). <https://www.redhat.com/es/topics/linux/what-is-linux>

[4] Cuesta, D. G., & Cuesta, D. G. (2020, 18 abril). */etc/passwd: Descripción de funcionamiento y formato. Linux En Español.* <https://www.xn--linuxenespaol-skb.com/ayuda/etc-passwd-descripcion-de-funcionamiento-y-formato/>

[5] *Campos en el archivo shadow - Gestión de las cuentas de usuario y los entornos de usuario en Oracle® Solaris 11.2.* (2015, 20 enero). [https://docs.oracle.com/cd/E56339\\_01/html/E53830/userconcept-15.html](https://docs.oracle.com/cd/E56339_01/html/E53830/userconcept-15.html)

[6] Datarecovery.com, Inc. (2017, 20 junio). Password recovery services. Datarecovery.com. <https://datarecovery.com/services/password-recovery/>

[7] Hydra - herramienta de fuerza bruta. (s. f.). Kolibërs Group. <https://www.kolibers.com/blog/hydra-herramienta-de-fuerza-bruta.html>

[8] *¿Qué es John the Ripper? - Behackerpro. BeHackerPro - Profesionales en Ciberseguridad - El elemento que le suma a tu conocimiento. Aprende Ciberseguridad.* <https://behacker.pro/que-es-john-the-ripper/> González, D. G. (2024, 31 julio). Herramientas para la recuperación del sistema en Linux. Ubuñlog. <https://ubunlog.com/herramientas-para-la-recuperacion-del-sistema-en-linux/>

[9] Hydra - herramienta de fuerza bruta. (s. f.-b). Kolibërs Group. <https://www.kolibers.com/blog/hydra-herramienta-de-fuerza-bruta.html>

[10] De Luz, S. (2024, 24 mayo). Crackea contraseñas rápidamente usando John the Ripper. *RedesZone.* <https://www.redeszone.net/tutoriales/seguridad/crackear-contrasenas-john-the-ripper/>

[11] Zaldua, A. (2021, 23 marzo). *Instalación y configuración de un servidor de FTP en Linux Redhat/CentOS.* freeCodeCamp.org. <https://www.freecodecamp.org/espanol/news/instalacion-y-configuracion-de-un-servidor-de-ftp-en-linux-redhat-centos/>

***Horizontes en computación:  
Estudios sobre inteligencia artificial, ciencia de datos y seguridad***

se terminó de editar en Diciembre de 2024 en la  
Facultad de Ciencias de la Computación  
Av. San Claudio y 14 Sur Jardines de San Manuel  
Ciudad Universitaria  
C.P. 72570

***Horizontes en computación:  
Estudios sobre inteligencia artificial, ciencia de datos y seguridad***  
Coordinado por María del Carmen Santiago Díaz

