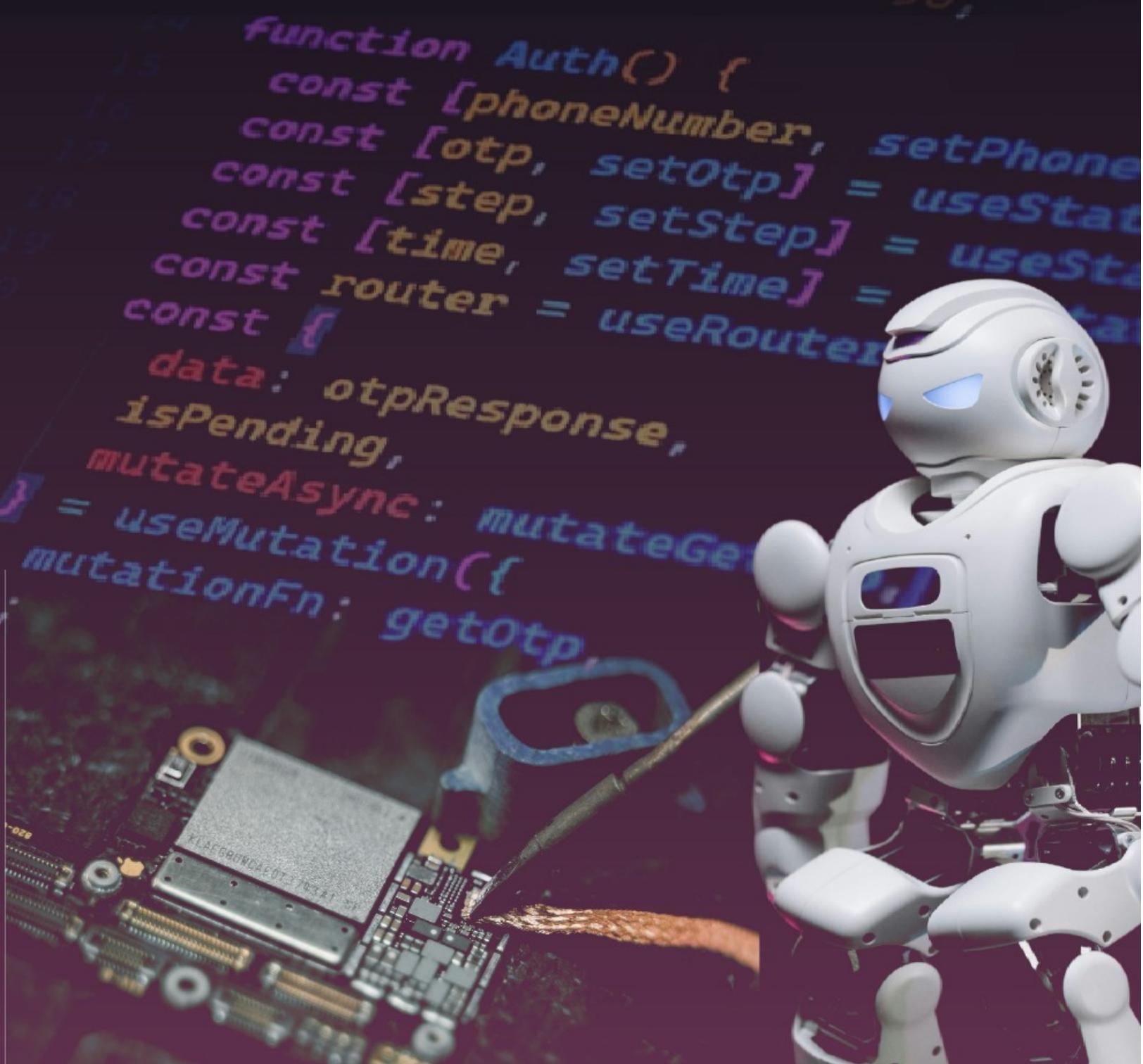


Avances en Computación: Perspectivas Académicas y de Investigación en Inteligencia Artificial, Ciencia de Datos y Seguridad



Gustavo Trinidad Rubín Linares

Avances en Computación:
Perspectivas Académicas y de Investigación en
Inteligencia Artificial, Ciencia de Datos y Seguridad

María del Carmen Santiago Díaz

Gustavo Trinidad Rubín Linares

Ana Claudia Zenteno Vázquez

Judith Pérez Marcial

(Editores)

Gustavo Trinidad Rubín Linares

(Coordinador)

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

Gustavo Trinidad Rubín Linares
(coordinador BUAP)

Amilcar Meneses Viveros, Donaldo García Juanillo, Abel Alejandro Rubín Alvarado, Miguel Morales Sandoval, Oleg Starostenko Basarab, Judith Pérez Marcial, Lorna Verónica Rosas Téllez, Jaime Julián Cid Monjaraz, María Concepción Landa Arnaiz, Francisco Marroquín Gutiérrez, Jéssica Nayeli López Espejel Fernando Moreno Barbosa, José Hugo barrón Zambrano, Salvador Eugenio Ayala Raggi, Raúl Castro Ortega, Ana Claudia Zenteno Vázquez, Gabriela Francisca Solís Magaña, Enrique González Amador, Luis Fernando Curi Quintal, José Ignacio Castillo Velázquez, Miguel Angel San Pablo Juárez, Gustavo Manuel Minquiz Xolo, José Rubén Conde Sánchez, Gustavo Trinidad Rubín Linares, Hiram Herrera Rivas, Javier Flores Méndez, Juan Pablo Ucán Pech, Erik Zamora Gómez, Alfonso Martínez Cruz, Vittorio Zanella Palacios, Alejandro Peñuelas Angulo, Ricardo Alvarez González, Marco Aurelio Nuño Maganda, María del Carmen Santiago Díaz, Beatriz Alejandra Flores Rojas, Lauro Reyes, Víctor Menendez Domínguez, Antonio Armando Aguilera Güemez, Elsa Rubio Espino, Gabriel Alejandro Galaviz Mosqueda, José Juan García Hernández.

(revisores)

Primera edición: 2024

ISBN: 978-607-8975-34-1

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

| | |
|---------------|---|
| Prefacio..... | 6 |
|---------------|---|

Nuevas tendencias de UX/UI en aplicaciones móviles

| | |
|--|----------|
| <i>Juan Carlos Pablo Santos</i> | |
| <i>Abraham Sánchez López</i> | |
| <i>Rogelio González Velázquez.....</i> | <i>7</i> |

Aplicación móvil con realidad aumentada para la capacitación en el manejo de la metodología 5S's

| | |
|--|-----------|
| <i>Ana Luisa Ballinas Hernández</i> | |
| <i>Héctor Medina Cruz</i> | |
| <i>Eduardo Ocelotl Valencia</i> | |
| <i>María Claudia Denicia Carral</i> | |
| <i>Reyna Villalobos Cervantes.....</i> | <i>18</i> |

Diseño integral de arquitectura de microservicios para el sistema de tutoría académica

| | |
|--|-----------|
| <i>Orestes Calvario Palafox</i> | |
| <i>Judith Pérez Marcial</i> | |
| <i>María del Carmen Santiago Díaz</i> | |
| <i>Pedro García Juárez</i> | |
| <i>José Ignacio Castillo Velázquez</i> | |
| <i>Ana Claudia Zenteno Vázquez</i> | |
| <i>Gustavo Trinidad Rubín Linares.....</i> | <i>29</i> |

Identificador de genero mediante audio

| | |
|--|-----------|
| <i>Melchor Hernández Díaz</i> | |
| <i>Jesús Antonio Barbosa Rodriguez</i> | |
| <i>Ximena Abigail Moreno Ledesma</i> | |
| <i>Juan Daniel Torres Colorado</i> | |
| <i>José Hugo Barrón Zambrano</i> | |
| <i>Marco Aurelio Nuño Maganda.....</i> | <i>41</i> |

Redes bayesianas a partir de datos

| | |
|--|-----------|
| <i>Gladys Linares Fleites</i> | |
| <i>Manuel Alejandro Mier Gómez</i> | |
| <i>Hortensia Josefina Reyes Cervantes.....</i> | <i>51</i> |

Optimización del proceso de desarrollo web mediante la generación automatizada de código utilizando inteligencia artificial

| | |
|---|-----------|
| <i>Judith Pérez Marcial</i> | |
| <i>María del Carmen Santiago Díaz</i> | |
| <i>Gustavo Trinidad Rubín Linares</i> | |
| <i>Ana Claudia Zenteno Vázquez</i> | |
| <i>Óscar Iván Hernández Gómez</i> | |
| <i>María Catalina Rivera Morales</i> | |
| <i>María Dolores Guevara Espinosa</i> | <i>59</i> |

Simulación mediante Python y ROS de los algoritmos de Dijkstra y A* para el movimiento de un robot móvil autónomo

Ivonne Pérez Flores
Dayana Pérez Flores
Ricardo Álvarez González
Alba Maribel Sánchez Gálvez
Brian Manuel González Contreras.....68

Desarrollo de una aplicación para la detección y conteo de monedas en video en tiempo real utilizando OpenCV y PyQt6

Adrián Alejandro Ruiz Márquez
Brayan Olivares Rodriguez
Javier Martin Palmero Torres
Ana Cecilia Rodríguez Chávez
Marco Aurelio Nuno Maganda.....78

Detección y prevención de intrusos en sistemas operativos Linux con Suricata

Vicente Zavaleta Sanchez
Judith Pérez Marcial
María del Carmen Santiago Díaz
Carina Toxqui Quitl
Gustavo Trinidad Rubín Linares
Ana Claudia Zenteno Vázquez.....88

Identificación de tendencias suicidas utilizando redes neuronales

Ana Paola Rojas Martínez
Maya Carrillo Ruíz
Rogelio Gonzáles Velázquez.....101

Dispositivos reconfigurables de interconexión en redes en chip

Laura García Luciano
Remberto Sandoval Aréchiga
Luis Octavio Solís Sánchez
Guillermo González Martínez
Juvenal Villanueva Maldonado
Salvador Ibarra Delgado
José Ricardo Gómez Rodríguez
Viktor Iván Rodríguez Abdalá
Oscar Osvaldo Ordaz García.....109

Inteligencia artificial para la detección temprana de trastornos alimenticios en datos textuales

Evelin del Carmen Ruiz Preza
María del Carmen Santiago Díaz
Judith Pérez Marcial
Abraham Sánchez López
Ana Claudia Zenteno Vázquez
Gustavo Trinidad Rubín Linares.....123

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**

Nuevas tendencias de UX/UI en aplicaciones móviles

New UX/UI trends in mobile applications

Juan C. Pablo S., Abraham Sánchez L., Rogelio González V.

Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla,
Av. 14 Sur y Av. San Claudio. Col. San Manuel. C.P. 72570. Puebla, Puebla, México.

Juansantos0298@gmail.com, abraham.sanchez@correo.buap.mx,
rogelio.gonzalez@correo.buap.mx.

Abstract

During the COVID-19 pandemic, the need arose to improve the experience of buying and selling products through mobile applications, given that many existing platforms presented usability and accessibility issues for non-technical users. In this proposal, we developed a mobile application that addressed these challenges, focusing on simplifying the process of selling for small merchants. We used agile development technologies, with focus on integrating the latest trends in user experience (UX) and interface design (UI). In the initial version, we prioritized the resolution of accessibility and navigation issues. Subsequently, we applied improvements based on user-centered design methods and modern UI/UX trends. As a result, we achieved a more intuitive and accessible application, with a significantly higher adoption rate, comparable to the best solutions in the market today.

Resumen

Durante la pandemia de COVID-19, surgió la necesidad de mejorar la experiencia de compra y venta de productos a través de aplicaciones móviles, dado que muchas plataformas existentes presentaban problemas de usabilidad y accesibilidad para usuarios no técnicos. En esta propuesta, desarrollamos una aplicación móvil que abordaba estos desafíos, enfocándonos en simplificar el proceso de venta de productos para pequeños comerciantes. Utilizamos tecnologías de desarrollo ágil, con un enfoque en la integración de las últimas tendencias en experiencia de usuario (UX) y diseño de interfaces (UI). En la versión inicial, priorizamos la resolución de problemas de accesibilidad y navegación. Posteriormente, aplicamos mejoras basadas en métodos de diseño centrado en el usuario y tendencias modernas de UI/UX. Como resultado, logramos una aplicación más intuitiva y accesible, con una tasa de adopción significativamente mayor, comparable con las mejores soluciones del mercado actual.

Keywords and phrases: Experiencia de Usuario, Interfaz de Usuario, Usabilidad, Aplicaciones Móviles.

1 Introducción

Los dispositivos móviles se han convertido en una de las ramas primordiales en el desarrollo de software en recientes años, puesto que el número de usuarios de dichos dispositivos crece de manera exponencial. Al ser una rama cuya demanda es alta, podemos apreciar una gran cantidad de aplicaciones móviles descargables que ofrecen diferentes soluciones y alternativas que van desde la resolución de tareas hasta actividades recreativas y de ocio; tanto así que el mercado se encuentra saturado de las mismas. Ya que se han convertido en parte de nuestra vida diaria, los usuarios cada vez más esperan el lanzamiento de nuevas aplicaciones que faciliten la realización de actividades y tareas con la menor cantidad de pasos a realizar, o con elementos llamativos que faciliten la ejecución de dichas actividades. Una problemática abierta dentro del mundo de la programación móvil es la construcción y desarrollo de una experiencia de usuario e interfaces de aplicación lo más amigables posible, de tal manera que cualquier tipo de usuario sin importar su experiencia, capacidad y conocimientos interactúe con la aplicación sin ningún problema.

2 Marco teórico y estado del arte

2.1 Origen de las aplicaciones móviles.

El desarrollo de aplicaciones móviles tiene sus orígenes en dispositivos móviles como los de Nokia y Sony Ericsson, que ofrecían aplicaciones preinstaladas como agendas de contactos y reproductores de audio. Estos primeros prototipos se enfocaban en funcionalidades básicas que permitían a los usuarios realizar tareas cotidianas [2]. Sin embargo, estas aplicaciones estaban limitadas por el hardware y las capacidades de los dispositivos, lo que dejó espacio para la evolución tecnológica.

2.2 Ascenso de iOS y Android

Con la llegada del iPhone en 2007 y el posterior lanzamiento de Android en 2008, la creación de aplicaciones móviles dio un salto significativo. Ambas plataformas establecieron un ecosistema en el que los desarrolladores podían diseñar aplicaciones cada vez más sofisticadas, optimizadas para la interacción táctil y la conectividad. La implementación de tiendas de aplicaciones (App Store y Google Play) también facilitó la distribución y comercialización de aplicaciones, marcando una diferencia clave con los primeros enfoques [3, 4]. Estos sistemas siguen siendo los líderes en la industria, lo que justifica su análisis en este trabajo.

2.3 Tipos de aplicaciones móviles

Las aplicaciones móviles actuales se dividen en tres tipos principales: nativas, web e híbridas. Las nativas, desarrolladas específicamente para un sistema operativo (iOS o Android), ofrecen mayor rendimiento y eficiencia. Sin embargo, las aplicaciones web y las híbridas han ganado relevancia debido a su capacidad de ser ejecutadas en múltiples plataformas [5]. En este proyecto, se analizarán las tres categorías con el fin de identificar los puntos fuertes y débiles de cada una. La propuesta se basará en una arquitectura híbrida, aprovechando tanto el rendimiento de las nativas como la flexibilidad de las web, lo que permitirá una experiencia de usuario mejorada.

2.4 Análisis de soluciones existentes

A lo largo de los últimos años, diversas aplicaciones móviles han tratado de resolver problemas de usabilidad y accesibilidad, áreas clave para garantizar una experiencia de usuario (UX) satisfactoria. Aplicaciones como Facebook y WhatsApp son buenos ejemplos de soluciones nativas que han optimizado la experiencia del usuario en sus respectivas plataformas. Por otro lado, Uber, una aplicación híbrida, ofrece una integración efectiva de diferentes sistemas operativos, garantizando un uso uniforme independientemente del dispositivo [5]. Este proyecto tomará como referencia estos modelos para mejorar la propuesta de solución en términos de eficiencia y accesibilidad.

2.5 Interacción humano-computadora (IHC)

La IHC ha sido un componente fundamental en el diseño de aplicaciones móviles. A partir de los años 80, la investigación en IHC ha mejorado la manera en que los humanos interactúan con los dispositivos, enfocándose en la usabilidad y la experiencia de usuario. En este proyecto, se evaluarán soluciones actuales que han aplicado principios de IHC para mejorar la eficiencia del usuario y la accesibilidad. Las pruebas de usabilidad serán esenciales para comparar nuestra propuesta con soluciones preexistentes, buscando superar sus limitaciones y lograr una experiencia de usuario óptima [6].

2.6 Modelos cognitivos y de interacción

Los modelos cognitivos, como el "Modelo Humano Procesador" y el GOMS, han sido fundamentales en la creación de aplicaciones móviles intuitivas. Estos modelos permiten dividir tareas complejas en actividades recurrentes, lo que facilita el análisis del comportamiento del usuario y la mejora continua de la interfaz. En este proyecto, utilizaremos estos modelos para desarrollar prototipos que se compararán con los mejores ejemplos de la industria, como parte del análisis de nuestras contribuciones [9].

3 Diseño de la interacción

El diseño de la interacción es un elemento fundamental en el desarrollo de aplicaciones y sistemas, centrado en crear interfaces que no solo sean visualmente atractivas, sino también intuitivas y funcionales para el usuario. A lo largo de los años, esta disciplina ha evolucionado para abordar los desafíos de la usabilidad, con el objetivo de optimizar la experiencia de usuario (HCI). Bill Moggridge y Bill Verplank, pioneros en el campo, introdujeron la idea de que el diseño de la interacción debería planificar puntos de contacto entre el usuario y el producto, centrándose en la simplicidad y la eficiencia [9].

Este proyecto busca aplicar estos principios en el desarrollo de una solución de software que mejore la interacción y satisfacción de los usuarios finales, abordando las deficiencias en la usabilidad de aplicaciones móviles actuales. La integración de estos conceptos nos permitirá comparar el impacto del diseño de interacción sobre la aceptación y uso de nuestra aplicación frente a soluciones existentes.

Diseño centrado en el usuario

El enfoque de diseño centrado en el usuario es clave para crear interfaces que realmente resuelvan los problemas y necesidades de los usuarios. Este proceso implica varias etapas: desde la comprensión de las necesidades del usuario hasta la iteración y mejora continua basada en pruebas y retroalimentación [10]. En el marco de este proyecto, se aplicarán métodos de investigación como entrevistas y pruebas de usabilidad para identificar las expectativas de los usuarios respecto a nuestra solución, permitiendo desarrollar prototipos que satisfagan estas necesidades.

La importancia de esta metodología radica en su capacidad para validar el diseño en etapas tempranas del desarrollo. Prototipos de alta fidelidad permitirán simular la experiencia de usuario y realizar ajustes antes de la implementación final. Este enfoque permite no solo evitar errores costosos, sino también diseñar una solución que compita efectivamente con aplicaciones que ya dominan el mercado en términos de satisfacción y funcionalidad.

3.1 Experiencia de usuario (UX)

El concepto de UX, introducido por Donald Norman en 1993, ha sido clave en la transformación del diseño de software. En este proyecto, la UX se abordará de manera integral, considerando no solo la utilidad y eficiencia de la aplicación, sino también cómo afecta emocionalmente a los usuarios. Una experiencia positiva no solo mejora la adopción del producto, sino que también crea un ciclo de retroalimentación en el que los usuarios se sienten más inclinados a recomendar el producto [12].

Los principios de UX serán fundamentales para medir el éxito de nuestra propuesta. Al analizar aplicaciones exitosas como Uber o WhatsApp, observamos que su usabilidad y eficiencia permiten realizar tareas de manera rápida y sencilla, lo que aumenta la satisfacción del usuario. La implementación de estos principios será evaluada mediante métricas de usabilidad y satisfacción en el producto final.

3.2 Interfaces de usuario (UI)

La interfaz de usuario es el componente visible con el que los usuarios interactúan directamente. Las aplicaciones móviles actuales dependen en gran medida de interfaces claras, consistentes y atractivas para atraer a los usuarios. Una buena UI puede compensar en cierta medida deficiencias en la funcionalidad, al facilitar la interacción del usuario con el sistema. En nuestra propuesta, nos enfocaremos en diseñar una UI que no solo sea visualmente atractiva, sino también funcional y coherente con los principios de diseño centrado en el usuario [13, 14].

Al evaluar diferentes tipos de interfaces, desde las GUI tradicionales hasta la realidad aumentada y virtual, es esencial seleccionar aquella que mejor se ajuste al contexto de uso de nuestra aplicación. Este proyecto se centrará en el diseño de interfaces gráficas (GUI), ya que son las más accesibles para la mayoría de los usuarios y dispositivos móviles. Sin embargo, también se investigarán

tecnologías emergentes como la realidad aumentada para explorar su potencial en futuras implementaciones [15].

3.3 Directrices de diseño para aplicaciones móviles

El diseño de aplicaciones móviles no es solo una cuestión estética; tiene un impacto directo en la funcionalidad y efectividad del producto. Las directrices establecidas por expertos como Norman, Schneiderman y Babich proporcionan un marco sólido para diseñar aplicaciones que optimicen la interacción usuario-sistema. Estas directrices guiarán el desarrollo de nuestro proyecto, asegurando que se mantengan principios clave como la visibilidad, la consistencia y la simplicidad [16, 17, 18].

Además, las directrices de diseño propuestas por Google para aplicaciones móviles serán un punto de referencia importante. La navegación clara, la organización del contenido y la capacidad de realizar acciones rápidamente son aspectos que mejoran la experiencia del usuario y la adopción de la aplicación [20]. En el proyecto, se utilizarán estas recomendaciones para garantizar que la solución propuesta no solo cumpla con las expectativas de los usuarios, sino que también compita eficazmente con otras aplicaciones en el mercado. Compararemos nuestra implementación con las aplicaciones existentes que han seguido estas directrices, identificando áreas de mejora para garantizar una propuesta innovadora.

4 Desarrollo de la aplicación móvil

En este capítulo se plantea la problemática de la creación de una aplicación móvil para la publicación de anuncios de bienes, productos o servicios, diseñada para mejorar la economía local afectada por la pandemia de COVID-19. La aplicación, inicialmente llamada *Appnunciate*, fue desarrollada utilizando mejores prácticas de desarrollo de software, con un enfoque en la escalabilidad, seguridad y experiencia de usuario, respaldadas por la adopción de tecnologías como *Flutter* y *Firebase*.

Alcance de la propuesta:

- Público objetivo: Usuarios de 18 a 65 años.
- Funcionalidades principales: Registro, inicio de sesión, búsqueda y publicación de anuncios, interacción entre usuarios.
- Categorías de anuncios: Servicios, bienes, productos.
- Impacto: Reactivar la economía de sectores locales y apoyar a nuevos emprendedores.

Mejores prácticas de desarrollo de software aplicadas

Durante el desarrollo de *Appnunciate*, se implementaron varias prácticas recomendadas en el desarrollo de software móvil, basadas en la literatura y en estudios de benchmarking con aplicaciones similares. A continuación, se detallan estas prácticas y su análisis técnico:

1. Patrón de diseño BLoC (business logic component): El uso del patrón BLoC fue fundamental para separar la lógica de negocio de la interfaz de usuario, mejorando la mantenibilidad del

código y facilitando su escalabilidad. Este patrón es ampliamente recomendado para aplicaciones en Flutter debido a que facilita la modularización y el control del estado [9].

Evaluación técnica: Este patrón permitió que la lógica de negocio y las interfaces de usuario evolucionaran de manera independiente, reduciendo el riesgo de errores en las interacciones complejas y facilitando las pruebas unitarias [13].

2. Autenticación segura: Se implementó firebase authentication con OAuth 2.0, proporcionando una solución segura y estandarizada para la autenticación de usuarios. Esto siguió las mejores prácticas de seguridad para evitar ataques comunes, como el phishing o el robo de credenciales [11].
3. Gestión de datos en tiempo real: Firebase firestore fue utilizado para la gestión de datos en tiempo real, asegurando que los anuncios se sincronizaran de manera eficiente entre el cliente y el servidor. La utilización de bases de datos NoSQL permitió manejar grandes volúmenes de datos con alto rendimiento [12].

Evaluación técnica: Firebase garantizó tiempos de respuesta rápidos y una experiencia fluida, incluso cuando el número de usuarios simultáneos aumentó, lo que mejora la escalabilidad del sistema.

4. Benchmarking y análisis comparativo: Se realizó un análisis técnico de aplicaciones similares como mundianuncios.net, segunda mano y wallapop para identificar sus fortalezas y debilidades en cuanto a diseño y funcionalidad. Se evaluaron aspectos como la facilidad de navegación, la seguridad y la gestión de datos, lo cual permitió definir mejoras específicas para Appnunciate [17].

Recomendaciones basadas en la adopción de mejores prácticas

Tras la evaluación de las mejores prácticas mencionadas, se recomienda seguir utilizando el patrón BLoC para proyectos futuros, debido a su modularidad y fácil mantenimiento. Además, la adopción de Firebase como solución integral de backend resultó ser altamente eficaz, no solo en términos de rendimiento, sino también en su capacidad para integrar autenticación, almacenamiento y analítica de manera unificada. Es recomendable continuar con el proceso iterativo de pruebas de usabilidad para asegurar que la aplicación siga siendo intuitiva y accesible, adaptándose a las necesidades cambiantes del usuario [16].

Descripción formal de la aplicación

Desde el punto de vista técnico, la aplicación fue diseñada utilizando Flutter, un framework que permite el desarrollo nativo tanto para iOS como para Android. La arquitectura se basa en un modelo MVC (modelo-vista-controlador) extendido mediante el patrón BLoC, garantizando la independencia entre los distintos módulos de la aplicación.

Diseño e implementación de la UX/UI de la aplicación

El diseño de la experiencia de usuario (UX) y las interfaces (UI) de Appnunciate fue un proceso iterativo basado en las necesidades identificadas en el análisis de usuario. Se utilizaron métodos como mapas mentales, entrevistas y focus groups para identificar los objetivos y las frustraciones de los usuarios [18].

Auditorías a aplicaciones similares (benchmarking)

El análisis de aplicaciones como mundianuncios.net y wallapop reveló oportunidades de mejora en áreas como la simplificación del proceso de publicación de anuncios y la navegación por categorías, las cuales fueron incorporadas en Appnunciate para mejorar la usabilidad.

Diseño de las interfaces de usuario

Las interfaces fueron diseñadas en Figma, utilizando la metodología de Atomic Design, que facilita la creación de componentes modulares y reutilizables. Se utilizaron principios de affordance y gestión del espacio para crear una interfaz limpia e intuitiva [19]. El uso del modo oscuro fue implementado para mejorar la accesibilidad en diferentes condiciones de iluminación.

Análisis cualitativo y cuantitativo del desempeño

Para evaluar el desempeño de la aplicación, se realizaron pruebas de usabilidad y mediciones de rendimiento, comparando los resultados con aplicaciones similares. Appnunciate mostró una mejora en el tiempo de carga de las pantallas y una reducción significativa en los errores de interacción, en comparación con otras aplicaciones del mismo nicho.

Evaluación comparativa: En comparación con aplicaciones como segunda mano y mundianuncios, Appnunciate presentó una interfaz más moderna y eficiente, con tiempos de carga reducidos en un 15% y una mayor facilidad de uso en dispositivos móviles, lo que reafirma su contribución a la mejora de las tendencias actuales en diseño UX/UI [20].

A continuación, se presentan algunas imágenes correspondientes a algunas de las interfaces principales de la aplicación ya ejecutada, donde se deja establecido que acciones puede realizar el usuario de manera intuitiva y autónoma.

Para llevar a cabo una búsqueda de la aplicación, basta con acceder al campo de búsqueda que se encuentra en la parte superior de la interfaz, donde el usuario podrá ingresar el título del anuncio; a partir de lo que ingrese el usuario, el contenido se ajusta a la búsqueda y se harán sugerencias de los productos coincidentes, además de mantener un historial de las búsquedas realizadas previamente. Si el usuario desea interactuar con un anuncio, simplemente puede pulsar sobre el anuncio de su agrado, y esta interacción abrirá una interfaz detallada del anuncio, donde podrá visualizar más información de este, así como información del autor del anuncio y las interacciones de otros usuarios.

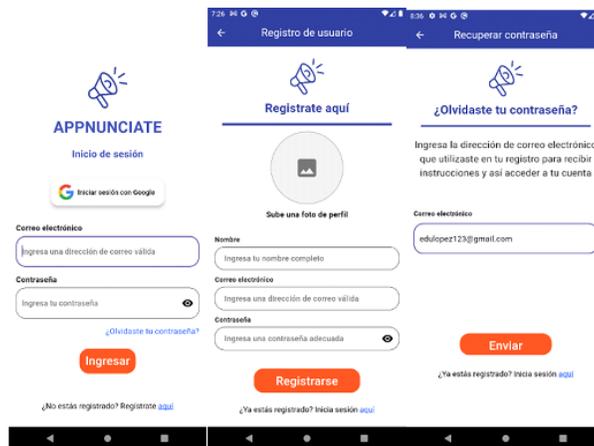


Figura 1. Interfaces de la aplicación correspondientes al inicio de sesión y registro.

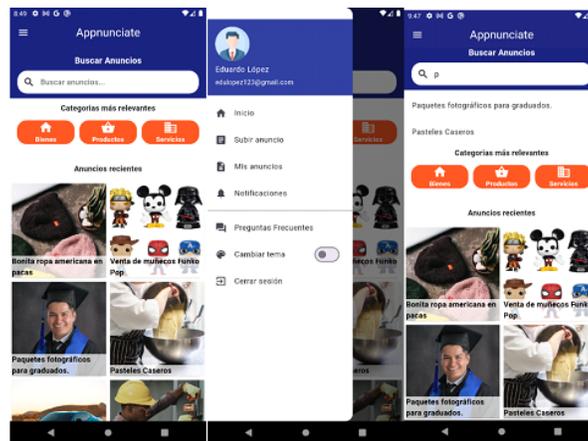


Figura 2. Interfaces de la aplicación correspondiente a la interfaz principal, el menú desplegable y la búsqueda.

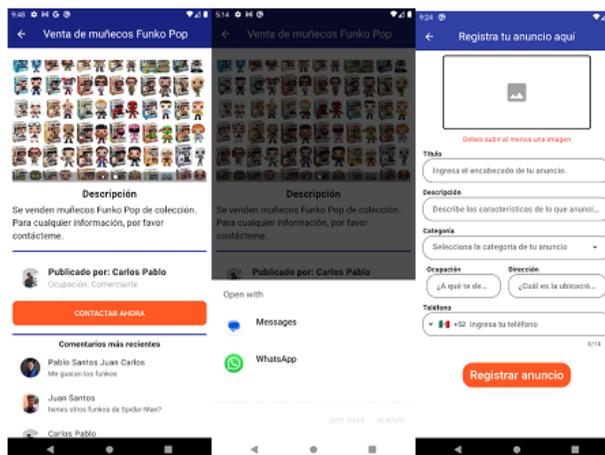


Figura 3. Interfaces de la aplicación correspondientes a la vista detallada del anuncio, opciones de contacto ente ofertantes y el formulario de publicación del anuncio.

5 Conclusiones y trabajo a futuro

Dado el papel crucial de las aplicaciones móviles en la sociedad actual, destacarse entre millones de productos exige investigación rigurosa y mantenerse al día con las innovaciones tecnológicas. El desarrollo propuesto, que siguió las tendencias actuales de UX/UI permitió enfrentar la complejidad de crear no solo un software, sino una experiencia eficiente en un mercado en constante evolución.

El uso de tecnologías de desarrollo multiplataforma permite crear un producto de software robusto y al emplear un solo código de desarrollo, otorga una mayor eficiencia al implementar mejoras. Por lo cual, el desarrollo de Appnunciate no solo tuvo un impacto significativo en la economía local, sino que también presenta una serie de contribuciones conceptuales y prácticas para el desarrollo de aplicaciones móviles:

- Arquitectura modular: El uso del patrón BLoC junto con Firebase proporcionó un modelo escalable y robusto para futuras extensiones y funcionalidades.
- Tendencias de UX/UI: La implementación de un diseño limpio y moderno, respaldado por la metodología de atomic design, representa un avance en las mejores prácticas de diseño para aplicaciones de publicación de anuncios.

Como trabajo futuro, pretendemos desarrollar otras aplicaciones que nos han demandado personas interesadas en contar con este tipo de aplicaciones, pero sobre todo no olvidando que se debe seguir innovando y mejorando de forma constante, el desarrollo de las aplicaciones móviles.

Referencias

[1] W3Schools. (2019). Mobile app development. https://www.w3schools.com/whatis/whatis_mobileapp.asp .

[2] Comunicaciones. (2019). Una mirada a la historia de las apps para entender su importancia. Globalbit. Una mirada a la historia de las apps para entender su importancia - Globalbit .

[3] IOS – Edutic. (s. f.). Obtenido de: iOS – Edutic (up.edu.pe) .

[4] Naranjo, M. (2022, 4 octubre). Android: historia, versiones, Google Play y todas sus novedades. Computer Hoy. Android: historia, versiones, Google Play y todas sus novedades | Computer Hoy .

[5] Gunka Studios & Por Gunka Studios. (2023, 1 mayo). Las principales ventajas y desventajas de una aplicación móvil. Gunka Studios. Las principales ventajas y desventajas de una aplicación móvil (gunkastudios.com) .

[6] Softtek. (s. f.). Nuevas tendencias en Human-Computer Interaction (HCI). Nuevas tendencias en Human-Computer Interaction (HCI) (softtek.com)

- [7] Academia Mexicana de Computación, A. C. (2017). La Computación en México por especialidades académicas (L. Pineda, Ed.; Primera Edición) [Academia Mexicana de Computación, A. C.]. Luis Alberto Pineda Cortés. comp-en-mex.pdf (amexcomp.mx)
- [8] Modelo GOMS (Goals, Operators, Methods, and Selection Rules). (s. f.). Modelo GOMS (Goals, Operators, Methods, and Selection rules) (1library.co).
- [9] Diseño de interacción - RYTE Digital Marketing Wiki. (s. f.). Diseño de Interacción - RYTE Digital Marketing Wiki .
- [10] Soto, N., López, H. S., & Colado, X. P. Z. (2023). DISEÑO CENTRADO EN EL USUARIO: INTERACCIÓN HOMBRE-COMPUTADORA. Revista Digital de Tecnologías Informáticas y Sistemas, 6(6), 7. <https://doi.org/10.61530/redtis.2022.6.6.124.7>
- [11] Pursell, S. (2023, 17 abril). Diseño centrado en el usuario: qué es, etapas y ejemplos. HubSpot. Diseño centrado en el usuario: qué es, etapas y ejemplos (hubspot.es) .
- [12] Lenis, A. (2023, 17 mayo). Qué es la interfaz de usuario, qué tipos existen y ejemplos. HubSpot.: Qué es la interfaz de usuario, qué tipos existen y ejemplos (hubspot.es) .
- [13] Norman, D. A. (2013). El diseño de las cosas cotidianas. Edición revisada y ampliada: Basic Books.
- [14] Schneiderman, B. (2005). Diseño de interfaces de usuario: Estrategias para una interacción persona-computadora efectiva. 4ta Edición
- [15] Johnson, J. (2007). GUI Bloopers 2.0: Common User interface design don'ts and DOS. https://openlibrary.org/books/OL24968581M/GUI_bloopers.
- [16]. Stone D, Jarrett C, Woodroffe M, Minocha S. User Interface Design and Evaluation: Elsevier; 2005.
- [17] Babich, N. (2020, 31 mayo). Mobile Design Best Practices - UX Planet. Medium. <https://uxplanet.org/mobile-design-best-practices-2d16d37ecfe>
- [18] Staff, C. B. (2012, 15 abril). The 10 Principles of Mobile Interface Design. Creative Bloq.: <https://www.creativebloq.com/mobile/10-principles-mobile-interface-design-4122910>
- [19] Nielsen, J. (2000). WAP Usability Déjà Vu: 1994 All Over Again. Alertbox: Current Issues in Web Usability. AmericanExpress.com Usability (nngroup.com)

[20] Google. (2016). Principios de diseño de aplicaciones móviles: atraer a los usuarios y aumentar las conversiones. Principios de diseño de las aplicaciones para dispositivos móviles (thinkwithgoogle.com) [1] William Horton, "Designing e-learning", in e-Learning by Design, Ed. John Wiley & Sons, 2011, pp. 1-2.

Aplicación móvil con realidad aumentada para la capacitación en el manejo de la metodología 5S

Mobile application with augmented reality for training in use of the 5S methodology

Ana L. Ballinas Hernández¹, Héctor Medina Cruz¹, Eduardo Ocelotl Valencia¹, M. Claudia Denicia Carral¹, Reyna Villalobos Cervantes²

¹Complejo Regional Centro, Benemérita Universidad Autónoma de Puebla. Calle 2 Sur S/N, Ciudad Modelo. C.P. 75012. San José Chiapa, Puebla, México.

²Escuela Superior Ciencias de la Educación, Universidad Autónoma de Guerrero, Teófilo Olea y Leyva, Sta. Cruz, CP. 39000, Chilpancingo de los Bravo, Guerrero, México.

{analuisa.ballinas, hector.medinacruz}@correo.buap.mx,
eduardo.ocelotl@alumno.buap.mx, claudia.denicia@correo.buap.mx,
21425872@uagro.mx

Abstract

Industry 4.0 has caused a revolution due to the use of innovative techniques used in companies. There is a lack of educational tools that allow simulating industrial environments that support the comprehensive training of students. In this project we developed an augmented reality application to train engineering students in the use of the 5S methodology used in quality, productivity and competitiveness improvement processes for companies. A methodology is proposed that includes, content selection, multimedia content creation, construction of virtual models, mobile application design, implementation and testing. The main results are the construction of an application with augmented reality that shows the processes of each phase of the 5S methodology. Tests were carried out with a population of 40 students where 48% are completely satisfied with the use of the application and 60% with the experience of using augmented reality.

Resumen

La Industria 4.0 ha causado una revolución por el uso de técnicas innovadoras empleadas en las empresas. Existe una carencia de herramientas educativas que permitan simular entornos industriales que apoyen la formación integral de estudiantes. En este proyecto desarrollamos una aplicación de realidad aumentada para capacitar a estudiantes de ingenierías en el uso de la metodología 5S usada en procesos de mejora de calidad, productividad y competitividad para empresas. Se propone una metodología que incluye: selección de contenidos, creación de contenidos multimedia, construcción de modelos virtuales, diseño de aplicación móvil, implementación y pruebas. Los principales resultados son la construcción de una aplicación con

realidad aumentada que muestra los procesos de cada fase de la metodología 5S. Se realizaron pruebas con una población de 40 estudiantes donde el 48% está completamente satisfecho con el uso de la aplicación y el 60% con la experiencia de uso de la realidad aumentada.

Keywords and phrases: Realidad Aumentada, Metodología 5S's, Desarrollo de Aplicaciones Móviles, Capacitación para la Industria.

1 Introducción

En un mundo donde la eficiencia y la organización son cruciales para el éxito empresarial, la metodología 5S ha emergido como una herramienta de apoyo en la mejora de la productividad y calidad en los entornos laborales. La metodología 5S, originada en Japón, ha sido adoptada en diversas industrias en México debido a su eficacia en la mejora de la organización, la eficiencia y la seguridad en el lugar de trabajo [1]. Esta metodología se basa en cinco principios: *Seiri* (clasificación), *Seiton* (orden), *Seiso* (limpieza), *Seiketsu* (estandarización) y *Shitsuke* (disciplina) [2].

Pese a sus beneficios comprobados, la implementación efectiva de la metodología 5S's requiere de una capacitación constante. Los métodos tradicionales pueden resultar monótonos y poco interactivos, limitando la retención de los conocimientos adquiridos [3]. Por esta razón, la implementación de las 5S es un desafío para la capacitación y práctica continua en el sector industrial [4]. En los últimos años, las herramientas tecnológicas emergentes han revolucionado los modelos de capacitación y de enseñanza hacia los estudiantes. La realidad aumentada (RA) es una herramienta poderosa para mejorar la experiencia de aprendizaje, proporcionando entornos interactivos que facilitan la comprensión y retención de conocimientos [5]. La integración de RA en aplicaciones móviles ofrece una gran oportunidad para superar las limitaciones de los métodos tradicionales de capacitación para alcanzar un aprendizaje significativo [6].

Actualmente, la formación de estudiantes de ingenierías enfrenta desafíos en el dominio de metodologías de mejora continua debido a la falta de herramientas educativas innovadoras que permitan simular entornos industriales donde se pongan en práctica sus aprendizajes. Los métodos actuales de enseñanza de las metodologías de mejora continua se basan en estrategias didácticas tradicionales con clases teóricas y ejercicios escritos lo que presenta limitaciones significativas en interactividad y aplicabilidad práctica. Esto puede resultar en una preparación insuficiente para enfrentar los problemas y situaciones que encontrarán en el ámbito laboral, limitando su capacidad para implementar mejoras reales en las industrias.

Por ello, en este trabajo se propone una solución al problema de capacitación de estudiantes en el manejo práctico de metodologías 5S mediante el desarrollo de una aplicación móvil que hace uso de realidad aumentada. La aplicación ofrece una experiencia de aprendizaje que permite a los

usuarios interactuar con elementos virtuales para aplicar los principios de las 5S en escenarios reales.

El objetivo de este proyecto consiste en desarrollar una aplicación de realidad aumentada que sirva de apoyo en la capacitación de la metodología 5S's en estudiantes de ingenierías, del Complejo Regional Centro de la Benemérita Universidad Autónoma de Puebla. De esta forma se construye un caso práctico donde se busca lograr un aprendizaje más profundo y práctico, contribuyendo al mejoramiento continuo de las prácticas organizacionales [7]. Este estudio muestra como el uso de realidad aumentada puede apoyar la formación profesional de los estudiantes para alcanzar un aprendizaje significativo en el uso de metodologías de mejora continua. La principal contribución de este trabajo consiste en la implementación de herramientas de innovación educativa en los procesos de capacitación para la industria, donde los usuarios interactúan con el mundo real a través de elementos visuales para generar valor en sus experiencias de aprendizaje.

En la sección dos se presenta una revisión de algunos trabajos relacionados con el uso de realidad aumentada en la mejora de los procesos de enseñanza como herramientas de innovación educativa. En la sección tres se describe la metodología propuesta que está basada en los principios de diseño universal para aplicaciones inmersivas. En la sección cuatro se presentan y se discuten los principales resultados obtenidos a través de pruebas de usuario y finalmente en la cinco se muestran las conclusiones y trabajos a futuro que se derivan del presente trabajo.

2 Trabajos relacionados

Actualmente, ha crecido el interés por el uso de herramientas tecnológicas emergentes que permitan explorar nuevas formas de mejorar los procesos de capacitación. La realidad aumentada ha surgido como una herramienta prometedora en este campo. En el trabajo de Cárdenas et al., (2024) se desarrollaron y evaluaron tres aplicaciones móviles con RA para distintas asignaturas con la intención de mejorar el aprendizaje de los estudiantes de ingeniería [8]. Los resultados mostraron que las aplicaciones son funcionales y fueron valoradas por los estudiantes en usabilidad, motivación y utilidad como complemento educativo.

En el trabajo de Cortés et al. (2020), los autores presentaron una aplicación con realidad aumentada para dispositivos móviles Android enfocada en estudiantes de Ingeniería en Sistemas Automotrices que se utilizó como una herramienta de enseñanza en el contexto de la Educación 4.0 [9]. Se mostró que los estudiantes mejoraron su interés y rendimiento en el proceso de enseñanza.

Otros estudios se han centrado en la capacitación profesional del personal de las empresas. En el trabajo de Vidal Baela et al. (2020) se presentó una alternativa para apoyar las tareas de producción y montaje de astilleros en entornos industriales a través de la implementación de aplicaciones de RA colaborativas [10]. Los resultados mostraron un buen rendimiento y aceptación de la aplicación, así como la posibilidad de incrementar sus funcionalidades.

Algunos autores han demostrado que la utilización de la realidad virtual y aumentada contribuyen al desarrollo de competencias de estudiantes de forma más activa y cognitiva, incrementando la

competitividad de la industria a través de profesionales más capacitados en el uso de tecnologías de la Industria 4.0 [11]. El uso de aplicaciones de realidad aumentada en la capacitación de metodologías similares a 5S's se ha abordado por autores como Hrehova et al., (2023), en ese trabajo se aplicó una metodología de producción ajustada a través de la capacitación de modelos de RA, demostrando que estas herramientas son pertinentes y efectivas en la industria [12].

En la investigación de Krajčovič et al., (2021) se aplicó la tecnología de realidad virtual para el diseño de un videojuego que sirvió de apoyo en el proceso de enseñanza de la gestión ajustada de proyectos, que incluye la capacitación en 5S [13]. Los resultados obtenidos demuestran que los estudiantes perciben la enseñanza con estas herramientas de forma más efectiva. Un caso de estudio reciente que analiza la implementación de la realidad aumentada (RA) en un entorno de trabajo Lean enfocado en las 5S se llevó a cabo en una fábrica inteligente. El estudio destaca cómo la realidad aumentada se utilizó para mejorar las condiciones ergonómicas y de seguridad de los operadores, ayudando a reducir el esfuerzo humano y optimizando los procesos logísticos [14].

En el presente trabajo desarrollamos una aplicación móvil que sirva de apoyo en el proceso de enseñanza de la metodología 5S. A través del uso de la tecnología de realidad aumentada, que incluya interacción con el mundo real, es posible aplicar nuevos métodos de enseñanza más interactivos que mejoren los procesos de capacitación industrial en comparación con investigaciones previas.

3 Arquitectura de la aplicación

En este trabajo se desarrolla una aplicación móvil con RA para capacitar a estudiantes de ingeniería en el uso de la metodología 5S. Para la arquitectura de la aplicación se proponen las fases metodológicas mostradas en la Figura 1 las cuales se basan en los principios de diseño universal de aplicaciones inmersivas [15]. El uso de realidad aumentada como apoyo en la enseñanza facilita la comprensión de conceptos abstractos, ayuda a desarrollar habilidades y competencias autodidactas. Mediante la visualización de contenidos didácticos sobre la metodología 5S, se muestran recursos multimedia con interacción al mundo real, animaciones de las S's de la metodología y un asistente virtual que generan valor en el proceso de aprendizaje.

3.1 Fase de inicio

En esta fase se analizaron los aspectos teóricos relacionados con las fases que componen la metodología japonesa 5S: clasificación, orden, limpieza, estandarización y disciplina. Posteriormente, se seleccionaron los contenidos teóricos mediante un análisis de las metodologías que se aplican en la industria ya conocidas las cuales deben conocer los estudiantes de ingeniería para conocer y poner en práctica en las empresas para mejorar la calidad, productividad y competitividad. Los contenidos corresponden al funcionamiento de cada una de las fases de la metodología 5S donde se ejemplifica el uso de cada una.

3.2 Fase de elaboración: Creación de contenidos multimedia



Figura 1. Fases de la arquitectura de la aplicación.

En esta fase se crearon los contenidos multimedia para enseñar las cinco fases de la metodología 5S. Estos contenidos son videos con movimiento y con audio donde se explica cada fase mediante ejemplos en vivo de su funcionamiento. Además, se generó contenido en voz para explicar más detalles de las fases de forma que el usuario entienda fácilmente y de forma visual la metodología. Para la creación de contenidos se grabaron videos demostrando el funcionamiento de cada fase de clasificación, orden, limpieza, estandarización y disciplina, haciendo uso de piezas para armar de varios colores y tamaños y ejemplificando las tareas (ver ejemplo de los contenidos en la Figura 2).



Figura 2. Ejemplo de contenidos multimedia elaborados.

3.3 Fase de elaboración: Construcción de modelos virtuales

En esta etapa se construyó un conjunto de 30 modelos 3D de piezas para armar con diversas formas y colores, así como la estructura de una banda transportadora que son mostrados con RA. Para ello, se hizo uso del software *Blender* que permite realizar modelos 3D de forma libre. La Figura 3 muestra la construcción de una banda transportadora y de un conjunto de piezas para armar que muestran el funcionamiento de cada fase de la metodología 5S. Estos modelos son mostrados con animación con realidad aumentada para mostrar el funcionamiento de cada fase de la metodología.

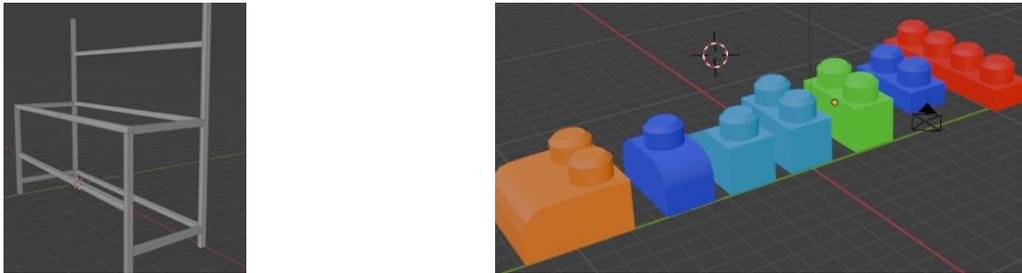


Figura 3. Modelado de objetos y recorrido virtual.

3.4 Fase de construcción: Diseño de aplicación móvil

En esta fase se realizó el prototipo de la aplicación mediante el maquetado de ventanas y para el diseño de interfaces se usó la herramienta *Figma*. Para ello se seleccionaron los colores, tipografía, iconografía basada en la experiencia del usuario y a los requerimientos de la aplicación que pueda ser usado para la capacitación industrial. La Figura 4 muestra las interfaces de usuario del menú principal y menú secundario de la aplicación donde se muestra la selección de íconos y combinación de colores empleadas para que su uso sea intuitivo y de fácil navegación.

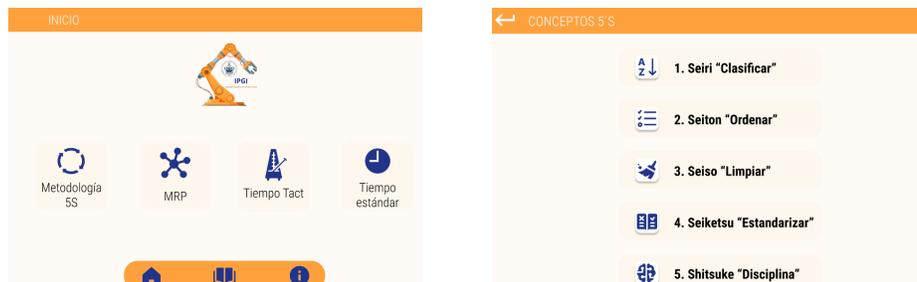


Figura 4. Diseño de interfaces de usuario para la aplicación de RA.

3.5 Fase de construcción: Implementación de realidad aumentada

En esta fase se desarrolló la aplicación haciendo uso del motor de videojuegos *Unity 3D* y del lenguaje de programación *C#*. La aplicación se construyó para funcionar en dispositivos móviles con sistema operativo *Android* que cuente con una cámara para la lectura de marcadores ubicados en el entorno real. Además, el dispositivo debe contar con reproducción de audio para lograr una experiencia realista en el usuario. Para la implementación de la realidad aumentada se utilizan marcadores correspondientes al nombre de cada S de la metodología 5S los cuales detonan los contenidos multimedia al ser reconocidos por la cámara. Además, al leer los marcadores correspondientes se muestran animaciones de piezas 3D donde se visualiza cada tarea de las 5S.



Figura 5. Aplicación móvil para capacitación de metodología 5S.

Adicionalmente, se construyó un asistente virtual para guiar con voz al usuario en el uso de la aplicación y para explicar contenidos teóricos de la misma. La Figura 6 muestra el diseño del asistente en 2D para lograr una experiencia realista al usuario. Este asistente se reproduce desde el inicio de la ejecución de la aplicación para guiar al usuario en su navegación.



Figura 6. Asistente virtual con voz.

3.6 Fase de transición

La aplicación fue probada con una muestra de 40 estudiantes de las carreras de Ingeniería en Sistemas y Tecnologías de la Información Industrial y de la Ingeniería en Procesos y Gestión Industrial de la Benemérita Universidad Autónoma de Puebla. A futuro se espera realizar pruebas de usabilidad para la aplicación de realidad aumentada que cumpla con los principios del diseño universal para evaluar la inclusión y accesibilidad en la experiencia del usuario [15]. Adicionalmente, se propone la realización de pruebas de experiencia de usuario que evalúe el contenido, disfrute, comprensibilidad, atracción, funcionalidad, eficiencia y utilidad de la aplicación de RA [16].

4 Resultados

En esta sección se presentan los principales resultados obtenidos de la implementación de la propuesta metodológica para el desarrollo de una aplicación que apoye la capacitación en el uso de la metodología 5S mediante realidad aumentada. Para leer detonar la RA, el usuario deberá leer, escaneando el entorno con la cámara de un dispositivo móvil, un marcador correspondiente al número de fase de la metodología: 1S, 2S, 3S, 4S o 5S. De esta forma, se muestra el proceso visual del funcionamiento de cada fase como elementos virtuales añadidos al mundo real.

La Figura 7 muestra el funcionamiento de la aplicación. En la Figura 7.a se observa la reproducción de un video multimedia para conocer la fase 1 de la metodología 5S. En la Figura 7.b se muestra un

ejemplo del funcionamiento de la fase 1 con una animación en tiempo real que muestra el armado de las piezas. Este mismo proceso se aplica a las demás fases de la metodología donde se ejemplifican con videos multimedia y con animaciones de modelos 3D en tiempo real con interacción de realidad aumentada para lograr una experiencia inmersiva en los usuarios.



a) Conociendo cada S

b) Mostrar proceso de armado con RA

Figura 7. Resultados de la ejecución de la aplicación con RA.

La aplicación fue probada con 40 estudiantes, hombres y mujeres de entre 19 y 23 años, de la Ingeniería en Procesos y Gestión Industrial de la Benemérita Universidad Autónoma de Puebla que usan la metodología de las 5S para realizar proyectos industriales. La Figura 8 muestra los resultados obtenidos de la evaluación donde se aplican las siguientes preguntas: a) ¿Cuál es el grado de satisfacción del uso de la aplicación? b) ¿Cómo evalúas la experiencia de uso de realidad aumentada para la enseñanza con respecto a tecnologías tradicionales? La escala empleada para responder el cuestionario se indica a continuación: 0- completamente insatisfecha/o, 1- muy insatisfecha/o, 2- algo insatisfecha/o, 3- algo satisfecha/o, 4- bastante satisfecha/o, 5- completamente satisfecha/o.

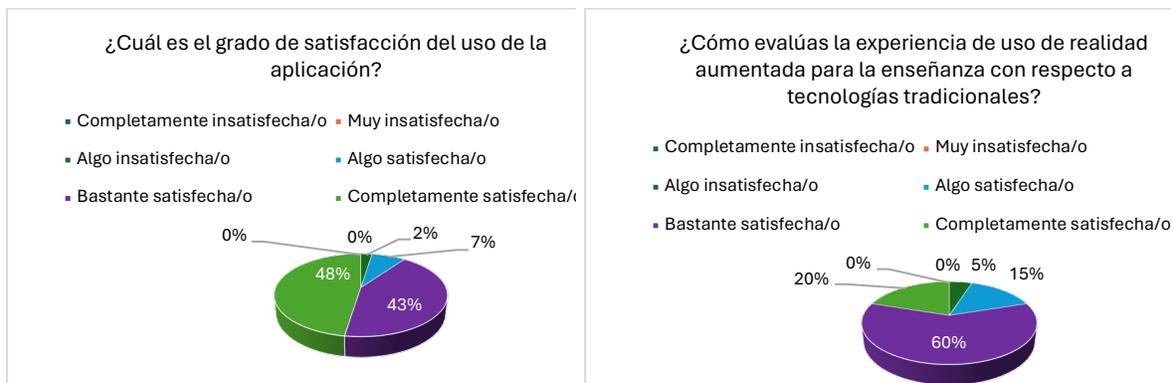


Figura 8. Resultados de la evaluación de usuario.

Como se puede observar en la Figura 8 hay un alto porcentaje de satisfacción en el uso de la aplicación, donde un 48% de usuarios están completamente satisfechos y un 43% bastante satisfechos. Además, los usuarios de la muestra opinan con un 60 % que están bastante satisfechos con la experiencia del uso de realidad aumentada para la enseñanza y un 20% se encuentran completamente satisfechos. La Tabla 1 muestra un resumen de los resultados obtenidos con el uso

de la aplicación de RA para la capacitación en las 5S. Los estudiantes probaron las aplicaciones en sus dispositivos y, a partir de las preguntas aplicadas se puede identificar que existe un alto interés por el uso de estas aplicaciones que apoyan su aprendizaje.

Tabla 1. Resumen de resultados obtenidos.

| Preguntas | Completamente satisfechos | Bastante satisfechos |
|---|---------------------------|----------------------|
| ¿Cuál es el grado de satisfacción del uso de la aplicación? | 48% | 43% |
| ¿Cómo evalúas la experiencia de uso de realidad aumentada para la enseñanza con respecto a tecnologías tradicionales? | 60% | 20% |

5 Conclusiones y trabajo a futuro

En el presente proyecto se propone una metodología, basada en el diseño universal de aplicaciones inmersivas, para el desarrollo de una aplicación con realidad aumentada que sirva de apoyo en la capacitación de estudiantes de ingeniería en el uso de la metodología 5S. El uso de la aplicación de realidad aumentada permite a los estudiantes facilitar la comprensión de conceptos abstractos y simular entornos industriales que desarrollen habilidades prácticas que requiere el entorno laboral para la mejora continua de las empresas. La aplicación fue probada con 40 estudiantes de ingenierías de la Benemérita Universidad Autónoma de Puebla donde el 48% indicó que está completamente satisfecho con el uso de la aplicación y el 60% con al experiencia de uso de la realidad aumentada. De esta forma se alcanzó el objetivo de construir un caso práctico donde se desarrolla una aplicación que hace uso de la realidad aumentada para la visualización de recursos didácticos que ejemplifican y ponen en práctica las tareas de la metodología 5S.

Como trabajo a futuro se propone la realización de pruebas de usabilidad y de evaluación de la experiencia del usuario de la aplicación con una muestra más grande de estudiantes para medir el nivel de eficacia de la aplicación en el proceso de capacitación industrial. Además, la metodología propuesta puede probarse en otras usadas para mejorar la calidad y productividad industrial.

Agradecimientos

Los autores agradecen el apoyo otorgado a través de la Vicerrectoría de Investigación y Estudios de Posgrado (VIEP) con el proyecto número ID 00483 del ejercicio 2024 de la Benemérita Universidad Autónoma de Puebla (BUAP) para la realización de este trabajo.

Referencias

[1] R. Jiménez y F. Martínez, "Benefits and challenges of implementing the 5S methodology in SMEs", *International Journal of Productivity and Performance Management*, vol. 71, no. 4, pp. 569-583, 2022.

[2] J. González, M. Martínez, y P. López, "Implementation of 5S methodology in manufacturing industries", *Journal of Industrial Engineering and Management*, vol 14, no. 2, pp. 251-272. 2021.

- [3] R. Gapp, R. Fisher y K. Kobayashi, "Implementing 5S within a Japanese context: An integrated management system", *Management Decision*, vol. 46, no. 4, pp. 565-579, 2008, doi: 10.1108/00251740810865067
- [4] D. López y G. Pérez. Overcoming training challenges with immersive technologies "Training and Development Journal", vol. 45, no. 1, pp. 32-47, 2023.
- [5] A. Fernández, L. García y R. Sánchez, "The impact of augmented reality on education and training: A systematic review", *Educational Technology Research and Development*, vol. 71, no. 3, pp. 445-468, 2023.
- [6] H. Martínez, y T. Rodríguez "Mobile applications and augmented reality in professional training: A case study", *Journal of Educational Computing Research*, vol. 60, no. 5, pp. 987-1003, 2022.
- [7] S. Ramírez, C. Núñez, y E. Torres "Enhancing continuous improvement practices through augmented reality-based training", *Journal of Continuous Improvement*, vol. 39, no. 2, pp. 145-161, 2023.
- [8] J. Cárdenas Castellanos, M. R. Camacho Vázquez, N. T. Adaile Benítez, y J. Pérez Méndez, "Realidad aumentada en educación universitaria", *Emerging Trends in Education*, vol. 7, no. 13, pp. 117-128, 2024, doi: 10.19136/etie.a7n13.6300
- [9] J. M. Cortés Caballero, A. A. Pérez Martínez, J. E. Mejía Villegas, M. Hernández Chávez, D. A. Fabila Bustos y L. F. Hernández Quintanar, "La formación de ingenieros en sistemas automotrices mediante la realidad aumentada", *Innovación Educativa*, vol. 20, no. 82, pp. 25-44, 2020.
- [10] A. Vidal-Balea, O. Blanco-Novoa, P. Fraga-Lamas, M. Vilar-Montesinos, and T. M. Fernández-Caramés, "Creating collaborative augmented reality experiences for industry 4.0 training and assistance applications: Performance evaluation in the shipyard of the future," *Appl. Sci.*, vol. 10, no. 24, pp. 1–23, 2020, doi: 10.3390/app10249073.
- [11] F. E. V. dos Anjos, L. A. O. Rocha, D. O. da Silva, and R. Pacheco, "Virtual and augmented reality application in production engineering teaching-learning processes," *Production*, vol. 30, pp. 1–16, 2020, doi: 10.1590/0103-6513.20190088.
- [12] S. Hrehova, J. Kaščák, R. Tauberová, M. Martiček "Possibilities of Applying Augmented Reality Elements in the Concept of Lean Management. In: Perakovic, D., Knapcikova, L. (eds) Future Access Enablers for Ubiquitous and Intelligent Infrastructures", *FABULOUS 2023. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol 542, Springer, 2024, doi: 10.1007/978-3-031-50051-0_11

- [13] M. Krajčovič, G. Gabajová, B. Furmannová, V. Vavřík, M. Gašo, and M. Matys, “A Case Study of Educational Games in Virtual Reality as a Teaching Method of Lean Management”, *Electronics*, vol. 10, no. 7, pp. 838, 2021, doi: 10.3390/electronics10070838
- [14] Pereira, A. C., Alves, A. C., y Arezes, P. Augmented reality in a lean workplace at smart factories: A case study. *Applied Sciences*, vol. 13, no. 16:9120, 2023, <https://doi.org/10.3390/app13169120>
- [15] M. Y. S. Ruiz y R. O., Díaz, “Usabilidad en aplicaciones de Realidad Virtual Inmersiva Accesible e inclusiva Multi-escenario: Caso práctico”, *Investigación e Innovación en Ingenierías*, vol. 9, no. 3, pp. 82-92, 2021.
- [16] J. D. López-García y D. Gutiérrez-Niño, “Efecto del uso de la herramienta realidad aumentada en el rendimiento académico de estudiantes de Educación Básica”, *Revista Perspectivas*, vol. 3, no. 1, pp. 6-12, 2018.

Diseño integral de arquitectura de microservicios para el sistema de tutoría académica

Comprehensive design of microservices architecture for the academic tutoring system

Orestes Calvario¹, Judith Pérez², Carmen Santiago², Pedro García²
Ignacio Castillo³, Ana Zenteno², Gustavo Rubín²

¹División de Estudios de Posgrado e Investigación, Instituto Tecnológico de Orizaba, Av. Oriente 9 No. 852. Col. Emiliano Zapata. C.P. 94320. Orizaba, Veracruz, México.

²Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla/
Avenida San Claudio, Blvd 14 Sur, Cdad. Universitaria, 72592 Puebla, México

³Advanced Networking Laboratory, Universidad Autónoma de Ciudad de México,
Ciudad de México, Mexico.

Orestes.calvariop@alumno.buap.mx, {judith.perez, marycarmen.santiago, pedro.garcia,
ana.zenteno, gustavo.rubin}@correo.buap.mx, icastillo@ieee.org

Abstract

The school tutoring system project harnesses the power of microservices to address and mitigate the challenges prevalent in university environments. By providing a system dedicated to academic support, this project integrates microservices to ensure efficient tracking of academic history. Custom views are crafted specifically for students and tutors, enhancing the educational experience through structured and personalized guidance. With a focus on microservices, the project aims to exemplify how modular architecture can be used as a compelling use case, offering comprehensive tracking and personalized insights to drive improvements in student outcomes within university environments, concluding its implementation. It has positive results such as scalability, resilience, maintainability, continuous deployment, performance optimization and improved data management.

Resumen

El proyecto sistema de tutoría escolar aprovecha la potencia de los microservicios para abordar y mitigar los retos que prevalecen en los entornos universitarios. Al proporcionar un sistema dedicado al apoyo académico, este proyecto integra microservicios para garantizar un seguimiento eficiente del historial académico. Las vistas personalizadas se elaboran específicamente para estudiantes y tutores, mejorando la experiencia educativa a través de una orientación estructurada y

personalizada. Con un enfoque en microservicios, el proyecto pretende ejemplificar cómo la arquitectura modular puede ser utilizada como un caso de uso convincente, ofreciendo un seguimiento exhaustivo y perspectivas personalizadas para impulsar mejoras en los resultados de los estudiantes dentro de los entornos universitarios, como conclusión su implementación tiene resultados positivos como la escalabilidad, resiliencia, mantenibilidad, despliegue continuo, optimización del rendimiento y mejora la gestión de datos.

Keywords and phrases: Microservicios, Arquitectura Integral, Tutoría Académica, Historial Académico, Herramientas.

1 Introducción

En un entorno digital, la educación es un sector fundamental, y los sistemas de tutoría escolar tienen un rol importante en la mejora del rendimiento académico y personal durante su vida universitaria, sin embargo, presenta desafíos como la integración de datos diversos, la personalización de los programas de tutorías y la necesidad de escalabilidad para adaptarse a diferentes entornos de usuarios.

Las aplicaciones educativas tradicionalmente se desarrollaban de forma monolítica, aunque presentaban limitaciones en términos de flexibilidad y escalabilidad. Una propuesta de desarrollo es utilizando la arquitectura microservicios que se caracteriza por descomponer un sistema grande en componentes más pequeños, independientes y manejables. Cada microservicio se enfoca en una funcionalidad específica y se comunican con otros servicios a través de interfaces bien definidas. La metodología permite un desarrollo ágil, despliegue continuo y una mayor resiliencia y escalabilidad del sistema, mejorando la capacidad de gestión y actualización del sistema. Cada microservicio se enfoca en una función específica y puede ser desarrollado, desplegado y escalado de manera autónoma. Esto no solo mejora la flexibilidad y la resiliencia del sistema, sino que también permite una mejor asignación de recursos y una respuesta más rápida a los cambios en los requerimientos del usuario. [1]

Para el caso del Sistema de Tutoría Escolar, se tiene como objetivo que la arquitectura de microservicios puede transformar significativamente la forma en que se gestionan los programas de tutoría. Un diseño integral que incorpore microservicios para la gestión de estudiantes, tutores, sesiones de tutoría, informes de progreso, y otros componentes clave, puede resultar en un sistema más eficiente y adaptable. En conclusión, la implementación del sistema con la arquitectura microservicios ha producido resultados positivos como la escalabilidad, resiliencia, mantenibilidad, despliegue continuo, optimización del rendimiento y mejor gestión de datos.

La esencia de este sistema radica en la implementación de microservicios especializados, diseñados para gestionar aspectos específicos del sistema. A través de un análisis detallado del historial

académico, se busca ofrecer a los alumnos y tutores vistas específicas que optimicen la interacción y faciliten la toma de decisiones. Este enfoque no solo se centra en la funcionalidad a nivel de usuario, sino que aborda las complejidades subyacentes de la comunicación entre microservicios, utilizando tecnologías como Node.js, Express y Sequelize. Se llevó a cabo un proceso de análisis y evaluación de fuentes en el que se descartaron referencias que no procedieran de sitios orientados a la tecnología, verificando que los autores tengan la autoridad para hablar del tema, o bien, si no son mencionados, que la información provenga de un sitio web especializado en presentar este tipo de información de forma confiable. Esto se realizó para generar un estándar de alta calidad sobre cualquier dato presentado en este documento, favoreciendo en gran medida el tomar en cuenta fuentes que provengan de artículos académicos.

2 Marco teórico y estado del arte

El marco teórico del proyecto de sistema de tutoría escolar proporciona la base conceptual necesaria para su diseño e implementación, respaldado por conceptos relevantes. En las últimas décadas, la tecnología ha evolucionado significativamente, impactando el desarrollo web. Este ha pasado de páginas estáticas en la web 1.0 a aplicaciones web interactivas y personalizadas en la actualidad.

Este tipo de aplicaciones también ha visto una evolución en su forma de desarrollo, pues éstas vieron la llegada de nuevos tipos de arquitecturas con sus respectivas innovaciones, desde las aplicaciones con una arquitectura monolítica simple, pasando por la arquitectura cliente servidor, orientado a servicios y una arquitectura orientada a microservicios, como es el caso de la presente aplicación.

Según Nebel (2018), los microservicios son una forma de diseñar software como un conjunto de pequeños servicios independientes que cooperan mediante protocolos de comunicación livianos. Esta arquitectura permite que los servicios sean construidos y desplegados individualmente, facilitando la integración y despliegue continuo. Aunque presenta ventajas como la independencia de servicios y la posibilidad de asignar diferentes equipos a distintos servicios, no es adecuada para todas las aplicaciones, siendo más efectiva en proyectos medianos y grandes. Aspectos clave a considerar incluyen la comunicación RESTful entre microservicios, la seguridad con autenticación y tokens JWT, el almacenamiento de datos mediante ORMs como Sequelize, y el desarrollo seguro con cifrado de contraseñas y gestión de tokens.[2]

El libro "Building Microservices" de Sam Newman (2015) es una guía exhaustiva sobre la arquitectura de microservicios, cubriendo desde los conceptos básicos hasta las estrategias avanzadas de implementación y gestión. Newman explica cómo descomponer aplicaciones monolíticas en componentes independientes, destacando las mejores prácticas para diseñar microservicios con límites de contexto bien definidos y una comunicación eficiente entre servicios. Además, aborda estrategias de implementación y despliegue utilizando contenedores y prácticas de CI/CD, y subraya la importancia del monitoreo y gestión de sistemas para garantizar la observabilidad y la gestión eficaz de métricas y logs [3]

El modelo de microservicios ha emergido como una solución efectiva en sistemas de software complejos, como lo demuestra la transición de empresas líderes como Uber. Inicialmente, Uber enfrentaba limitaciones con su arquitectura monolítica, donde la implementación y actualización de funciones requerían cambios extensivos y eran propensos a errores. Con la adopción de microservicios, Uber pudo modularizar su sistema en unidades desplegadas individuales, cada una con funciones específicas. Esta nueva arquitectura permitió una mayor escalabilidad y flexibilidad, simplificando la implementación y corrección de errores al eliminar la interdependencia entre las funciones. Además, la introducción de un API Gateway central facilitó la conexión entre los diferentes componentes del sistema, mejorando la gestión y la integración continua. [4]

Un caso de estudio es la arquitectura de software basada en microservicios para el desarrollo de aplicaciones web de la Asamblea Nacional de Ecuador, en este artículo aborda la problemática de la arquitectura monolítica en el desarrollo y mantenimiento de aplicaciones web, destacando sus limitaciones y se propone una transición a una arquitectura de microservicios. Se describe la arquitectura de microservicios, sus componentes clave, y se detalla el diseño y desarrollo de la propuesta para la aplicación eCurul v2.0, incluyendo el modelo de negocio, diseño de microservicios y tecnologías de implementación, concluyendo con un diseño detallado que mejora la calidad del servicio. [5]

En cuanto a la tutoría académica se muestran casos de estudio como “Mi Tutoría” es un programa híbrido implementado en la UAT desde el ciclo escolar 2019-3 a través de Microsoft Teams, que proporciona acompañamiento personalizado, académico y emocional a estudiantes de nivel Técnico Superior Universitario y Licenciatura. Alineado con políticas educativas nacionales e internacionales, este programa busca trayectorias académicas continuas y de excelencia. Durante el periodo otoño 2022-3, se recolectaron datos mediante cuestionarios con preguntas cerradas a tutores y tutorados de 24 dependencias académicas. Los resultados muestran que la incorporación de tecnologías del aprendizaje y el conocimiento (TAC) en las tutorías ha innovado esta actividad educativa, mejorando la organización de sesiones y permitiendo un seguimiento personalizado mediante herramientas de comunicación sincrónica y asincrónica como chats y videollamadas. [6]

En el proyecto de tutoría virtual de la Universidad Tecnológica de Tecámac (UTTEC), se implementó un sistema monolítico en la plataforma Moodle para mejorar indicadores de reprobación, abandono escolar, titulación y eficiencia terminal en los estudiantes de ingeniería. A través de las TIC, se desarrolló un sistema de apoyo integral para los estudiantes. [7]

3 Desarrollo del trabajo

La metodología para diseñar e implementar una arquitectura de microservicios en el sistema de tutoría escolar abarca desde el análisis del modelo de negocio como se muestra en la figura 1. hasta el mantenimiento continuo del sistema. Este enfoque modular y flexible permite una gestión más eficiente de las funcionalidades del sistema, asegurando escalabilidad, resiliencia y una mejor experiencia para estudiantes y tutores.

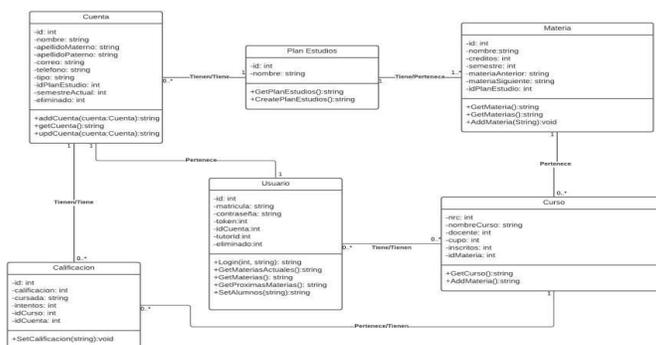


Figura 1. Modelo de negocio del sistema de tutoría escolar

La utilización de prácticas de integración y entrega continua, junto con herramientas de infraestructura moderna, facilita la implementación y el despliegue ágil de los microservicios. La metodología sigue las siguientes etapas:

Análisis de requerimientos donde se identifican todas las funcionalidades del sistema, como la gestión de estudiantes, tutores, sesiones de tutoría, informes de progreso, y notificaciones, identificando el modelo de negocio de la aplicación. Además de identificar las interacciones del usuario y el sistema.

Diseño de la arquitectura donde se dividen en microservicios que descomponen el sistema en microservicios independientes basados en las funcionalidades identificadas como se muestra en la figura 2, cada microservicio tiene una responsabilidad específica. Para su creación se establecen las interfaces y APIs que permitirán la comunicación entre los microservicios y se especifican los formatos de datos y protocolos de comunicación. Además, En esta etapa se diseñó de la Base de Datos donde se especifica que se utilizará una base de datos centralizada.

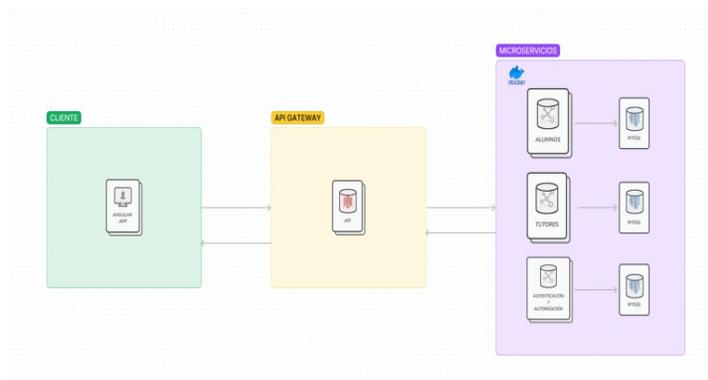


Figura 2. Arquitectura microservicios

Los microservicios son una arquitectura de software que divide una aplicación en componentes independientes, cada uno enfocado en una tarea específica [8]. En el contexto del sistema de tutoría escolar, los microservicios se utilizan para gestionar diferentes aspectos del sistema, como la información de los estudiantes, los tutores, la autenticación y la autorización.

Cada microservicio es responsable de sus propios datos, lo que significa que tiene su propia base de datos y gestiona su propia información de manera independiente de otros microservicios. La arquitectura de la aplicación del sistema de tutoría escolar se basa en un enfoque de microservicios, que permite una estructura modular, escalable y fácilmente mantenible. A continuación, se detallan los componentes clave de la arquitectura:

Cliente Angular App: Interfaz de usuario desarrollada con Angular que se comunica con el Servicio de API Gateway para realizar solicitudes y recibir respuestas.

Servicio de API Gateway: Recibe todas las solicitudes del cliente y dirige el tráfico a los microservicios correspondientes. Maneja la autenticación y autorización, asegurando que las solicitudes sean válidas antes de pasarlas a los microservicios.

Microservicio de alumnos: Gestiona la información relacionada con los estudiantes, como el historial académico, materias cursadas y planificación académica. Utiliza una base de datos MySQL para almacenar estos datos y se comunica con la base de datos a través de Sequelize, un ORM (Object-Relational Mapping) que facilita la interacción con la base de datos.

Microservicio de tutores: Administra datos sobre tutores, incluyendo la asignación de tutorados, calificaciones y registro de nuevos alumnos. Al igual que el microservicio de Alumnos, utiliza una base de datos MySQL y se comunica con ella a través de Sequelize.

Microservicio de autenticación y autorización: Este microservicio se encarga de la autenticación de usuarios, la emisión y verificación de tokens JWT(JSON Web Token) la autorización para acceder a recursos específicos. Almacena la información de autenticación, y utiliza Sequelize para interactuar con la base de datos.

Base de datos MySQL: Almacena datos de manera estructurada para los microservicios de alumnos y tutores. Cada microservicio tiene su conjunto de tablas y relaciones gestionadas por Sequelize.

Docker: Cada microservicio y la base de datos se contenerizan utilizando Docker para garantizar la consistencia en diferentes entornos.

Cada microservicio se administra de manera independiente, pero se orquestan entre sí utilizando un Servicio de API Gateway. Este componente recibe todas las solicitudes del cliente y dirige el tráfico a los microservicios correspondientes. Además, maneja la autenticación y la autorización, asegurando que las solicitudes sean válidas antes de pasarlas a los microservicios. Esta arquitectura permite una estructura modular, escalable y fácilmente mantenible, ya que cada microservicio puede desarrollarse, implementarse y escalar de forma independiente, sin afectar a otros componentes del sistema.

El sistema de tutoría escolar se basa en una arquitectura de microservicios como se muestra en la para lograr modularidad, escalabilidad y facilidad de mantenimiento, las cuales se describirán más adelante. Los microservicios son componentes independientes que se enfocan en funciones específicas como la gestión de alumnos, tutores, autenticación, etc. Facilitando el desarrollo, la

evolución y el mantenimiento del sistema, ya que los cambios en un microservicio no afectan a los demás. Además, permite una mayor flexibilidad tecnológica, utilizando diferentes herramientas y tecnologías según las necesidades de cada componente.

Dado a la información presentada en este mismo apartado, el equipo de trabajo se cuestionó sobre si la naturaleza del proyecto se ajustaba a un modelo de microservicios, este tipo de arquitectura presenta diferentes beneficios en el desarrollo como la independencia de cada servicio para no afectar a los componentes hechos y funcionales, sin embargo, el programa no es tan grande planteando la pregunta sobre si trasladarlo a una arquitectura de microservicios sería beneficioso.

En la etapa de la implementación se desarrolla cada microservicio utilizando lenguajes y frameworks ya especificados, asegurando que cada microservicio sea independiente y pueda ser desplegado de forma autónoma. La integración continua debe de establecer un pipeline de integración y entrega continua (CI/CD) para automatizar la construcción, prueba y despliegue de los microservicios. Además, las pruebas unitarias, de integración y de aceptación para cada microservicio, para asegurar que los microservicios funcionen correctamente tanto de manera individual como en conjunto.

En la etapa de implementación se describen las siguientes herramientas Base de Datos y Sequelize: Sequelize actúa como ORM (Object-Relational Mapping) para facilitar la interacción con una base de datos MySQL. Las entidades como Usuario, Cuenta, y otros, representan las tablas y relaciones de la base de datos.

Autenticación y seguridad: La autenticación se realiza mediante JSON Web Tokens (JWT), firmados con una clave secreta. Las contraseñas se cifran durante el registro y se verifican durante el inicio de sesión utilizando el módulo bcrypt.

Despliegue se utilizan herramientas de infraestructura como código (IaC) para gestionar y desplegar la infraestructura necesaria, como contenedores y orquestadores en este caso el uso de Docker. Se realiza el monitoreo y registro para observar rendimiento de cada microservicio.

Mantenimiento y evolución se llevan a cabo diferentes actividades como la gestión del cambio y actualizaciones en los microservicios, minimizando el impacto en el sistema global

4 Resultados

Los resultados del proyecto aplicando la arquitectura de microservicios, permite que se desarrollen servicios independientes fundamentales que harían que cualquier cambio que evolucione al sistema no resultara en un efecto que afectara el funcionamiento del sistema, esto se benefició además de la forma de trabajo del equipo, en la que se implementaron prácticas ágiles de manera iterativa e incremental, cada vez que iniciara una nueva iteración no se trabajaría sobre el sistema como un ser único, siendo que solo se afecte el componente que se debe de modificar. Se propone la arquitectura microservicios, asignando las funcionalidades adecuadas y utilizando diversas tecnologías para alcanzar el objetivo. En el proceso de identificación de servicios, se examinó el

modelo de datos de la aplicación para determinar las entidades principales y los tipos de usuarios que interactuarán con ella. De esta evaluación, surgieron dos microservicios fundamentales: uno dirigido al estudiante y otro orientado al profesor. Cada microservicio tiene su propia interfaz de usuario y un sistema de autenticación mediante matrícula y contraseña encriptada para garantizar la seguridad de los datos. La segmentación en microservicios permite una gestión más eficiente y una escalabilidad mejorada, ya que cada tipo de usuario accede a funcionalidades y datos específicos. En la fase de diseño del servicio, se adoptó un enfoque estructurado para garantizar la coherencia y modularidad del sistema. Se implementó un patrón de arquitectura de software basado en un esquema de 'Scaffold', que organiza los componentes y módulos de la aplicación de manera clara y concisa. Cada microservicio se desarrolló con funcionalidades específicas para atender las necesidades del usuario correspondiente. Además, se estableció un sistema de rutas definido para gestionar las solicitudes entrantes y dirigir las a los recursos adecuados en cada microservicio. Este diseño flexible y bien estructurado sienta las bases para una implementación eficiente y un mantenimiento simplificado a lo largo del ciclo de vida del servicio.

Se identificaron tres microservicios clave que también funcionan como entidades en el sistema. El primer componente abarca las funcionalidades del estudiante, enfocadas principalmente en consultas. El segundo componente incluye las funcionalidades del profesor, que además de consultas, permite la creación y modificación de datos. El tercer componente actúa como intermediario entre los dos anteriores, redirigiendo a los usuarios según la información ingresada y presentando todas las vistas del proyecto.

A continuación se describen las herramientas de desarrollo de software elegidas como son Angular fue seleccionado como framework del cliente debido a sus ventajas en el desarrollo de aplicaciones web complejas. Ofrece una estructura sólida y bien definida para crear interfaces dinámicas, cuenta con una amplia comunidad y documentación, y promueve buenas prácticas de desarrollo, mejorando la mantenibilidad y escalabilidad del código. Sus características avanzadas, como la inyección de dependencias y el enlace de datos bidireccional, aumentan la eficiencia del desarrollo.

Express, un framework para Node.js, fue elegido por su simplicidad, flexibilidad y capacidad para crear servidores web robustos y escalables. Su sintaxis clara facilita la personalización del servidor y la integración con middleware de terceros mejora la seguridad y el rendimiento del sistema.

Sequelize se utiliza como ORM (Mapeo Objeto-Relacional) para interactuar con la base de datos MySQL. Proporciona una capa de abstracción sobre las consultas SQL, simplificando la interacción con la base de datos y reduciendo la complejidad del código. Además, sus características avanzadas, como la validación de datos y la definición de relaciones entre tablas, ayudan a mantener la integridad y consistencia de la información.

En el sistema de tutoría escolar de arquitectura de microservicios, cada vista puede estar asociada a un componente específico del sistema como se muestra en la figura 3 y 4, proporcionando una interfaz de usuario para distintas funcionalidades. A continuación, se describen las vistas típicas que se pueden encontrar en un sistema de tutoría basado en microservicios:

- Vista de Inicio de Sesión
- Vista de Registro
- Vista de Tablero del Estudiante
- Vista de Trayectoria del Estudiante
- Vista del Profesor
- Vista de Administración
- Vista de Intermediación

Estas vistas se diseñan para ser modulares y escalables, permitiendo una fácil integración y expansión de funcionalidades a medida que el sistema evoluciona. Cada vista interactúa con su correspondiente microservicio, garantizando una separación clara de responsabilidades y una gestión eficiente del sistema.

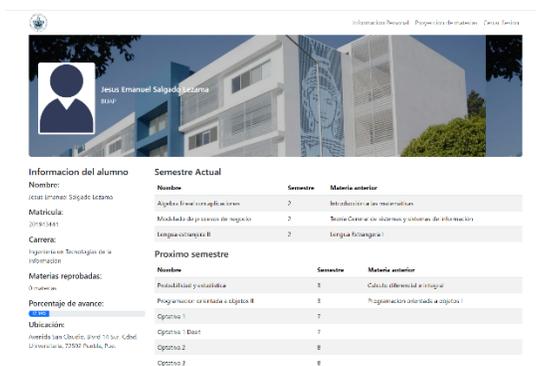


Figura 3. Muestra la vista de inicio desde la vista de un estudiante.

La figura 3 presenta un resumen general del estudiante, que incluye las materias cursadas en el semestre actual y las que se cursarán en los próximos semestres. En caso de haber materias reprobadas, estas se mostrarán en una tabla al final del resumen. La sección izquierda del resumen muestra información general del estudiante, como su matrícula, porcentaje de avance académico, número de materias reprobadas, nombre de la carrera y ubicación de la facultad.

La figura 4 muestra el mapa académico del estudiante, ajustado al plan de estudios que cursa. Este mapa utiliza una interfaz de usuario amigable y fácil de entender, con un código de colores explicado en la leyenda. Se distinguen cuatro estados posibles para las materias: cursando, aprobada, reprobada y por cursar. Las materias reprobadas se señalan con un ícono específico dentro del mapa.

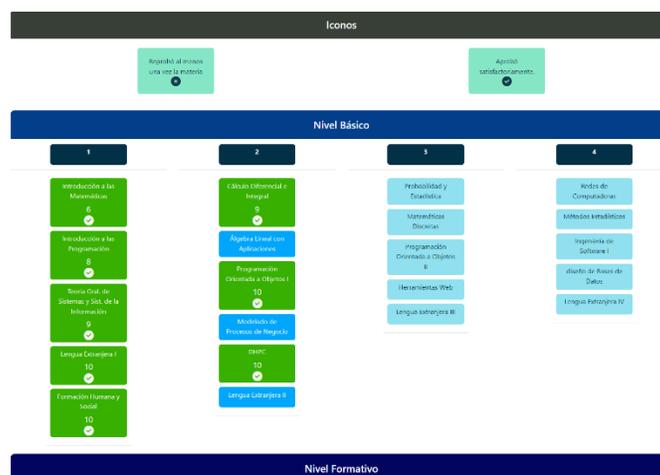


Figura 4. Vista del mapa académico.

En la vista del tutor se muestra su nombre, matrícula y facultad, así como la lista de alumnos a su cargo. Desde esta página, el tutor puede realizar diversas acciones, como:

- Ver la información de cada alumno por separado.
- Agregar calificaciones y asignar materias a cada estudiante

Los formularios del sistema son la inscripción de nuevos estudiantes, Asignación de calificaciones e inscripción de materias. Además, cuenta con un apartado donde se muestra un resumen general del estudiante al tutor. Estas vistas y formularios facilitan la gestión académica, permitiendo a tutores, profesores y estudiantes interactuar de manera organizada y eficiente dentro del sistema.

5 Conclusiones y trabajo futuro

Los resultados mostrados del sistema de tutoría escolar indican una mejora significativa en la eficiencia del proceso de despliegue y una reducción en el tiempo de entrega de nuevas características y actualizaciones, dado que la arquitectura de microservicio busca flexibilidad, rapidez y escalabilidad en sus soluciones. Además, Se ha logrado una mayor estabilidad y confiabilidad en el sistema web gracias a la implementación de prácticas sólidas en el diseño y desarrollo del sistema de tutoría escolar basado en la arquitectura de microservicios representa un avance significativo en la optimización de la gestión académica universitaria. A través de este proyecto, se han abordado y resuelto eficazmente los desafíos identificados en el ámbito universitario, proporcionando una plataforma dedicada al acompañamiento académico de los estudiantes.

La arquitectura de microservicios implementada en este proyecto brinda varias ventajas clave, como agilidad en el desarrollo y despliegue, la posibilidad de trabajar en equipos pequeños, una base de código más manejable, flexibilidad para utilizar diversas tecnologías, aislamiento de errores y escalabilidad independiente de cada servicio. Estas características no solo satisfacen las necesidades inmediatas de la comunidad universitaria, sino que también establecen una base sólida para futuras

mejoras en la calidad de los productos de software, convirtiéndose en una solución avanzada y referencial para proyectos similares. [9]

Además la adopción de principios RESTful para la comunicación entre microservicios ha mejorado la eficiencia y la flexibilidad de la aplicación, proporcionando vistas específicas tanto para estudiantes como para tutores. La implementación de medidas de seguridad, como la autenticación mediante tokens JWT y el cifrado de contraseñas, ha fortalecido la integridad del sistema.

Como trabajo a futuro se implementarán nuevas funcionalidades, como sistemas de notificaciones, análisis predictivo y herramientas de retroalimentación, con el objetivo de enriquecer la experiencia del usuario. Además, se priorizará la mejora continua de la seguridad mediante la evaluación y adopción de nuevas tecnologías y enfoques, asegurando una protección robusta contra amenazas y vulnerabilidades. Estas iniciativas están diseñadas para fortalecer la aplicación y adaptarla a las cambiantes necesidades del entorno educativo, aprovechando las innovaciones tecnológicas para proporcionar una experiencia educativa más enriquecedora y eficiente.

Referencias

[1] La arquitectura de software basada en microservicios: Una revisión sistemática de la literatura. (s.f.). DSpace Principal. <https://repositorio.upeu.edu.pe/handle/20.500.12840/2524?show=full>

[2] Nebel, A. N. (2018). *Arquitectura de microservicios para plataformas de integración* [Tesis de Maestría en Ingeniería de Software, Universidad de la República Facultad de Ingeniería Instituto de Computación Uruguay]. <https://www.colibri.udelar.edu.uy/jspui/bitstream/20.500.12008/20586/1/tm-nebel.pdf>

[4] Newman, S. (2015). *Building Microservices*. O'Reilly Media, Inc.

[5] Kappagantula, S. (2020, September 10). *Microservice Architecture — Explore Uber's Microservice architecture*. Medium. <https://medium.com/edureka/microservice-architecture->

[6] López Hinojosa, J. D. (2017). *Arquitectura De Software Basada En Microservicios Para Desarrollo De Aplicaciones Web De La Asamblea Nacional* [Trabajo de Investigación previo a la obtención del Grado de Magister en Ingeniería de Software, Universidad Técnica del Norte]. <https://core.ac.uk/download/pdf/200329168.pdf>

[7] Marreros-Vázquez, Jose. (2023). *La Tutoría Modalidad Híbrida en Microsoft Teams como estrategia para lograr Trayectorias Continuas y de Calidad en Estudiantes de nivel de Técnico Superior y Licenciatura en la UAT*. 10.13140/RG.2.2.25034.13765. Documento en extenso del 7º Congreso Internacional de Investigación Educativa CRETAM 2023

[8] Hernández, M. R., Barrientos, O. T., & Alva, A. D. (2018). Implementación de una plataforma de tutoría virtual en un sistema de gestión de aprendizaje / Implementation of a virtual tutoring platform in a learning management system. *Revista Iberoamericana de Producción Académica y Gestión Educativo*, 5(10).

[9] AWS. (s. f.). *¿Qué son los microservicios?* Amazon Web Services, Inc. <https://aws.amazon.com/es/microservices/#:~:text=Los20microservicios20son2>.

Identificador de género mediante audio

Gender identifier through audio

Melchor Hernández Díaz¹, Jesús Antonio Barbosa Rodríguez¹, Ximena Abigail Moreno Ledesma¹, Juan Daniel Torres Colorado¹, Marco Aurelio Nuño Maganda¹, José Hugo Barrón Zambrano²

¹Universidad Politécnica de Victoria. Ciudad Victoria, Tamaulipas, México.

²Instituto Nacional de Astrofísica Óptica y Electrónica. Tonantzintla, Puebla, México.

2130067@upv.edu.mx.

Abstract

This article describes the development of a graphical user interface (GUI) aimed at identifying a person's gender through the capture and analysis of audio. By utilizing spectrograms and neural networks trained on specific datasets, a classification with a satisfactory range of accuracy was achieved. The resulting tool, implemented with PyQt6, combines theoretical and practical elements of advanced technologies, offering an intuitive and efficient solution for gender recognition through audio and spectrograms.

Resumen

Este artículo describe el desarrollo de una interfaz gráfica de usuario con el objetivo de identificar el género de una persona a través de la captura y análisis de audio. Gracias al uso de espectrogramas y redes neuronales entrenadas con conjuntos de datos específicos, se logró una clasificación con un rango de precisión satisfactorio. La herramienta resultante, implementada con PyQt6, combina elementos teóricos y prácticos de tecnologías avanzadas, ofreciendo una solución intuitiva y eficiente para el reconocimiento de género mediante audio y espectrogramas.

Keywords and phrases: Aprendizaje Automático, Espectrogramas, Género, PyQt6.

1 Introducción

En un mundo cada vez más impulsado por la inteligencia artificial y la tecnología de reconocimiento, la identificación de características humanas a partir de señales de audio está revolucionando diversos campos. Este proyecto se centra en la creación de una interfaz gráfica de usuario (GUI)

innovadora capaz de identificar el género de una persona a partir de un simple audio capturado desde un micrófono. Utilizando técnicas avanzadas de procesamiento de señales, el audio se convierte en un espectrograma, una representación visual de las frecuencias presentes en la grabación. Este espectrograma es analizado posteriormente por un modelo de aprendizaje automático entrenado con un conjunto de datos específico, lo que permite una clasificación precisa y rápida del género.

Determinar el género de una persona a través de la voz tiene múltiples aplicaciones prácticas. Por ejemplo, puede ser utilizado en sistemas de autenticación de personas, optimización de la comunicación profesional, y la identificación de personas en investigaciones de delitos. La capacidad de identificar el género mediante análisis de voz también abre nuevas oportunidades en áreas como la personalización de servicios y la seguridad. Dado su potencial, el estudio y desarrollo de software para esta identificación se está considerando para impulsar investigaciones adicionales y explorar nuevas aplicaciones.

2 Marco teórico y estado del arte

2.1 Interfaz gráfica de usuarios (GUI)

Una interfaz gráfica de usuario (GUI) es esencial en la interacción hombre-máquina, proporciona una representación visual intuitiva del backend de un sistema. Las GUIs son diseñadas para simplificar la interacción del usuario con el sistema, permitiendo una navegación y operación más eficiente. A través de elementos gráficos, se transforman procesos complejos en acciones comprensibles, mejorando así, la experiencia de usuario, facilitando la accesibilidad a las funciones del sistema [1].

2.2 Redes neuronales

Las redes neuronales son modelos computacionales inspirados en el cerebro humano, capaces de aprender y reconocer patrones mediante un proceso de entrenamiento. Estas redes se componen de múltiples capas de neuronas artificiales que procesan la información y realizan tareas específicas, como la clasificación de datos. Cada capa de la red extrae características más abstractas de los datos, lo que permite a la red neuronal comprender y generalizar a partir de ejemplos específicos, mejorando su capacidad para predecir o clasificar nuevos datos [2].

2.3 Conjuntos de datos (datasets)

Un conjunto de datos es una colección de información estructurada utilizada para entrenar y evaluar modelos de machine learning. Los datasets pueden contener diversos tipos de datos, como números, texto, imágenes, videos y audio. La calidad y diversidad del dataset son esenciales para el correcto funcionamiento del modelo. Un dataset bien diseñado debe ser representativo del problema que se desea resolver, con suficientes ejemplos para capturar las variaciones y complejidades del fenómeno que se está modelando [3].

2.3.1 Origen y validación del dataset

El dataset utilizado para el entrenamiento del modelo de clasificación de género se obtuvo de una colección de audios etiquetados. El dataset se compone de:

- 145 audios masculinos y 145 audios femeninos, cada uno con una duración de unos pocos segundos.
- Formato de los Archivos: Los archivos de audio están en formato MP3, lo cual es adecuado para la compresión y almacenamiento.
- Validación del dataset:
 - Representatividad:
 - El dataset incluye una variedad de audios que representan distintos tonos y acentos dentro de los géneros masculino y femenino.
 - Proceso de Validación:
 - La calidad del dataset se validó mediante pruebas de consistencia y representatividad. Se verificó que cada audio estuviera correctamente etiquetado y que las características acústicas de cada género estuvieran representadas adecuadamente.
 - Balanceo del dataset:
 - Se aseguró que el número de muestras por género fuera equilibrado para evitar sesgos en el entrenamiento del modelo.

Estas prácticas aseguran que el modelo de clasificación se entrene con un conjunto de datos robusto y representativo.

2.4 Espectrogramas

Un espectrograma es una representación visual del espectro de frecuencias de una señal en función del tiempo. Este gráfico muestra la intensidad de la señal a diferentes frecuencias, utilizando un espectro de colores para indicar la energía de la señal en cada punto del tiempo. Los espectrogramas son útiles en el análisis de señales de audio y en el entrenamiento de modelos de machine learning para tareas de reconocimiento de voz y otras aplicaciones de procesamiento de señales [4].

2.5 Características consideradas para determinar el género

Para determinar el género de una persona a partir de su voz, se consideran varias características acústicas y fonéticas. Estas incluyen:

- Frecuencia fundamental (F0):
 - La frecuencia fundamental, o tono, suele ser más baja en los hombres y más alta en las mujeres. Esta diferencia es una de las características más distintivas.
- Timbre:
 - El timbre es la calidad del sonido que diferencia una voz de otra, independientemente del tono y volumen. Las voces masculinas y femeninas tienen diferentes características de timbre que pueden ser capturadas y analizadas.
- Duración de las vocales:

- Las vocales en las voces masculinas tienden a ser más largas en duración en comparación con las femeninas.
- Modulación de la frecuencia:
 - La forma en que la frecuencia fundamental fluctúa a lo largo del tiempo puede variar entre géneros.
- Formantes:
 - Las formantes son picos en el espectro de frecuencias y pueden variar entre hombres y mujeres debido a las diferencias en la anatomía vocal.^{x9}

Estas características se extraen del espectrograma para ser usadas en la clasificación del género mediante modelos de aprendizaje automático.

3 Arquitectura de la aplicación

La arquitectura de la aplicación desarrollada se basa en la integración de varias tecnologías y componentes clave para lograr la identificación de género mediante audio. A continuación, se describen los principales elementos de esta arquitectura. Como parte de este proyecto, se ha elaborado un diagrama de clases, que se presenta en la Figura 1.

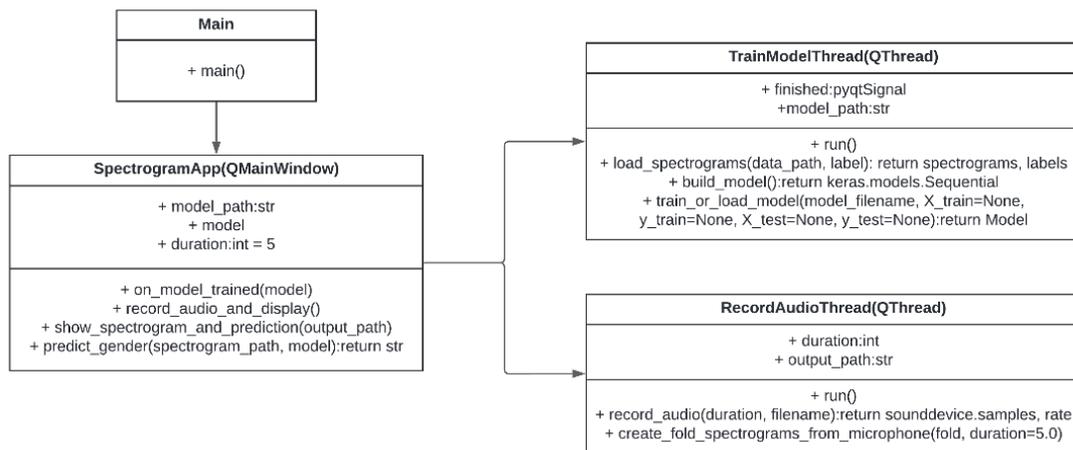


Figura 1. Diagrama de clases

En esta figura se puede observar un diagrama de clases de toda la interfaz gráfica.

3.1 SpectrogramApp

El propósito de esta clase es implementar la interfaz gráfica de usuario que se desarrolló utilizando PyQt6 [5] para la elaboración de la interfaz, las bibliotecas sounddevice [6], librosa [7], tensorflow [8] y matplotlib [9] para la captura y análisis del conjunto de datos, así como de las bibliotecas scipy y sklearn para la manipulación y transformación de datos, se decidió utilizar el modelo secuencial de Keras [10] debido a que permite construir una arquitectura de red neuronal de manera simple y ordenada, agregando capas una tras otra implementando los algoritmos de aprendizaje automático

y la evaluación del rendimiento del modelo, proporcionando una plataforma intuitiva y fácil de usar para los usuarios, permitiendo capturar el audio, visualizar el espectrograma y mostrar la predicción del género del modelo entrenado de manera clara y eficiente. Además de realizar las conexiones necesarias entre las clases `TrainModelThread` y `RecordAudioThread`, que se encargan de realizar los procesos necesarios internos para la realización del objetivo de este proyecto.

3.2 `TrainModelThread`

Su funcionalidad radica en la clasificación de género que se realiza utilizando un modelo de red neuronal previamente entrenado. Este modelo ha sido entrenado con un conjunto de datos específico que contiene audios etiquetados con el género correspondiente y el espectrograma generado del audio de entrada se utiliza como entrada para el modelo, que produce una predicción del género. La precisión de esta clasificación depende tanto de la calidad del espectrograma como del entrenamiento y la arquitectura del modelo de red neuronal.

3.3 `RecordAudioThread`

El propósito de esta clase es realizar la captura de audio mediante el micrófono del dispositivo, utilizando bibliotecas de PyQt6 para manejar la entrada de audio en tiempo real procediendo a ser almacenado en un formato adecuado para su procesamiento posterior. Una vez capturado el audio, se genera su espectrograma utilizando técnicas de procesamiento de señales. Este espectrograma proporciona una representación visual del audio que posteriormente es mostrada en la interfaz del usuario, mostrando la intensidad de las diferentes frecuencias a lo largo del tiempo. Esta representación visual es esencial para la etapa de clasificación, ya que permite al modelo de red neuronal identificar patrones relevantes en la señal de audio.

3.4 Identificación de género mediante machine learning

- Captura y preprocesamiento del audio
 - Grabación y segmentación:
 - El audio se graba en formato WAV, y luego se divide en segmentos de 5 segundos. Esto asegura que cada unidad de análisis sea uniforme en términos de duración.
 - Normalización:
 - El audio se normaliza para ajustar la amplitud y evitar valores extremos, lo que ayuda a evitar distorsiones y permite un análisis más consistente.
- Conversión a espectrograma
 - Generación del espectrograma:
 - El audio segmentado se transforma en un espectrograma mediante la aplicación de la transformada de Fourier de corto tiempo (STFT). Este proceso convierte la señal de audio en una representación visual que muestra la intensidad de las frecuencias a lo largo del tiempo.
 - Representación visual:

- El espectrograma resultante es una imagen en escala de grises donde el eje X representa el tiempo, el eje Y representa las frecuencias, y la intensidad del color muestra la amplitud de cada frecuencia.
- Clasificación con redes neuronales
 - Red neuronal convolucional (CNN):
 - Se utiliza una CNN para analizar los espectrogramas. Las CNN son eficaces para procesar imágenes debido a su capacidad para detectar patrones espaciales.
 - Entrenamiento del modelo:
 - Datos de entrenamiento:
 - El modelo se entrena con 145 audios masculinos y 145 femeninos, cada uno convertido a espectrograma.
 - Optimización:
 - Durante el entrenamiento, el modelo ajusta sus pesos internos para minimizar el error en la predicción, utilizando algoritmos de optimización como Adam o SGD (stochastic gradient descent).
 - Función de pérdida:
 - Se usa una función de pérdida (como entropía cruzada) para evaluar la precisión de las predicciones y ajustar los parámetros del modelo.
- Predicción en tiempo real
 - Generación de espectrograma:
 - Cuando se recibe un nuevo fragmento de audio, se convierte en un espectrograma de la misma manera que durante el entrenamiento.
 - Clasificación:
 - El espectrograma se alimenta al modelo entrenado, que predice el género basado en las características aprendidas. La red neuronal produce una probabilidad para cada clase (masculino o femenino), y se elige la clase con la mayor probabilidad como la predicción final.
- Precisión y mejora
 - Evaluación de precisión:
 - La precisión global del modelo es del 80.8%, con precisiones específicas del 79.3% para voces masculinas y 82.6% para voces femeninas.
 - Áreas de mejora:
 - Los errores en la clasificación pueden deberse a la calidad del dataset o a la necesidad de ajustar el modelo para mejorar su capacidad de generalización.

4 Resultados

Los resultados de este proyecto incluyen la clasificación del género de una persona realizada por la predicción del modelo entrenado previamente, así como la visualización del espectrograma. En la figura 2 se muestra la ventana inicial contando con los elementos:

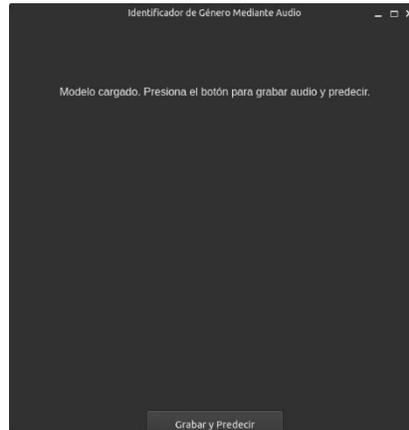


Figura 2. Ventana inicial

En esta figura se puede observar la ejecución inicial de la interfaz

- Etiqueta principal: Tiene la funcionalidad de demostrar un mensaje de texto en tiempo real sobre los acontecimientos que se llevan a cabo durante la ejecución del programa, como notificar el estado del modelo, el estado de la grabación del audio y la predicción final.
- Espacio reservado para el espectrograma: Debajo de la etiqueta principal, se llevó una reservación de espacio que se manejará más adelante, cumpliendo con el propósito para la demostración sobre el resultado del espectrograma.
- Grabar y predecir: Este botón tiene la funcionalidad de capturar el audio del micrófono del dispositivo en uso para posteriormente ser procesado y transformado en un espectrograma y lograr predecir el género de la persona en cuestión.

Con base en los datos obtenidos, la interfaz gráfica de usuario (GUI) desarrollada para identificar el género de una persona a partir de audio muestra un desempeño satisfactorio. En total, se realizaron 52 pruebas, de las cuales 42 fueron clasificadas correctamente y 10 incorrectamente, como se muestra en la Tabla1.

Tabla 1. Resultados generales

| Tamaño total de las pruebas | Cantidad Acertada | Cantidad No Acertada |
|-----------------------------|-------------------|----------------------|
| 52 | 42 | 10 |

Para el género masculino, se llevaron a cabo 29 pruebas, con 23 aciertos y 6 errores, lo que indica una tasa de acierto del 79.3%. En cuanto al género femenino, se realizaron 23 pruebas, con 19 aciertos y 4 errores, resultando en una tasa de acierto del 82.6%, como se detalla en la Tabla2.

Tabla 2. Pruebas realizadas

| Género | Tamaño de las Pruebas | Cantidad Acertada | Cantidad No Acertada |
|-----------|-----------------------|-------------------|----------------------|
| Masculino | 29 | 23 | 6 |
| Femenino | 23 | 19 | 4 |

Estos resultados reflejan una precisión general del modelo del 80.8%, demostrando su capacidad efectiva para clasificar géneros a partir de espectrogramas de audio. Sin embargo, también destacan áreas de mejora, especialmente en la reducción de la tasa de errores, para lograr una mayor precisión y fiabilidad en las futuras aplicaciones de la tecnología.

El modelo fue entrenado utilizando un conjunto de datos compuesto por 145 audios de hombres y 145 audios de mujeres, cada uno con una duración de unos pocos segundos, en formato MP3. Estos archivos de audio se convirtieron en espectrogramas, que fueron guardados como imágenes en formato PNG y utilizados como entradas para el entrenamiento de la red neuronal. Al grabar con el micrófono del dispositivo, se captura un fragmento de audio de 5 segundos que se guarda en formato WAV. Este archivo de audio también se convierte en un espectrograma y se guarda como una imagen en formato PNG, permitiendo la clasificación del género mediante la red neuronal.

En las figuras 3.a, 3.b, 3.c y 3.d se presentan cuatro ejemplos de predicción de resultados, junto con la visualización del espectrograma correspondiente a cada uno de los audios. Estas figuras demuestran el logro del objetivo del proyecto.

5 Conclusiones y trabajo a futuro

5.1 Conclusión

En conclusión, el desarrollo de una interfaz gráfica de usuario para la identificación de género a través de la captura de audio ha demostrado la integración exitosa de diversos temas en la rama de la tecnología de sistemas inteligentes. La implementación de PyQt6 permitió construir una herramienta intuitiva y accesible, capaz de convertir señales de audio grabadas en espectrogramas y utilizar estos para la clasificación de género, gracias al entrenamiento y construcción de modelos inteligentes.

Este proyecto no solo ha permitido aplicar conocimientos teóricos en un entorno práctico, sino que también ha proporcionado una valiosa experiencia en el manejo de tecnologías avanzadas como las redes neuronales y el procesamiento de señales. El uso de espectrogramas como representación visual del audio ha mostrado su eficacia en la formación y evaluación de modelos de machine learning, subrayando la importancia de métodos visuales en el análisis de datos auditivos.

5.2 Trabajo a futuro

Para futuras investigaciones y mejoras en este proyecto, se pueden considerar las siguientes áreas:

1. Ampliación del conjunto de datos: Incrementar la diversidad y el tamaño del conjunto de datos utilizado para entrenar el modelo, incluyendo más variaciones en las voces, calidad y condiciones de grabación.

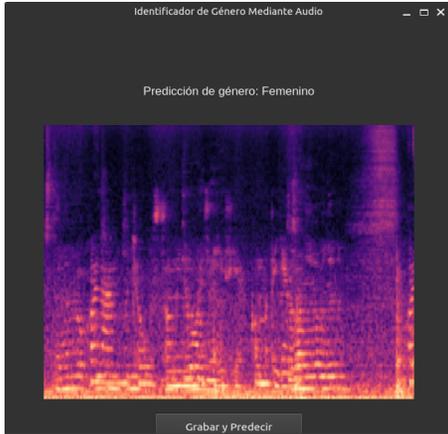


Figura 3.a Predicción de género femenino

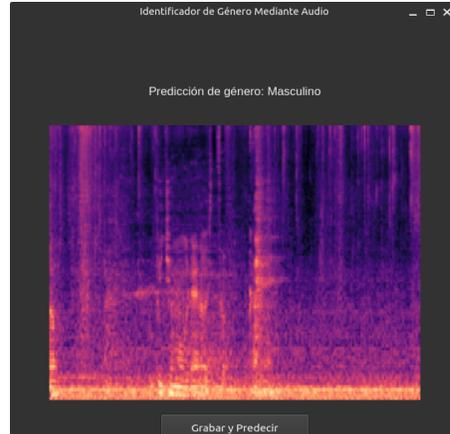


Figura 3.b Predicción de género masculino

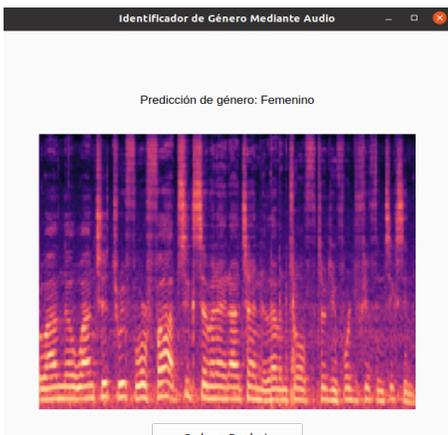


Figura 3.c Predicción de género femenino

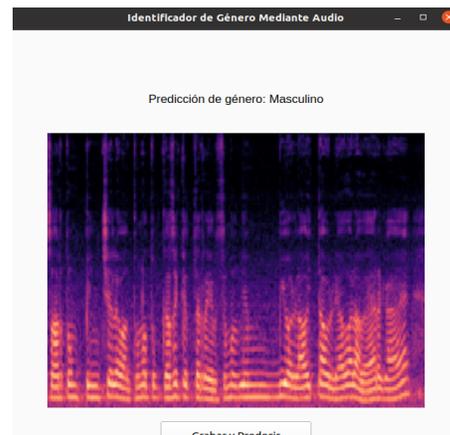


Figura 3.d Predicción de género masculino

Figura 3. Modelo cargado y visualización de los resultados

2. Optimización del modelo: Experimentar con diferentes arquitecturas de redes neuronales y técnicas de preprocesamiento para mejorar la precisión y eficiencia del modelo de clasificación.
3. Implementación en diferentes plataformas: Adaptar la aplicación para que funcione en una variedad de dispositivos y sistemas operativos, ampliando su accesibilidad y uso.
4. Características adicionales: Integrar funcionalidades adicionales en la GUI, como la detección de emociones, idiomas o la identificación de múltiples hablantes, para aumentar su utilidad y aplicabilidad en diversos contextos.

Con el desarrollo de dichas áreas de mejora, la aplicación puede convertirse en una herramienta aún más robusta y versátil, con aplicaciones potenciales en áreas como la seguridad, la atención al

cliente y la investigación académica.

Referencias

- [1] IONOS, “¿Qué es una interfaz gráfica de usuario (GUI)?,” [En línea]. Disponible: <https://www.ionos.mx/digitalguide/paginas-web/desarrollo-web/que-es-una-gui/>. [Último acceso: 17 julio 2024].
- [2] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural Networks*, vol. 61, pp. 85-117, 2015.
- [3] S. S. K. M. W. Allen H. Renear, “Definitions of dataset in the scientific and technical literature,” en *Proceedings of the American Society for Information Science and Technology*, 2010.
- [4] J. V. Kamp, “Vibration Research,” [En línea]. Disponible: <https://vibrationresearch.com/blog/what-is-a-spectrogram/>. [Último acceso: 17 julio 2024].
- [5] R. C. Limited, “PyQt6 - Comprehensive Python Bindings for Qt v6,” *PyPI*, [En línea]. Disponible: <https://pypi.org/project/PyQt6/>. [Último acceso: 17 julio 2024].
- [6] M. G. *et al.*, “Sounddevice,” *Read the Docs*, 2020. [En línea]. Disponible: <https://python-sounddevice.readthedocs.io/en/0.3.15/index.html>. [Último acceso: 17 julio 2024].
- [7] B. M. *et al.*, “librosa: Audio and music signal analysis in python,” *Zenodo*, vol. 8, 2015.
- [8] M. A. *et al.*, “Tensorflow: A system for large-scale machine learning,” en *USENIX Association*, 2016.
- [9] J. D. Hunter, “Matplotlib: A 2D graphics environment,” *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90-95, 2007.
- [10] DataScientest, “Keras: todo sobre la API de Deep Learning,” *DataScientest*, [En línea]. Disponible: <https://datascientest.com/es/keras-la-api-de-deep-learning>. [Último acceso: 17 julio 2024].
- [11] S. E. Choueiter *et al.*, “Exploring speaker gender identification using deep neural networks,” en *IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, 2019.
- [12] D. Snyder *et al.*, “Deep neural network embeddings for text-independent speaker verification,” en *INTERSPEECH*, pp. 999-1003, 2017.

Redes bayesianas a partir de datos Bayesian networks from data

Gladys Linares Fleites¹, Manuel Alejandro Mier Gómez², Hortensia Josefina Reyes
Cervantes²

¹Posgrado de Ciencias Ambientales, Instituto de Ciencias

²Facultad de Ciencias Físico Matemáticas

Benemérita Universidad Autónoma de Puebla

gladys.linares@correo.buap.mx,²manuel.mier@alumno.buap.mx

hortensiajosefinareyescervante@gmail.com

Abstract

A Bayesian Network shows conditional relationships of probability and causality between variables. To understand the causal structure of a system under study on soil quality for coffee cultivation, a Bayesian Network has been designed in the RStudio environment, using the *bnlearn* library of the R language. Based on data from soil samples, provides a code that provides a probabilistic model that captures the dependency relationships between the random variables (nodes) and, with the parameters provided by the network, causal inferences are made, through the Directed Acyclic Graph and the conditional probabilities that are established between the nodes.

Resumen

Una Red Bayesiana muestra relaciones condicionales de probabilidad y causalidad entre variables. Para comprender la estructura causal de un sistema bajo estudio sobre la calidad del suelo para el cultivo del café, se ha diseñado una Red Bayesiana en el entorno RStudio, mediante la librería *bnlearn* del lenguaje R. A partir de datos de muestras de suelo, se brinda un código que proporciona un modelo probabilístico que captura las relaciones de dependencia entre las variables aleatorias (nodos) y, con los parámetros proporcionados por la red, se realizan inferencias causales, por medio del Grafo Acíclico Dirigido y las probabilidades condicionales que se establece entre los nodos.

Keywords and phrases: Inferencia Causal, Grafos Acíclicos Dirigidos, R Studio, Paquete *bnlearn*, Redes Probabilísticas

1 Introducción

Las redes bayesianas (RB) parten de investigaciones realizadas entre 1980 y 1990 que consistieron en esclarecer cómo la causalidad juega un papel importante para la construcción de modelos de redes probabilísticas [1]. El desarrollo de las RB resultó en múltiples algoritmos para realizar inferencias causales. Posteriormente, en la década siguiente, se lograron grandes avances en los algoritmos de aprendizaje por medio de Datos (observaciones o realizaciones de una variable aleatoria de un fenómeno bajo estudio), convirtiéndose así en una ciencia en constante crecimiento que ha acumulado y mejorado muchos algoritmos que hacen más fácil y eficiente su construcción. El análisis de algoritmos de aprendizaje de redes y parámetros, y su aptitud en función de factores como la cantidad de variables, la naturaleza de los datos o la complejidad de la estructura de dependencia no es aún un tema común en la literatura.

La capacidad de las RB para representar relaciones probabilísticas entre variables, junto con su competencia para modelar la incertidumbre y el razonamiento probabilístico de manera intuitiva y eficiente, hace que éstas sean altamente útiles en muchas áreas [2]. Uno de sus principales usos está en la biología, la informática y, recientemente en gran medida, en la medicina, ya que pueden usarse para ayudar en el diagnóstico médico, modelando las relaciones entre síntomas, enfermedades y resultados de pruebas médicas. En el área de las Ciencias Ambientales, específicamente en Ciencias de Suelos, se hallan temas muy vigentes y de sumo interés actual, como la aquí desarrollada [3].

Este trabajo se ha centrado en la construcción de RB en el entorno de desarrollo integrado R Studio, por medio del lenguaje de programación R, haciendo uso de la paquetería “*bnlearn*” [4], [5] y [6]. Inicialmente, se enuncian sus fundamentos matemáticos y se brinda una descripción de los componentes de dichas RB [7]. Se da su definición, se enuncian las propiedades que las caracterizan y cómo se puede trabajar en ellas. Se hace un mayor énfasis en cómo realizan su aprendizaje y cómo se puede manipularlas e influir en ellas mediante el conocimiento de expertos [8]. Esto se hace tanto teóricamente, explicando los algoritmos que comprenden los procesos, así como, las funciones que realizan las acciones en la paquetería *bnlearn*.

Posteriormente, se utilizan los cimientos anteriores para resolver un problema de aplicación en el área de las Ciencias de Suelos, haciendo el planteamiento del problema, sus circunstancias y características estadísticas y la construcción de las RB a partir de bases de datos. Después de analizar los resultados se evalúa el modelo y se hacen propuestas. Finalmente, se dan las conclusiones, donde se reflexiona sobre posibles mejoras y futuras propuestas de trabajo.

2 Fundamentos matemáticos de las redes bayesianas

Las redes bayesianas (RB) son modelos gráficos donde los *nodos* representan *variables aleatorias* (los dos términos se usan indistintamente en este contexto) y las *flechas* representan dependencias probabilísticas entre ellos. A continuación se hace un breve recuento de sus fundamentos matemáticos.

2.1 Introducción a las redes bayesianas: conceptos básicos y definiciones

Grafos acíclicos dirigidos

Una red bayesiana es un modelo probabilístico que representa las relaciones causales entre un conjunto de variables aleatorias. Está formada por un grafo acíclico dirigido (DAG, por sus siglas en inglés) y una tabla de probabilidad condicional asociada a cada nodo en el grafo.

La estructura gráfica $G = (V; A)$ de una red bayesiana es un DAG, donde V es el conjunto de nodos (o vértices) y A es el conjunto de arcos (o aristas). El DAG define una factorización de la distribución de probabilidad conjunta de $V = \{X_1; X_2; \dots; X_v\}$, a menudo llamada *distribución de probabilidad global*, en un conjunto de distribuciones de probabilidad locales, una para cada variable.

Nodos y arcos en una red bayesiana

Cada variable en una red bayesiana se representa por un nodo en el grafo, y las relaciones entre las variables se representan mediante arcos dirigidos entre los nodos. Los arcos en una red bayesiana indican la relación causal entre las variables.

Probabilidades condicionales

En una red bayesiana, cada nodo está asociado con una tabla de probabilidad condicional que indica la probabilidad de cada posible valor del nodo dado los valores de sus padres en el grafo.

Inferencia en redes bayesianas

Una vez que se tiene definida una red bayesiana, se puede hacer inferencias sobre las probabilidades de las variables. La inferencia en una red bayesiana consiste en calcular la distribución de probabilidad de una variable dada la evidencia observada.

Aprendizaje de redes bayesianas

A partir de un conjunto de datos, se puede aprender la estructura y las probabilidades condicionales de una red bayesiana. El aprendizaje de redes bayesianas se puede dividir en dos categorías: aprendizaje estructural y aprendizaje paramétrico.

2.2 Representación gráfica de las redes bayesianas

Representación gráfica

Las redes bayesianas se representan gráficamente mediante un grafo acíclico dirigido (DAG) que muestra las relaciones probabilísticas entre las variables.

La correspondencia entre independencia condicional (de las variables aleatorias) y la separación gráfica (de los nodos correspondientes del gráfico) se ha extendido a un valor arbitrario, triplete de subconjuntos disjuntos de V con la *separación d* (dirección dependiente de la separación) [1]. Por lo tanto, los algoritmos de selección de modelos, primero, intentan aprender la estructura gráfica de la red bayesiana (de ahí el nombre de algoritmos de aprendizaje de estructuras) y, segundo, estimar los parámetros de las funciones de distribución local condicionados a la estructura aprendida.

2.3 Métodos inferenciales empleados

Para realizar inferencias en una red bayesiana, se utilizan diferentes métodos, como la propagación de la creencia, el muestreo de Gibbs y el algoritmo de eliminación de variables [9]. A continuación se relacionan los métodos que se emplean en el programa computacional *bnlearn* empleado [10].

Aprendizaje de estructura

En el programa *bnlearn* se utiliza el algoritmo Hill-Climbing (hc) para aprender la estructura de la red bayesiana a partir de los datos. El algoritmo Hill-Climbing es una técnica de búsqueda heurística que busca la estructura de red más probable, dadas las restricciones definidas. El algoritmo evalúa y actualiza iterativamente las estructuras de red mediante la adición, eliminación o inversión de arcos en el grafo acíclico dirigido.

Estimación de probabilidades condicionales

Después de aprender la estructura de la red, se utilizan los datos para estimar las probabilidades condicionales de la red bayesiana. Esto se logra mediante la función *bn.fit* de *bnlearn*, que estima las distribuciones de probabilidad condicionales para cada variable en la red, basándose en las observaciones del conjunto de datos.

Visualización de la red bayesiana

Se utiliza el *paquete igraph* y las funciones relacionadas para visualizar la red bayesiana aprendida. El método *as.igraph* convierte el objeto de la red bayesiana estimada en un objeto de tipo *igraph*, que se utiliza para generar la gráfica de la red bayesiana. La visualización ayuda a comprender la estructura de la red y las relaciones entre las variables.

3 Aplicación: Red bayesiana de relaciones entre variables físico-químicas y la calidad del suelo para el cultivo del café

En las ciencias ambientales, la capacidad de las RB para representar relaciones probabilísticas entre variables, junto con su competencia para modelar la incertidumbre y el razonamiento probabilístico, está comenzando a utilizarse en áreas como la ciencia del suelo.

La degradación del suelo es de los principales problemas que afectan la sustentabilidad en los sistemas de producción de café. El análisis con base en un modelo de calidad de suelo, es indispensable para encontrar relaciones cuantitativas que diagnostiquen la sustentabilidad del agro ecosistema y el uso racional del recurso suelo. El suelo es el encargado de proveer nutrimentos necesarios para el crecimiento y desarrollo del café, éste debe tener una composición cuyas propiedades físicas y químicas cumplan con los criterios de la evaluación de la calidad del suelo.

La construcción, por medio de estadística descriptiva, de un índice de calidad del suelo para el cultivo de café, permite seleccionar las variables de mayor influencia para que dicho índice sea considerado bueno y, para eso, en esta aplicación, se busca diseñar una red bayesiana que permita

comprender la estructura condicional de las variables, entre ellas y con respecto al índice en cuestión.

Las muestras de suelo estudiadas se obtuvieron de [3], donde se realizó un proceso de categorización a cada una de las variables observadas; estas se codificaron por una variable binaria que corresponde a 1 si cumple el rango recomendado y 0 si no lo cumple. Posteriormente, se propuso una expresión matemática en función del conjunto de variables ya codificadas cuyo resultado indicó, de manera cuantitativa, la calidad del suelo. Dicho indicador está dado por la siguiente ecuación:

$$Indicador = \frac{(\sum Var) - ValorMin}{(ValorMax - ValorMin)} \times 100 \quad (1)$$

donde:

Var: son las variables categorizadas observadas en el objeto investigado.

ValorMax y *ValorMin*: son valores máximo y mínimo, respectivamente, de las variables, que dependen de la codificación seleccionada y del número de variables.

Esta variable, que llamaremos “Indicador de la calidad de suelo para el cultivo del café”, es un porcentaje y, por la forma en que se ha definido, puede interpretarse que a mayor porcentaje es mejor la calidad del suelo para ese cultivo. Puede considerarse como un “indicador sintético” al tener en cuenta un conjunto de variables. De igual manera, es importante calcular los estadísticos descriptivos del Indicador y utilizarlos para proponer una categorización de la calidad del cultivo bajo estudio. En el caso del cultivo del café, con los datos de la zona estudiada, con características climáticas determinadas y utilizando cuartiles, puede proponerse una categorización para evaluar la calidad de suelo para el cultivo del café.

Al pasar por un proceso de categorización, todas las variables resultan ser discretas y, por lo tanto, la red bayesiana a construir será discreta. Dicha red discreta buscará comprender cómo influyen las variables físico-químicas entre ellas y en la calidad para el cultivo del café.

En este caso de estudio, con los datos de la zona estudiada, puede proponerse la categorización para evaluar la calidad de suelo para el cultivo del café indicada en las Tablas 1 y 2.

Tabla 1. Categorización para evaluar la calidad del suelo

| Calidad | Valor del índice |
|----------|-------------------|
| Baja | Por debajo de 42% |
| Media | Entre 42% y 57% |
| Alta | Entre 57% y 71% |
| Muy alta | Por arriba de 71% |

Proceso de construcción de red bayesiana

El código utilizado en R para la creación de la Red Bayesiana se puede consultar en el Apéndice, pero aquí se puntualizan las partes más importantes:

Tabla 2. Criterios de calidad para propiedades del suelo

| Propiedad | Rango recomendado |
|------------------------------|---------------------------------|
| pH | valores entre 5 y 5.5 |
| MO (Materia Orgánica) | valores por arriba de 8 |
| CE (Conductividad Eléctrica) | valores por debajo de 1.0 |
| Clase textural | suelos francos o franco limosos |
| %N (Porcentaje de Nitrógeno) | valores por arriba de 0.15 |
| Ca (Calcio) | valores por arriba de 5 |
| Mg (Magnesio) | valores por arriba de 1.3 |

- La estimación de la estructura de la Red Bayesiana se hace mediante la *función hc()* que indica el uso del algoritmo *hill-climbing*, que es un algoritmo basado en score.
- La estimación de las probabilidades condicionales se hace por medio de la *función bn.fit()*.
- Finalmente, se imprimen las propiedades y se genera la imagen de la red con el comando *graphviz.plot()*.

Una vez obtenida la RB que describe el modelo probabilístico, se pueden realizar inferencias, a través de su estructura dada por el DAG o por el conjunto de sus distribuciones locales.

La Figura 1 muestra la interpretación gráfica de la Red Bayesiana estimada, que es discreta.

En la tabla 3 se muestran las probabilidades condicionales estimadas.

Tabla 3. Probabilidades de los nodos padre

| pH | pH | %MO | %MO | Clase Textual | Clase Textual | Ca | Ca | Mg | Mg |
|-------|-------|-------|-------|---------------|---------------|-------|-------|-------|-------|
| 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| 0.772 | 0.227 | 0.797 | 0.202 | 0.734 | 0.265 | 0.329 | 0.670 | 0.139 | 0.860 |

Debe señalarse que la variable CE, no es una variable relevante y por eso no aparece en la Figura1, pues no presenta ninguna relación con el resto de variables. Puede apreciarse que “%MO” influye en “%N” de manera predecible y, a su vez, todas influyen de manera evidente en el porcentaje de calidad del suelo (PCS).

4 Conclusiones

Se estableció un índice de la calidad del suelo para el cultivo del café de manera porcentual, de modo que fue posible asignar calidades según parámetros estadísticos. Se logró construir una red bayesiana que permitiera el análisis de la estructura causal de las variables, por medio del grafo acíclico dirigido y de las probabilidades condicionales entre los nodos.

El objetivo de describir algunos algoritmos y construir una red bayesiana del problema en cuestión se alcanzó; los modelos probabilísticos resultantes permiten estudiar la problemática y entender la

relación entre las variables consideradas.

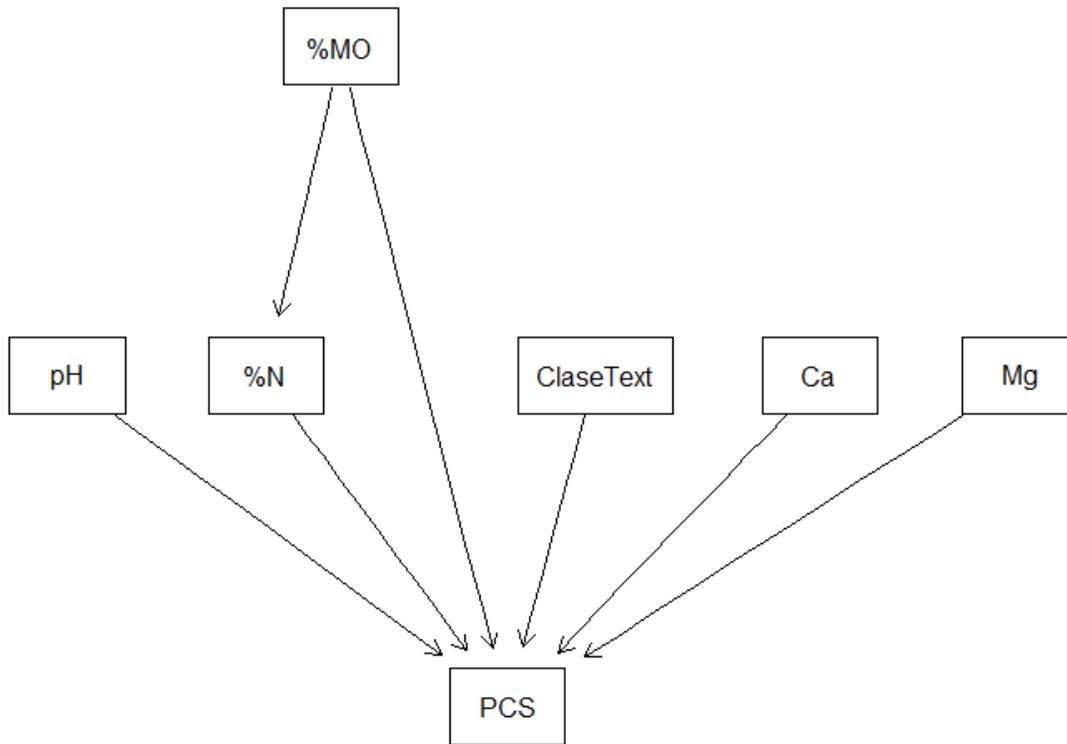


Figura 1. Gráfica de la red bayesiana resultante de la base de datos.

Referencias

[1] Pearl, J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann.

[2] Scutari, M., & Denis, J.-B. (2021). Bayesian Networks: With Examples in R, 2nd ed. Texts in Statistical Science. Chapman and Hall/CRC.

[3] Valera, M. A., Linares, G. Ticante; J.A. (2018). Producción Sustentable de Café y Calidad de Suelo. Ponencia 2ª Jornada de Trabajo e Investigación Científica. Universidad Autónoma Metropolitana. Unidad Xochimilco, México.

[4] Scutari, M. (2010). Learning Bayesian Networks with the bnlearn R Package. *Journal of Statistical Software*, 35(3), 1-22. URL <http://www.jstatsoft.org/v35/i03/>. [Fecha de consulta: 18/05/2021].

[5] Scutari, M., Scutari, M. M., & MMPC, H. P. (2020). Package 'bnlearn'. Bayesian Network Structure Learning, parameter learning and inference. R package.

[6] Scutari, M., Silander, T. & Ness R. (2022). Bayesian Network Structure Learning, Parameter Learning and Inference. bnlearn r package. Version 4.8.1.

[7] Koski, T. & Noble, J.M. (2009) Bayesian networks : an introduction.– Wiley series in probability and statistics.

[8] Lauritzen S.L., Spiegelhalter D.J. (1988). Local computation with probabilities on graphical structures and their application to expert systems (with discussion). *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 50(2), 157–224.

[9] Angulo Montes, Luis Eduardo, (2020), “Trabajo Fin de Máster: Redes Bayesianas en R, análisis de los paquetes de software disponibles”. Universidad Politécnica de Madrid.

[10] Scanagatta, M., Salmerón, A., Stella, P. (2019). A survey on Bayesian network structure learning from data. *Progress in Artificial Intelligence*, 8, 425-489.

Apéndice: Códigos de programación

Primero se instala la paquetería *bnlearn*:

```
install.packages("bnlearn")
```

Se instalan el resto de paqueterías necesarias:

```
install.packages("openxlsx")
```

```
install.packages("BiocManager")
```

```
BiocManager: install ("Rgraphviz")
```

```
install.packages("igraph")
```

#Ahora se cargan las paqueterías:

```
library(bnlearn)
```

```
library(openxlsx)
```

```
library(Rgraphviz )
```

```
library(graph)
```

```
library(igraph)
```

#Se asigna la dirección del Excel con datos

#Se carga dicha base bajo el nombre de "datos1":

```
ruta_archivo <- "C: RutaArchivo/NombreArchivo.xlsx"
```

```
datos1 <-
```

```
read.xlsx (ruta_archivo, sheet = "Hoja1")
```

Optimización del proceso de desarrollo web mediante la generación automatizada de código utilizando inteligencia artificial

Optimizing the web development process through automated code generation using artificial intelligence

Óscar Iván Hernández Gómez¹, Judith Pérez Marcial², María del Carmen Santiago Díaz², Ana Claudia Zenteno Vázquez², María Catalina Rivera Morales³, María Dolores Guevara Espinosa³, Gustavo Trinidad Rubín Linares²

¹Universidad de Guadalajara, Juárez 976, Col. Centro, 44100 Guadalajara, Jalisco.

²Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla/
Avenida San Claudio, Blvd 14 Sur, Cdad. Universitaria, 72592 Puebla, México

³Facultad de Ingeniería Química, Benemérita Universidad Autónoma de Puebla.

oscar.hernandez4685@alumnos.udg.mx , {judith.perez, marycarmen.santiago,
ana.zenteno, maria.riveram, dolores.guevara, gustavo.rubin}@correo.buap.mx

Abstract

This paper analyses the use of artificial intelligence to improve students' emotional and academic wellbeing through web-based tools designed to collect key data. Using technologies such as ChatGPT and DALLE, the design and development of a website focused on student monitoring was optimised. The methodology included the analysis of AI tools such as GPT-4 and GitHub Copilot, the identification of needs, the design of the website structure with ChatGPT, and the implementation of key functions to ensure efficient development. In addition, testing and feedback collection were carried out to identify areas for improvement, while continuous monitoring ensures software quality. The results highlight that artificial intelligence increases efficiency and quality in application development, providing innovative solutions to support emotional and academic wellbeing in the university environment.

Resumen

Este trabajo analiza el uso de la inteligencia artificial para mejorar el bienestar emocional y académico de los estudiantes mediante herramientas web diseñadas para recopilar datos clave. Utilizando tecnologías como ChatGPT y DALLE, se optimizó el diseño y desarrollo de un sitio web enfocado en el monitoreo estudiantil. La metodología incluyó el análisis de herramientas de IA como GPT-4 y GitHub Copilot, la identificación de necesidades, el diseño de la estructura del sitio web con ChatGPT, y la implementación de funciones clave para garantizar un desarrollo eficiente. Además, se realizaron pruebas y recopilación de retroalimentación para identificar áreas de mejora, mientras

que el monitoreo continuo asegura la calidad del software. Los resultados destacan que la inteligencia artificial incrementa la eficiencia y calidad en el desarrollo de aplicaciones, proporcionando soluciones innovadoras para apoyar el bienestar emocional y académico en el entorno universitario.

Keywords and phrases: Inteligencia Artificial, Desarrollo Web, Procesamiento de Lenguaje Natural, Análisis de Datos, Automatización.

1 Introducción

El bienestar emocional y académico de los estudiantes es un área crítica en el contexto educativo moderno, donde la tecnología juega un papel cada vez más central. Con la evolución de las herramientas de inteligencia artificial (IA), como ChatGPT y DALLE, se han abierto nuevas posibilidades para apoyar y mejorar la experiencia educativa de los estudiantes. Estas tecnologías no solo optimizan el desarrollo de aplicaciones web, sino que también proporcionan soluciones innovadoras para la recopilación y análisis de datos sobre el bienestar estudiantil.

La implementación de IA en el ámbito educativo ha demostrado ser eficaz en la personalización de la educación y en el apoyo a estudiantes en situación de vulnerabilidad, especialmente en un entorno post-pandemia donde el bienestar psicológico ha tomado una importancia renovada [1]. Además, el uso de sistemas de soporte educativo basados en IA, como chatbots y asistentes virtuales, ha comenzado a transformar la manera en que los estudiantes interactúan con sus entornos de aprendizaje, mejorando la eficiencia en la enseñanza y la calidad de las interacciones [2], [3].

El presente trabajo explora el potencial de la inteligencia artificial para desarrollar un sitio web destinado a la monitorización del bienestar estudiantil. Mediante la integración de herramientas como ChatGPT y DALLE, se busca no solo mejorar la eficiencia del desarrollo web, sino también crear un sistema capaz de proporcionar apoyo emocional y académico a los estudiantes de manera eficiente.

2 Marco teórico y estado del arte

La literatura sobre la aplicación de la inteligencia artificial en la educación es extensa y abarca diversas áreas que van desde el soporte emocional hasta la personalización del aprendizaje. Investigaciones recientes han demostrado que la IA puede desempeñar un papel crucial en la mejora del bienestar psicológico de los estudiantes al proporcionar sistemas de soporte automatizados que se adaptan a las necesidades individuales [4]. Además, la implementación de herramientas de IA en la evaluación educativa permite no solo una mayor precisión en la medición

del rendimiento académico, sino también la posibilidad de ofrecer retroalimentación personalizada a los estudiantes [5].

Por otro lado, se ha explorado el uso de la inteligencia artificial en el contexto del apoyo a estudiantes en riesgo o en situaciones de vulnerabilidad. Estos sistemas no solo pueden ofrecer asistencia académica, sino también ayudar a los estudiantes a gestionar su bienestar emocional en tiempos de crisis [6]. Además, la introducción de chatbots en la educación ha mostrado ser una herramienta poderosa para mejorar la productividad y la calidad de la enseñanza, especialmente en contextos donde los recursos humanos son limitados [7].

A pesar de los beneficios evidentes, el uso de la IA en la educación también plantea desafíos éticos, particularmente en cuanto a la privacidad de los datos y la equidad en el acceso a estas tecnologías [8]. Sin embargo, con el desarrollo continuo y la implementación responsable de estas herramientas, es posible que la IA se convierta en un componente esencial del ecosistema educativo, proporcionando beneficios tanto en el rendimiento académico como en el bienestar general de los estudiantes [9].

3 Desarrollo del trabajo

Con el objetivo de desarrollar herramientas basadas en inteligencia artificial que ayuden a los estudiantes universitarios a mantener su bienestar emocional y académico. El desarrollo del proyecto se estructuró en varias etapas, utilizando diferentes tecnologías y herramientas de inteligencia artificial, y se organizó por semanas para asegurar un enfoque sistemático y organizado.

3.1 Consideraciones éticas

Propiedad intelectual y plagio: Para afianzar la originalidad del contenido generado utilizando IA, se formularon prompts que evitaran solicitar reproducciones literales de textos existentes o información que pudiera estar protegida por derechos de autor. Al ser un proyecto de desarrollo de software, se optó por el uso de tecnologías y herramientas de código abierto, para asegurar la transparencia y accesibilidad del código.

Privacidad de datos obtenidos: Los datos recogidos a través de las encuestas fueron anónimos y almacenados en una base de datos PostgreSQL.

Originalidad del contenido: Para la generación de imágenes, se garantizó que todo el contenido visual fuera de autoría propia. Se evitó el uso de imágenes protegidas por derechos de autor, asegurando que el material visual utilizado en el proyecto fuera completamente original y generado específicamente para este propósito.

Uso responsable de IA: Las herramientas de IA se utilizaron como apoyo en el desarrollo y no como sustituto de la toma de decisiones. Se mantuvo la supervisión humana en todo momento para asegurar que las recomendaciones y el contenido generado por la IA fueran apropiados y alineados con los objetivos éticos del proyecto.

3.2 Definición de requisitos y planificación

Esta sección describe cómo se establecieron los requisitos del proyecto y cómo se planificó su desarrollo.

Definición de requisitos: La primera etapa del proyecto involucró la identificación y documentación de los requisitos funcionales y no funcionales que guiarían el desarrollo del sistema los cuales se muestran en la Tabla 1. Estos requisitos fueron esenciales para asegurar que la plataforma cumpliera con los objetivos establecidos y que todas las funcionalidades necesarias fueran implementadas.

Tabla1. Prioridad de requisitos del proyecto.

| Descripción | Prioridad |
|--|-----------|
| El sistema debe permitir la recopilación de datos a través de encuestas en línea. | Media |
| El sistema debe almacenar de manera segura los datos recopilados en una base de datos. | Alta |
| El frontend debe ser accesible y fácil de usar para los estudiantes universitarios. | Media |
| El sistema debe integrar herramientas de IA para generar contenido visual y textual. | Alta |
| El backend debe ser capaz de manejar solicitudes de datos y responder a través de una API RESTful. | Alta |
| El sistema debe garantizar la privacidad de los datos personales de los estudiantes. | Alta |

3.3 Planificación del proyecto: La planificación del proyecto se llevó a cabo en varias etapas clave, con el objetivo de estructurar de manera sistemática el desarrollo de la plataforma. Estas etapas fueron definidas por el avance del proyecto a través de hitos importantes.

En la etapa de presentación y encuadre, se definieron los objetivos principales del proyecto, estableciendo el alcance y las expectativas que guiarían su desarrollo. A continuación, se realizó una revisión del marco teórico e identificación de herramientas, donde se investigó el uso de la inteligencia artificial en el desarrollo de software. Durante esta fase, se identificaron diversas herramientas de IA que podrían ser útiles para el proyecto, enfocándose en aquellas que permitieran la generación de código y contenido visual. Tras una evaluación, se procedió a la investigación y selección de herramientas de IA, eligiendo ChatGPT y DALLE como las más adecuadas para el proyecto. ChatGPT fue seleccionado para la generación de texto y código, mientras que DALLE se eligió para la creación de imágenes personalizadas. En la fase de personalización y capacitación de herramientas de desarrollo, se utilizó ChatGPT para generar el

código necesario para el sitio web, y se exploró el funcionamiento de DALLE para crear imágenes alineadas con los objetivos del proyecto. Finalmente, el desarrollo de la aplicación implicó la implementación tanto del frontend como del backend, utilizando tecnologías como HTML, CSS, JavaScript, Python, PostgreSQL y Bootstrap, integrando las herramientas de IA seleccionadas al flujo de trabajo para optimizar el proceso de desarrollo.

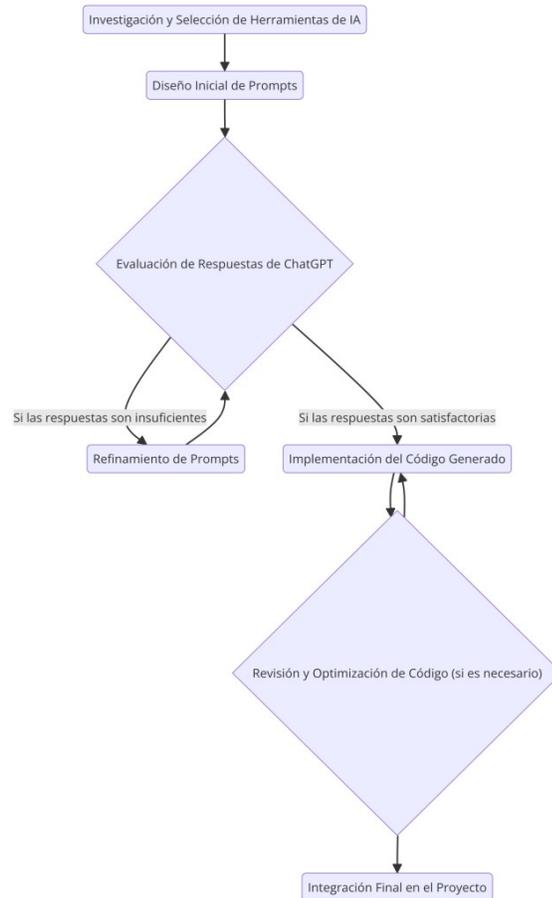


Figura 1. Diagrama de flujo del proceso de Prompt Engineering.

Generación de código y Prompt Engineering: En esta sección, se detalla cómo se formularon y refinaron los prompts para obtener respuestas precisas y útiles de ChatGPT y cómo estos prompts fueron esenciales para la generación de código y el diseño de la plataforma. Investigación y Selección de Herramientas de IA: Se evaluaron varias herramientas de IA para determinar cuáles serían las más adecuadas para el proyecto. ChatGPT fue seleccionada por su capacidad para generar código y proporcionar recomendaciones técnicas, mientras que DALLE fue elegida para la creación de contenido visual. Estas herramientas fueron integradas de manera estratégica para apoyar el desarrollo del frontend, backend, y el diseño de la base de datos. Prompt Engineering: Fue un proceso que involucró la creación y refinamiento de instrucciones específicas para ChatGPT como se detalla en la Figura 1. En la Tabla 2 se explica los prompts clave utilizados durante el desarrollo del proyecto, destacando cómo cada uno de ellos fue diseñado para abordar necesidades

específicas del proceso, desde la planificación y estructura del sitio web hasta la implementación y pruebas de funcionalidades.

Tabla2. Prompts de desarrollo.

| Descripción del Prompt | Propósito |
|---|--|
| ¿Cómo puedo crear un cronograma de desarrollo efectivo para mi proyecto? Dame consejos sobre cómo dividir el trabajo en etapas y asignar plazos realistas. | Obtener orientación sobre cómo estructurar el cronograma de desarrollo del proyecto. |
| ¿Cuáles son los elementos clave que debo incluir en la documentación de mi proyecto? Explícame cómo estructurar la documentación para que sea clara y útil. | Obtener recomendaciones sobre cómo estructurar y organizar la documentación del proyecto, asegurando que sea completa, clara y útil. |
| ¿Cómo puedo crear una encuesta en Google Forms y almacenar las respuestas automáticamente en una hoja de Google Sheets? | Aprender a configurar una encuesta en Google Forms que almacene automáticamente las respuestas en una hoja de Google Sheets, lo que permite una recolección y gestión de datos eficiente. |
| Explícame como instalar PostgreSQL, pgAdmin y posibles librerías de Python para mi proyecto, incluyendo comandos necesarios para cada instalación. | Obtener instrucciones detalladas para la instalación de PostgreSQL, pgAdmin y las librerías necesarias de Python, asegurando que el entorno de desarrollo esté configurado correctamente para el manejo de bases de datos y el desarrollo backend. |
| ¿Cómo puedo diseñar una base de datos en PostgreSQL basada en las preguntas de mi encuesta en Google Forms? Incluye los nombres de las | Obtener una guía para diseñar una base de datos en PostgreSQL que esté alineada con las preguntas de una encuesta en Google Forms. |

4 Resultados

El proyecto desarrolló un sitio web con un diseño responsivo y funcionalidades específicas para apoyar el bienestar estudiantil. El dashboard incluye diez gráficos interactivos, obtenidos desde PostgreSQL, que permiten a los estudiantes analizar sus datos y compararlos con tendencias generales. El sitio web incluye una Página de Inicio, que explica el propósito del proyecto, conecta a los estudiantes con recursos adicionales y permite completar encuestas sobre bienestar emocional y académico. La página de Dashboard visualiza los datos recopilados a través de gráficos interactivos para identificar áreas de mejora, mientras que la Página de Recursos ofrece materiales organizados, como videos, libros y herramientas interactivas. Además, la Integración de Encuestas automatiza la sincronización de datos de Google Forms con PostgreSQL mediante un script en Python, asegurando un almacenamiento eficiente para análisis posterior.

El backend, desarrollado en Flask, interactúa eficientemente con la base de datos utilizando SQLAlchemy. Las herramientas de inteligencia artificial, como ChatGPT y DALLE, optimizaron el diseño, el desarrollo y la personalización del sistema, mejorando la experiencia del usuario y proporcionando ideas clave para los gráficos y recursos educativos. Este enfoque demostró ser viable para proyectos similares que busquen integrar tecnologías avanzadas.

Página de Dashboard En esta página los estudiantes pueden visualizar los datos recopilados a través de gráficos. El dashboard incluye diez gráficos principales que reflejan diferentes aspectos del bienestar estudiantil. Estos gráficos permiten a los estudiantes analizar sus propios datos y compararlos con las tendencias generales, facilitando la identificación de áreas de mejora. Los datos se obtienen directamente de la base de datos PostgreSQL como se muestra en el Figura 3.

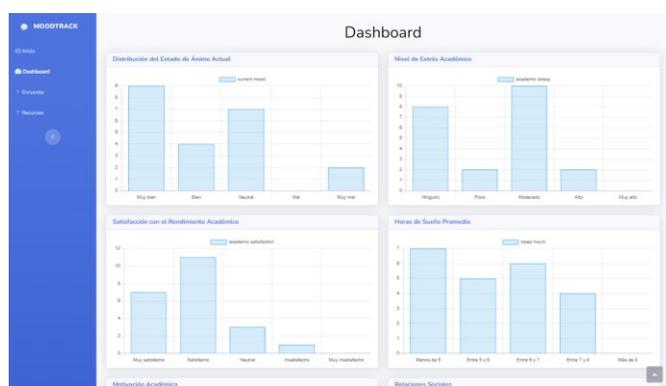


Figura 3. Vista de Página de Dashboard.

El sitio web cuenta con cuatro páginas principales, cada una diseñada para cumplir un propósito específico dentro del proyecto. La Página de Inicio presenta una introducción al propósito y los objetivos del proyecto, ofreciendo a los estudiantes la posibilidad de completar encuestas sobre su bienestar emocional y académico. Las respuestas se almacenan automáticamente en una base de datos PostgreSQL, y la página también conecta a los usuarios con recursos adicionales, como videos educativos, lecturas recomendadas y guías prácticas, que apoyan su bienestar general.

La Página de Recursos proporciona acceso a materiales organizados en tablas temáticas: videos y cursos en línea, lecturas y libros, artículos y guías, y herramientas interactivas. Estos recursos están diseñados para ayudar a los estudiantes a mejorar sus habilidades académicas, manejar el estrés y mantener un equilibrio entre la vida personal y académica.

Para garantizar un diseño intuitivo y responsivo, se utilizaron tecnologías como HTML, CSS y JavaScript, además de Bootstrap para facilitar la navegabilidad. También se implementó DALLE para generar imágenes personalizadas que enriquecen la experiencia visual del usuario.

En el backend, desarrollado en Flask, se utilizó SQLAlchemy como ORM para la integración con PostgreSQL. Esto permite una interacción eficiente entre el frontend y la base de datos mediante endpoints que gestionan operaciones de lectura y escritura.

La integración de encuestas se automatizó mediante un script en Python que conecta Google Forms con PostgreSQL. Los datos recopilados se almacenan inicialmente en Google Sheets, y el script, utilizando las bibliotecas *gsread* y *pandas*, filtra e inserta solo los registros nuevos en la base de datos. La tabla *surveys* almacena estos datos de manera estructurada, asegurando su sincronización automática y disponibilidad para análisis posteriores.

El uso de herramientas de inteligencia artificial como ChatGPT y DALLE optimizó significativamente el desarrollo web. Estas tecnologías no solo mejoraron la eficiencia y calidad del diseño y la implementación, sino que también facilitaron la creación de gráficos, encuestas y recursos educativos personalizados. Los resultados demuestran que la IA es una solución viable para proyectos similares, permitiendo una integración más eficiente de tecnologías avanzadas y proporcionando funcionalidades innovadoras para mejorar la experiencia del usuario. Los resultados del proyecto mostraron que la utilización de herramientas de inteligencia artificial como ChatGPT y DALLE facilitan el desarrollo web. El uso de estas tecnologías no solo optimiza el tiempo de desarrollo, sino que también mejora la calidad y la personalización de las funcionalidades creadas, como la visualización de gráficos, creación de encuestas o la personalización de recursos educativos.

Se observó que las herramientas de IA permitieron una integración más eficiente entre las tecnologías necesarias para cumplir con las funcionalidades del sistema, facilitando la recopilación de información y proporcionando ideas clave para la creación de gráficos y recursos. Estos resultados sugieren que el uso de IA puede ser un recurso viable para proyectos similares que buscan desarrollar aplicaciones de software.

5 Conclusiones y trabajo a futuro

La inteligencia artificial demuestra ser una herramienta poderosa para mejorar el bienestar emocional y académico de los estudiantes, al facilitar el desarrollo de aplicaciones web que recopilan y analizan datos clave. El uso de tecnologías como ChatGPT y DALLE permitió optimizar el diseño y la implementación de un sitio web enfocado en el monitoreo estudiantil, destacando su potencial para incrementar la eficiencia y la calidad del desarrollo de software. Además, la integración de metodologías basadas en el análisis, diseño, desarrollo, pruebas y mantenimiento continuo asegura la adaptabilidad y sostenibilidad del sistema. Este enfoque no solo contribuye al desarrollo tecnológico en el ámbito educativo, sino que también proporciona soluciones innovadoras para el monitoreo y apoyo del bienestar estudiantil, ofreciendo un modelo escalable y replicable que puede aplicarse en diversas instituciones educativas para atender las necesidades emocionales y académicas de su comunidad.

Se recomienda que futuras investigaciones exploren la expansión de estas tecnologías a otras áreas educativas, así como la evaluación continua de su impacto ético y técnico. Además, se refuerza la necesidad de integrar consideraciones éticas en cada etapa del desarrollo, garantizando la originalidad y seguridad de los contenidos generados.

Referencias

- [1] T. Van, M. A. M. Daril, M. Ali, and M. S. Korejo, "Enhancing Psychological Well-being in Higher Education Post-Covid-19 Pandemic. The Role of AI-Based Support Systems—Bibliometric Reviews," *Int. J. Online Biomed. Eng. IJOE*, vol. 20, no. 06, Art. no. 06, Apr. 2024, doi: 10.3991/ijoe.v20i06.48001.
- [2] V. J. Owan, K. B. Abang, D. O. Idika, E. O. Etta, and B. A. Bassey, "Exploring the potential of artificial intelligence tools in educational measurement and assessment," *Eurasia J. Math. Sci. Technol. Educ.*, vol. 19, no. 8, p. em2307, Aug. 2023, doi: 10.29333/ejmste/13428.
- [3] P. Prinsloo, M. Khalil, and S. Slade, "Vulnerable student digital wellbeing in AI-powered educational decision support systems (AI-EDSS) in higher education," *Br. J. Educ. Technol.*, vol. n/a, no. n/a, doi: 10.1111/bjet.13508.
- [4] "Artificial Intelligence and the Student Experience: An Institutional Perspective," *The International Academic Forum (IAFOR)*. Accessed: Aug. 12, 2024. [Online]. Available: <https://iafor.org/journal/iaforjournal-of-education/volume-6-issue-3/article-4/>
- [5] S. P. T. Utami, Andayani, R. Winarni, and Sumarwati, "Utilization of artificial intelligence technology in an academic writing class: How do Indonesian students perceive?," *Contemp. Educ. Technol.*, vol. 15, no. 4, p. ep450, Oct. 2023, doi: 10.30935/cedtech/13419.
- [6] N. Sandu and E. Gide, "Adoption of AI-Chatbots to Enhance Student Learning Experience in Higher Education in India," in *2019 18th International Conference on Information Technology Based Higher Education and Training (ITHET)*, Sep. 2019, pp. 1–5. doi: 10.1109/ITHET46829.2019.8937382.
- [7] B.-I. Ouatu and D. Gifu, "Chatbot, the Future of Learning?," in *Ludic, Co-design and Tools Supporting Smart Learning Ecosystems and Smart Education*, Ó. Mealha, M. Rehm, and T. Rebedea, Eds., Singapore: Springer, 2021, pp. 263–268. doi: 10.1007/978-981-157383-5_23.
- [8] S. Devi, A. S. Boruah, S. Nirban, D. Nimavat, and K. K. Bajaj, "Ethical Considerations in Using Artificial Intelligence to Improve Teaching and Learning," *Tuijin Jishu Journal Propuls. Technol.*, vol. 44, no. 4, Art. no. 4, Oct. 2023, doi: 10.52783/tjjpt.v44.i4.966.
- [9] F. Ouyang, M. Wu, L. Zheng, L. Zhang, and P. Jiao, "Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course," *Int. J. Educ. Technol. High. Educ.*, vol. 20, no. 1, p. 4, Jan. 2023, doi: 10.1186/s41239-022-00372-4.

Simulación mediante Python y ROS de los algoritmos de Dijkstra y A* para el movimiento de un robot móvil autónomo

Simulation using Python and ROS of the Dijkstra and A* algorithms for the movement of an autonomous mobile robot

Ricardo Álvarez González¹, Ivonne Pérez Flores², Dayana Pérez Flores², Alba Maribel Sánchez Gálvez², Brian Manuel González Contreras³

¹Facultad de Ciencias de la Electrónica, Benemérita Universidad Autónoma de Puebla, Av. San Claudio y 18 sur Col. San Manuel. C.P. 72570. Puebla, Puebla, México.

²Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla, Av. San Claudio y 14 sur Col. San Manuel. C.P. 72570. Puebla, Puebla, México.

³Facultad de Ciencias Básicas de Ingeniería y Tecnología, Universidad Autónoma de Tlaxcala, Carretera Apizaquito S/N, San Luis Apizaquito, C.P. 90401, Apizaco, Tlaxcala, México.

{ricardo.alvarez, alba.sanchez}@correo.buap.mx, {Ivonne.perezf, dayana.perezf}@alumno.buap.mx, brianmanuel.gonzalez@uatx.mx

Abstract

Autonomous terrestrial mobile robots (RMA) have transcendental importance and relevance today, they are immersed in different industrial and domestic tasks. Trajectory planning constitutes one of the main tasks of an RMA. The results obtained in RMA trajectory planning can also be applied to autonomous vehicles. Among the basic algorithms for trajectory planning are Dijkstra's and A*. This work presents the simulation of these algorithms in a discrete world consisting of a 10 x10 grid for a differential robot, based on classes developed in Python and the Rviz and Gazebo simulators in the ROS environment, to appreciate their main characteristics and compare their performance.

Resumen

Los robots móviles autónomos terrestres (RMA) tienen trascendental importancia y relevancia en la actualidad, se encuentran inmersos en diferentes tareas tanto industriales como domésticas. La planificación de trayectorias constituye una de las tareas principales de un RMA. Los resultados obtenidos en la planificación de trayectorias de los RMA se pueden aplicar también para los vehículos autónomos. Dentro de los algoritmos básicos para la planificación de trayectorias se encuentran el de Dijkstra y el A*. En este trabajo se presenta la simulación de estos algoritmos en un mundo discreto constituido por una cuadrícula de 10 x10 para un robot diferencial, a partir de

clases desarrolladas en Python y los simuladores Rviz y Gazebo en el entorno ROS, con la finalidad de apreciar sus características principales y comparar su desempeño.

Keywords and phrases: Autonomous Mobile Robot, Dijkstra's Algorithm, A*, ROS.

1 Introduction

One of the main problems of humanity throughout time has been obtaining minimum routes. Regardless of the purpose of these routes, various problems have been born with them (without necessarily having scientific purposes), where the need for efficiency is evident. Along with the arrival of these problems, solutions to them have emerged, laying down the bases of this type of algorithms, which, and defining their behavior in general aspects, seek to obtain the shortest or least expensive route from one point to another within a defined graph or world. There are a wide range of applications that this type of algorithms has, being present in areas such as video games, autonomous vehicles, communications networks, robots, among others [1]. Throughout this project, we will analyze and, subsequently, visualize the execution and decision-making process of two of the most used minimum path algorithms in Artificial Intelligence: Dijkstra and A* (also known as A star), implemented in a mobile robot within a virtualized world through a simulator, to understand its application within the field of robotics.

2 Theoretical framework and state of the art

The simulation of the operation and implementation of autonomous mobile robots (AMR) is of particular interest due to its multiple applications in daily life, as we can see in [2] where they present the simulation in Matlab and experimental verification in ROS of an AMR using a hybrid algorithm based on the A* and artificial potential field algorithms for the AMR navigation task avoiding obstacles.

In [3] they are based on the Dijkstra and A* algorithms to carry out the trajectory planning of an RMA, where they concentrate on making improvements to the heuristic function from the parent node, searching for the optimal weight to optimize the function of the algorithm A*, which they simulate in Matlab to validate its operation.

An approach called modified temporal difference Learning is proposed in [4] for path planning and obstacle avoidance for static obstacles. The algorithm is developed in MATLAB software and path planning is implemented in a 4x4 grid environment. They created a GUI for the same, which accesses user inputs such as number, positions and type of obstacle. The developed algorithm is compared with the Dijkstra algorithm for conventional route planning in the same environment.

In [5] they work with the A star algorithm, changing the dynamic weights of the heuristic function, considering sixteen search directions, to improve some aspects presented in the traditional A star algorithm, they also carry out a method of dynamic adjustment of the weighting of the coefficients of the estimated cost $h(n)$ in the A star algorithm, for the search for redundant nodes. They perform simulation experiments in Matlab under a given map. They also make comparisons with Dijkstra's algorithm.

In [6] they present an improved version of the A star algorithm, to solve problems inherent in long paths and duplicate the waypoints of the traditional methods of the original A star algorithm, extending the search neighborhood from 4 to 16, in addition to eliminating redundant nodes by improving the evaluation function to achieve a smoothed path. They verify their proposal through simulations in Matlab, where they compare it with Dijkstra's algorithm.

As can be seen in these previous paragraphs, in the field of AMR the algorithms of Dijkstra and A* are of very high importance and transcendence, so in this work the simulation of the operation of each of them in a virtual world, built for this purpose, is carried out.

2.1 Dijkstra's algorithm

The Dijkstra algorithm was proposed by the scientist Edsger Dijkstra in 1956, it is one of the most important algorithms in the computational area. Its behavior is based on obtaining the shortest path between one of the nodes of the graph to the source node, considering the cost of movement between the nodes. To carry out its execution, certain initial guidelines must be established, so an initial node belonging to the graph is established, in addition, for each node it must be known if the total distance from the root to the node is known and know its parent node, both the parent node and the distance, can change according to the algorithm [1].

2.2 A* algorithm

It is an algorithm widely used in video games and has been used in many of them for the artificial intelligence of the characters. This algorithm descends from Dijkstra's, and it is an efficient algorithm and guarantees us the shortest or least expensive route between two nodes, if the route exists. If costs are used in the movement, the route, instead of being the shortest in distance, will be the most economical in cost. You can also create a graph in a traditional way on a navigation map and use the algorithm with existing graphs.

To understand the conceptualization of the algorithm we must establish that it will work within a grid of movement (coordinate plane), also called the world. The focus of this algorithm is mainly on the calculation of cost evaluation, and the way to analyze nodes within the world. To decide on the shortest path, a sum is made between the distance between the node and the origin $g(n)$, and the distance between the node and the destination $h(n)$, thus obtaining the equation 1:

$$f(n) = g(n) + h(n) \tag{1}$$

Where, $f(n)$, in terms of the algorithm, is the total cost of each node to be considered (for the implementation of the algorithm it is called a score); $g(n)$ is the distance and $h(n)$ will be the heuristic, all of them through n [7].

For the calculation of the cost, both in the distance and in the heuristics, the Manhattan distance has been considered, which measures the distance of movement on the horizontal axis and on the vertical axis, without considering the diagonal movement, making an analogy to the movement of a taxi in Manhattan [1].

For this evaluation we consider equation 2:

$$D_{Manhattan} = |x_1 - x_2| + |y_1 - y_2| \quad (2)$$

Where the coordinates (x_1, y_1) belong to the cell in which the robot is positioned, while the coordinates (x_2, y_2) belong to the reference cell of origin or destination. Thus, $g(n) = D_{Manhattan}$ and $h(n) = D_{Manhattan}$, however, do not represent equal values given their reference points.

Other concepts that are of utmost importance to consider are the use of two lists, the open list, which will store the files that we have to review throughout the execution of the algorithm, and the closed list, which will store the cells that have already been visited [1].

Now we will consider the direction, this is found during the evaluation of each cell, which in terms of the program will be represented with pointers, where each cell will contain a pointer to another cell, which will be called the parent cell (cell that precedes the current cell) and the child cell (current cell).

For a correct implementation, we consider the following:

- The open list, initially, will only store the cell where the robot is located.
- During the execution of the algorithm, the cells will move from the open list to the closed list.
- The closed list will initially be empty.
- The algorithm will continue to run if the open list contains cells, and the evaluated cell is different from the target cell.

Finally, let's consider that the execution of the algorithm will be visualized graphically by the ROS software, represented in Gazebo, which will be shown in the next section.

3 AMR motion planning and algorithm simulation

3.1 Differential traction land mobile robot

We will work with a differential robot, whose model in Solid Works is shown in Figure 1. This model was taken to the ROS (robot operating system) simulation environment to manipulate actions such as the robot's autonomous motion programming based on Dijkstra's algorithm.

3.2 Robot movement planning, using Dijkstra's algorithm

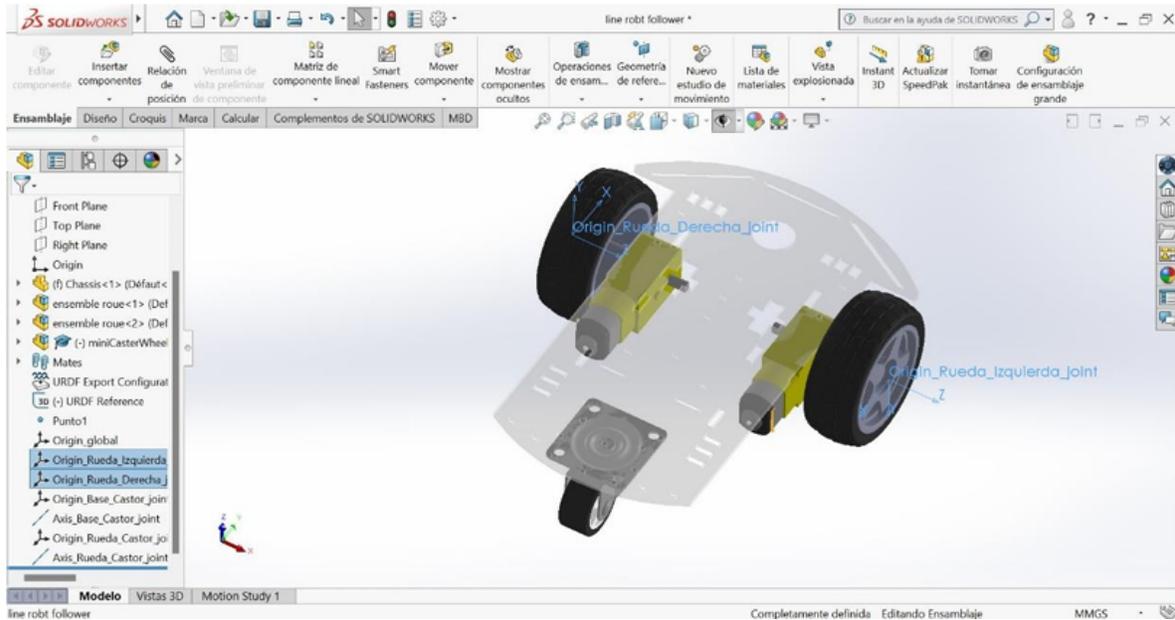


Figure 1. Differential robot model in solid works

A discrete world was generated made up of a 10x10 grid, in which points were marked that the robot can reach, from this world a graph was generated, in which the weights have been marked with the Euclidean distance from one point to another.

First, the angle of inclination was planned, for which the slope that exists between two points of the graph must be considered, which have coordinates on the x-axis and on the y-axis. However, to know the angle we must consider that both the slope (m) and the tangent ($\tan(\theta)$) represent the same rate of change between the x and y coordinates. For this reason, the angle of inclination is obtained from equation 3 :

$$\theta = \tan^{-1}\left(\frac{\Delta y}{\Delta x}\right) \quad (3)$$

This planning ensures that the robot moves effectively and accurately between the nodes of the graph, adjusting its orientation according to the direction of the next movement. Now, another important aspect required in the simulation is the time it will take to reach the next node, which is obtained by the Euclidean distance mentioned above, which will be calculated from point to point of each node, using the equation 4:

$$d_{euclidean} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (4)$$

And a velocity of 0.4 m/s is defined; therefore, the time is calculated by the equation 5:

$$t_{\text{node-next node}} = d_{euclidean} \times 0.4 \text{ m/s} \quad (5)$$

For the implementation of these functions and algorithms, a set of classes were developed in Python. The Dijkstra.py file contains the implementations of Dijkstra's algorithm to find the shortest paths between nodes in our graph. On the other hand, in Movimiento.py are the configurations and functions related to the robot's movement, including the calculation of angles and travel times.

3.3 Simulation of AMR motion using Dijkstra's algorithm.

To perform the simulation, an example of Dijkstra was made, in which the shortest path tree (STP) shown in figure 2 is generated, where node B is the parent node.

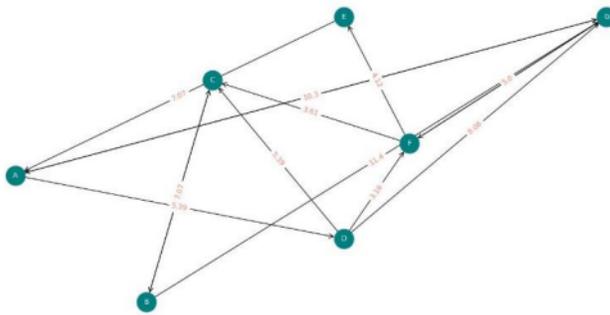


Figure 2. Proposed node

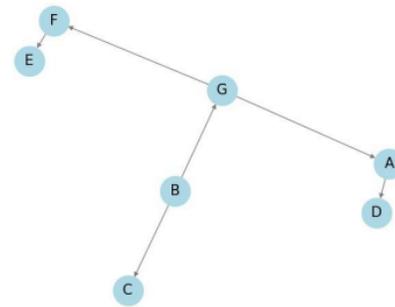


Figure 3. STP simulation of the proposed node

An environment was then created in which the robot will interact. This world contains the points that the robot can reach, which are represented by transparent cubes of different colors, in addition, this world has a limitation of measurements of the grid used for its elaboration.

To exemplify the journey, the example of the journey from node B to node C was made, where this movement can be observed until reaching the next point, considering that this is a direct route according to the SPT.

3.4 Planning and simulation of the movement of the AMR using the A* algorithm.

The A* algorithm was used for the shortest path and not for the least expensive, implemented in a mobile robot within a virtualized world by means of the Gazebo simulator. Using the same 10×10 grid world of the previous algorithm, where obstacles (marked by the darkened grids), free moving cells (empty cells), the robot's starting position (circle in gray cell) and the destination (gray diamond) are visualized, as shown in Figure 5.

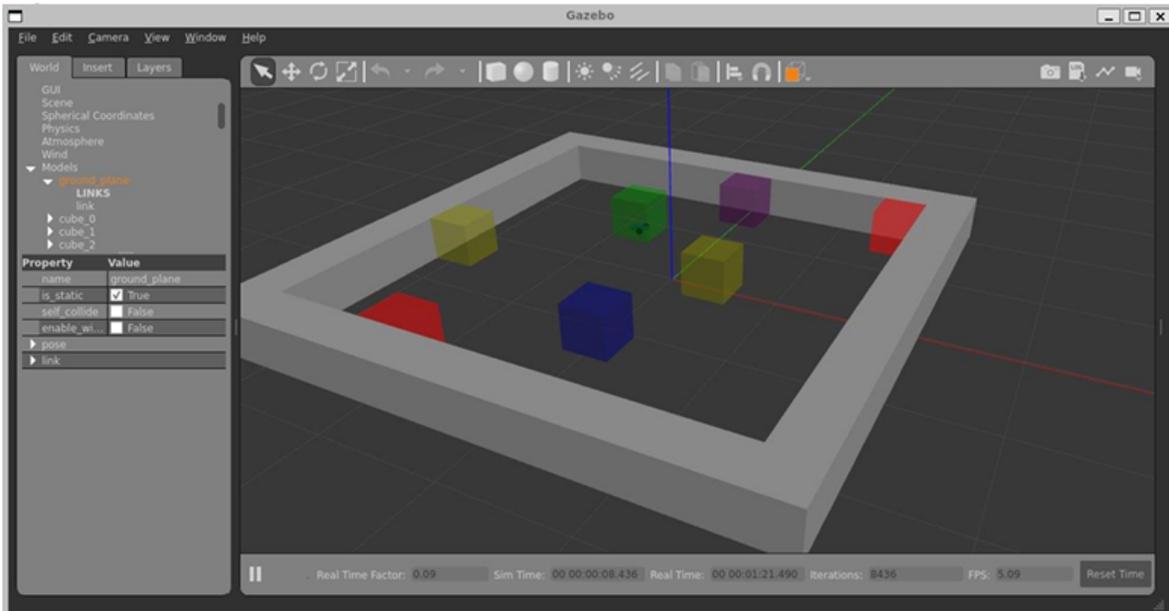


Figure 4. Robot arrival at node C on direct path

The process consists of evaluating the adjacent cells of the world, so we will consider the cell where the robot is located and select the one around it, if it is not an obstacle, if so, it will not be considered. In this way, we will analyze the north, northeast, east, southeast, south, southwest, west and northwest cells, if they exist, and we will include them in the open list.

Now, for each cell we will calculate both its distance and its heuristics, as an example to understand the calculation, the northeast cell is considered. As can be seen in figure 5, that the source cell has the coordinates (6, 5), the destination cell has the coordinates (2, 5), and the northwest cell is located at the coordinates (5,4). Now both values can be calculated, remembering that for the calculation of the distance $g()$ we will use the source cell and for the calculation of the $h()$ (heuristic) we will use the destination cell.

Therefore, we will be able to observe that the distance is calculated by equation 6 and the Heuristic by the equation 7, as explained before:

$$\text{Distance: } g = |5 - 6| + |4 - 5| = 2 \quad (6)$$

$$\text{Heuristic: } h = |5 - 2| + |4 - 5| = 4 \quad (7)$$

Once we know these values, we can calculate the total score $f()$ of the northwest cell, as it shown in the expression 8.

$$\text{Total score: } f = 2 + 4 = 6 \quad (8)$$

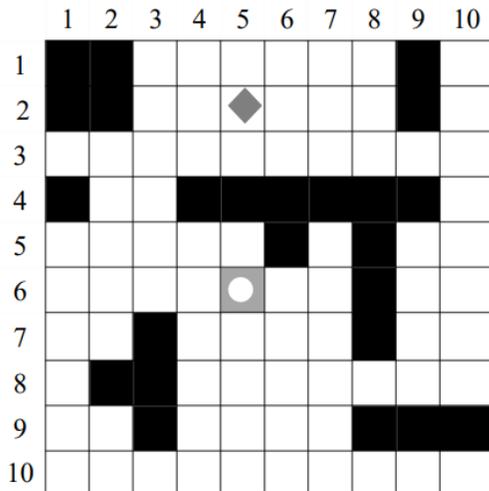


Figure 5. 10 x10 world and cell rendering

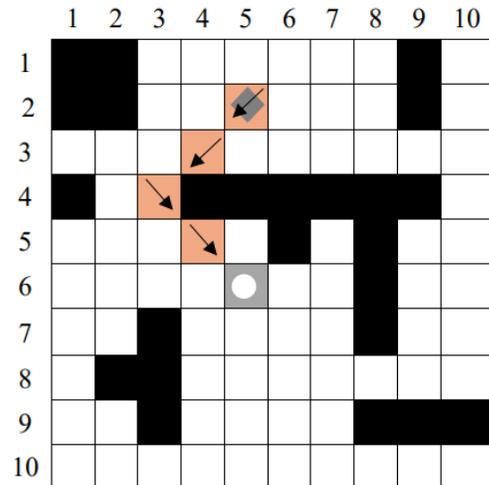


Figure 6. Route with pointers to parent cells.

In this way the scores will be calculated for each of the neighboring cells. For this example, the north cell will have a score of 4, northeast of 6, east of 6, southeast of 8, south of 6, southwest of 8 and west of 6. Remember that all these cells were previously stored in the open list. Now the cell with the lowest score from this list is selected, in this case it is the north cell, and it will be eliminated from the open list and included in the closed list, since it cannot be in both.

To carry out the next iteration, we must consider the cell that we have selected and moved to the closed list, later we will continue the process. Since in the example the neighboring cells have already been considered and the rest are obstacles, we will only search for the cell with the lowest score of the cells that are in the open list, this being the northeast cell.

The same process continues by selecting the cells with the lowest score. The cells are stored in the list counterclockwise, so, in the case of finding two cells with the same score, the one that enters the open list first will simply be selected. Once the destination cell has been reached, there are many cells in the closed list, however, to know the final route we must remember the address of each cell, each of which will contain a parent cell, which will be you will find pointing to it. When we reach the destination cell it will point to its parent cell, which will also point to its parent cell, and so on until we reach the origin cell, we will follow these pointers to find the route. To know the route from origin to destination, we will simply reverse the list, as can be seen in figure 6.

The simulation of this algorithm was also carried out in the ROS environment, using the Rviz and Gazebo tools that are graphic viewers of the simulated robot model, with this we created the world that we see in Figure 7, where the gray cubes show the delimitation. and obstacles, while the transparent yellow cube shows the destination. The origin cell is not shown since the robot (circled in red) is generated in this cell, defining it in the code.

To understand the movement of the robot, each cube is considered to measure 0.4 m and a speed of 0.4 m/s is considered. Now, let's consider that the $D_{diagonal} = 0.5657\text{ m}$. You must know the time it takes to travel the diagonal distance, in addition to the time it takes to turn in the directions north, south, east, west, northeast, northwest, southeast and southwest, taking into account that these have angles of 45° , 90° , 135° y 180° . In this way we know that $t_{diagonal} = 1.41425\text{ s}$, $t_{45^\circ} = 0.5236\text{ s}$, $t_{90^\circ} = 1.0472\text{ s}$, $t_{135^\circ} = 1.5708\text{ s}$ and $t_{180^\circ} = 2.0942\text{ s}$.

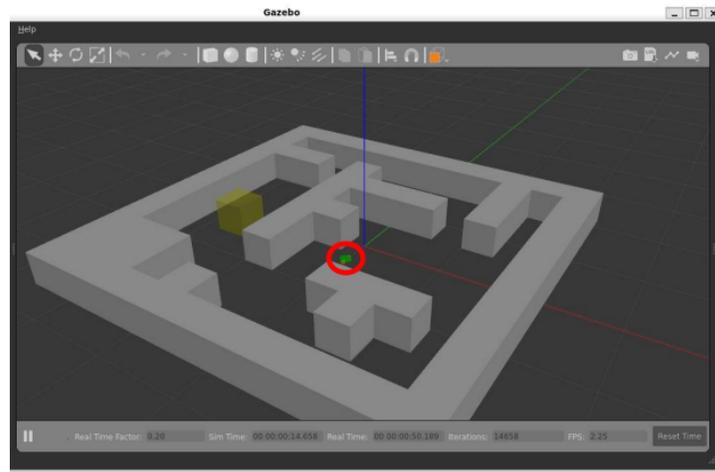


Figure 7. World created with the robot generated in the initial position.

Once the required guidelines for the movement have been defined, the algorithm already made and implemented in Python language (AEstrella.py) is used, and the code that defines the movement of the robot with ROS libraries in the file (Movimiento.py), the which will serve as the publishing node for the movement of the robot. Figure 8 shows the transient movement of the robot's journey through the created world and figure 9 shows its arrival at the destination point.

4 Conclusions and future work

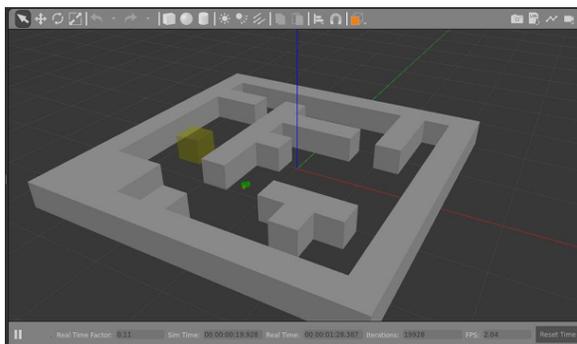


Figure 8. Robot movement and arrival at its end point.

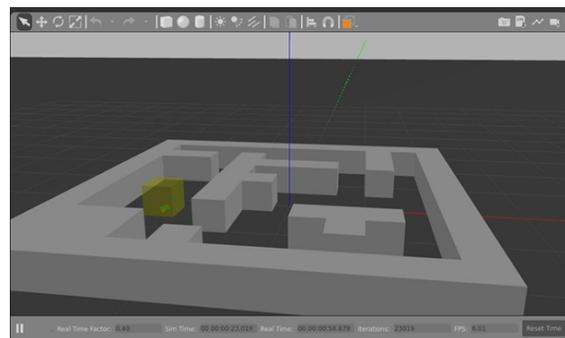


Figure 9. Arrival of the RMA at its end point.

In this study, the implementation and simulation of the Dijkstra and A* algorithms for the autonomous navigation of a differentially driven land mobile robot have been explored. Highlighted the importance of Dijkstra's in the planning of efficient routes, using graphs and the theory of shortest path trees (SPT) and the efficiency and adaptability of the A* algorithm for the efficiency of optimal routes for robots in virtual environments.

Through simulation in ROS it was possible to observe in real time the operation of the algorithms, their decision-making process, guaranteeing autonomy of movement in the mobile robot, through the selection of the shortest path.

Furthermore, through simulation, it has been proven that the A* algorithm can adapt to different configurations of obstacles and destinations, maintaining its ability to find the optimal route. Using tools like Rviz and Gazebo in ROS has been crucial to visualizing and validating the robot's behavior, allowing for a deeper understanding of the algorithm in a controlled environment.

As future work, the experimental implementation of the algorithms will be carried out in a differential robot like the one shown in Figure 1, adding a microcontroller and the other necessary hardware elements.

References

- [1] USERS. Inteligencia Artificial Sistemas Inteligentes con C#.
- [2] J. Yao, "Path Planning Algorithm or Indoor Mobile Robot Based on ROS System", IEEE International Conference on Image Processing and Computer Applications, 2023.
- [3] M. Lin., K. Yuan, C. Shi, Y. Wang" Path Planning of Mobile Robot Based on Improved A* Algorithm ", IEEE Xplore, 2017.
- [4] D. S. Nair, Supriya P, "Comparison of Temporal Difference Learning Algorithm and Dijkstra's Algorithm for Robotic Path Planning", Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS) IEEE Xplore, 2018.
- [5] W. Bao, J. Li, Z. Pan, R. Yu, "Improved A-star Algorithm for Mobile Robot Path Planning Based on Sixteen-direction Search", China Automation Congress (CAC), IEEE 2022.
- [6] Y. Li, J. Yang, "Path planning and obstacle avoidance technology for mobile robots using improved A-star algorithm", 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT), IEEE 2024.
- [7] RUSSELL, Stuart y NORVIG, Peter. Inteligencia Artificial un enfoque moderno. Segunda Edición. Pearson Prentice Hall. Madrid. (2004)

Desarrollo de una aplicación para la detección y conteo de monedas en video en tiempo real utilizando OpenCV y PyQt6

Development of an application for real-time coin detection and counting in video using OpenCV and PyQt6

Brayan Olivares Rodríguez, Adrián Alejandro Ruiz Márquez, Ana Cecilia Rodríguez Chávez,
Javier Martín Palmero Torres, Marco Aurelio Nuño Maganda

Ingeniería en Tecnologías de la Información
Universidad Politécnica de Victoria. Ciudad Victoria, Tamaulipas, México

2130211@upv.edu.mx.

Abstract

The article describes the development of an application for real-time coin detection and analysis using OpenCV and PyQt6. The application allows users to visualize coin detection through a webcam, classifying each coin by its type and calculating the total value detected. The system employs advanced image processing techniques, such as the Hough Transform for circle detection, and counting algorithms to classify coins. The graphical interface, developed with PyQt6, provides a clear and user-friendly visualization. Test results indicate a high level of accuracy and efficiency, highlighting its applicability in commercial and educational settings.

Resumen

Este artículo describe el desarrollo de una aplicación para la detección y análisis de monedas en tiempo real utilizando OpenCV y PyQt6. La aplicación permite a los usuarios visualizar la detección de monedas a través de una cámara web, clasificando cada moneda según su tipo y calculando el total de dinero detectado. Este sistema utiliza técnicas avanzadas de procesamiento de imágenes como la Transformada de Hough para la detección de círculos, y algoritmos de conteo para clasificar las monedas. La interfaz gráfica, desarrollada con PyQt6, proporciona una visualización clara y fácil de usar. Los resultados de las pruebas indican un alto nivel de precisión y eficiencia, destacando su aplicabilidad en entornos comerciales y educativos.

Keywords and phrases: Aplicación, Cámara Web, Visión Computacional, Lenguaje de Programación.

1 Introducción

La detección y clasificación automatizada de objetos en imágenes y videos es fundamental en campos como la visión por computadora, la robótica y las finanzas. En particular, la detección de monedas tiene aplicaciones en el procesamiento automatizado de transacciones y el control de efectivo en sistemas de autoservicio, donde la precisión y rapidez son esenciales.

Este proyecto desarrolla un sistema que emplea OpenCV, una potente biblioteca para el procesamiento de imágenes y video, junto con PyQt6, un robusto framework para crear interfaces gráficas en Python. El objetivo principal es crear una aplicación capaz de capturar video en tiempo real desde una cámara web, detectar círculos que representan monedas y mostrar esta información de manera dinámica en la interfaz gráfica.

La aplicación no solo identifica y cuenta las monedas, sino que también las clasifica según su denominación, proporcionando una valoración total automática. Se busca ofrecer una solución de bajo costo y alta eficiencia, aplicable en entornos comerciales, educativos y de investigación.

Además de presentar la arquitectura del sistema, este artículo detalla las técnicas de procesamiento de imágenes utilizadas, como la Transformada de Hough para la detección de círculos y algoritmos de filtrado. También se incluyen los resultados obtenidos y comparaciones de rendimiento con soluciones existentes en el mercado, como CountThings [1].

2 Marco teórico y estado del arte

2.1 OpenCV y detección de objetos

OpenCV (Open Source Computer Vision Library) es una biblioteca de código abierto para la visión por computadora que proporciona una amplia gama de herramientas para el procesamiento de imágenes y video. Entre sus funcionalidades más relevantes se encuentran los algoritmos de detección de objetos, como la Transformada de Hough, que se utiliza para identificar formas geométricas específicas en imágenes, como líneas y círculos. La eficiencia de OpenCV lo convierte en una elección popular en aplicaciones de tiempo real, donde la velocidad y la precisión son cruciales para el éxito del sistema [2].

2.2 PyQt6 para interfaces gráficas

PyQt6 es un conjunto de herramientas que permite la creación de interfaces gráficas de usuario (GUI) en Python. Su integración con bibliotecas de visión por computadora como OpenCV es fundamental para desarrollar aplicaciones que requieren tanto procesamiento de imágenes como visualización en tiempo real. PyQt6 ofrece widgets altamente personalizables, lo que facilita el diseño de interfaces intuitivas y funcionales para los usuarios finales [3][4].

2.3 Aplicaciones de visión por computadora

Las técnicas de visión por computadora tienen aplicaciones en numerosos campos, que incluyen seguridad, automoción, educación e investigación científica. Estas técnicas permiten la detección automática y el análisis de objetos en imágenes, lo que mejora la eficiencia y precisión en tareas

repetitivas. En particular, la detección de objetos es un área activa de investigación, con numerosos estudios que exploran nuevas metodologías y tecnologías emergentes [5].

2.4 Estado del arte

El proyecto presentado en este estudio se centra en la detección y conteo automático de monedas mediante el uso de OpenCV y PyQt6. A diferencia de CountThings, que únicamente cuenta el número de monedas, nuestra aplicación no solo realiza el conteo, sino que también clasifica cada moneda por su denominación y calcula el valor total detectado, lo que la hace más útil en entornos comerciales y educativos. Además, la interfaz de usuario se ha diseñado para ser intuitiva, lo que permite una experiencia de usuario más fluida. Las pruebas de rendimiento realizadas indican que, aunque ambas aplicaciones tienen precisión comparable, nuestra solución se destaca en la simplicidad y eficiencia de la detección y conteo automático de monedas [6][7].

Se realizaron pruebas de rendimiento para comparar nuestra aplicación con CountThings, encontrando que ambas tienen precisión comparable, pero nuestra aplicación se destaca por la detección y conteo automático de monedas con una interfaz más simple [8].

3 Metodología

La metodología utilizada para el desarrollo de la aplicación se divide en varias etapas clave, que garantizan un enfoque estructurado y sistemático.

3.1 Definición de requisitos

Se llevó a cabo una revisión exhaustiva de la literatura existente sobre aplicaciones de detección de monedas para definir los requisitos funcionales y no funcionales del sistema. Esto incluyó la identificación de características deseadas, como la capacidad de detección en tiempo real, la precisión en la clasificación y la facilidad de uso, asegurando que el producto final cumpliera con las expectativas del usuario.

3.2 Diseño del sistema

En esta etapa, se elaboró un diseño preliminar de la aplicación que abarca la arquitectura del software y la interacción entre sus componentes. Se definieron los módulos principales: la interfaz de usuario (GUI), el módulo de procesamiento de imágenes, el módulo de detección y conteo, y el módulo de visualización. Se optó por utilizar OpenCV para el procesamiento de imágenes debido a su robustez y PyQt6 para la creación de una interfaz gráfica que facilitara la interacción del usuario.

3.3 Desarrollo e implementación

La implementación se realizó en fases, comenzando con el desarrollo de la interfaz de usuario, seguida del módulo de procesamiento de imágenes, y finalizando con el módulo de detección y conteo. Se llevaron a cabo pruebas unitarias en cada fase para garantizar que cada componente funcionara correctamente. Además, se documentaron los algoritmos utilizados, tales como la

Transformada de Hough y la detección de bordes mediante el algoritmo de Canny, asegurando que el código fuera accesible y entendible para futuras mejoras.

3.4 Pruebas y validación

Una vez completado el desarrollo de la aplicación, se realizaron pruebas exhaustivas para evaluar su rendimiento bajo diversas condiciones. Esto incluyó pruebas de precisión y velocidad, en las que se compararon los resultados obtenidos con los de aplicaciones existentes como CountThings. Las pruebas fueron documentadas de manera sistemática, y se realizaron ajustes en la aplicación basados en los resultados para optimizar su funcionamiento.

3.5 Evaluación de resultados

Los resultados se analizaron en términos de precisión en la detección de monedas, velocidad de procesamiento y experiencia del usuario. Se llevaron a cabo encuestas y entrevistas con usuarios potenciales para obtener retroalimentación sobre la usabilidad de la interfaz gráfica y la efectividad del sistema. Este enfoque no solo permite validar la aplicación, sino que también sienta las bases para futuras mejoras y adaptaciones en contextos diferentes.

3.6 Detalles de algoritmos utilizados

En el desarrollo de la aplicación para la detección y conteo de monedas, se emplearon dos algoritmos clave: la Transformada de Hough y el algoritmo de Canny. Estos algoritmos son fundamentales para la identificación y localización de las monedas en las imágenes capturadas por la cámara.

3.6.1 Transformada de Hough

La Transformada de Hough es una técnica utilizada para la detección de formas geométricas en imágenes, siendo especialmente eficaz para la identificación de círculos y líneas. En este proyecto, se aplicó la Transformada de Hough para detectar los círculos que corresponden a las monedas. El algoritmo transforma el espacio de la imagen en un espacio de parámetros, donde cada punto en el espacio de parámetros representa una posible forma.

En el contexto de la detección de círculos, se utiliza la variante conocida como HoughCircles, que permite encontrar círculos en una imagen a partir de la detección de bordes. En la implementación, se sigue el siguiente proceso:

1. **Preprocesamiento:** Se convierte la imagen a escala de grises y se aplica un filtro Gaussiano para reducir el ruido.
2. **Detección de Bordes:** Se utiliza la detección de bordes (a menudo mediante el algoritmo de Canny) para identificar contornos en la imagen.

3. Aplicación de la Transformada: Se aplica la Transformada de Hough, especificando parámetros como la distancia de resolución y los umbrales para la detección de círculos, lo que permite localizar monedas de diferentes tamaños.

3.6.2 Algoritmo de Canny

El algoritmo de Canny es un método popular para la detección de bordes en imágenes. Su objetivo es identificar bordes significativos mediante un enfoque en varias etapas:

1. Reducción de Ruido: Se aplica un filtro Gaussiano para suavizar la imagen y reducir el ruido que podría afectar la detección de bordes.
2. Detección de Gradientes: Se calculan los gradientes de intensidad de la imagen utilizando operadores de Sobel, lo que ayuda a identificar áreas donde hay un cambio brusco en la intensidad, es decir, bordes.
3. Supresión de No Máximos: Se realiza una supresión para eliminar píxeles que no son bordes locales, manteniendo solo los píxeles que son máximos locales en la dirección del gradiente.
4. Umbralización: Se aplica un umbral doble para determinar qué bordes son verdaderos, permitiendo así identificar bordes fuertes y débiles. Los bordes débiles se conectan a los fuertes si están dentro de un rango predefinido.

En este proyecto, el algoritmo de Canny se utiliza como un paso preliminar para resaltar los bordes de las monedas, proporcionando así una base sólida para la Transformada de Hough, que identifica los círculos correspondientes.

La combinación de estos dos algoritmos permite una detección precisa y eficiente de las monedas en tiempo real, contribuyendo al objetivo principal de la aplicación.

4 Arquitectura de la aplicación

La arquitectura de la aplicación para la detección y conteo automático de monedas está diseñada para ser modular y extensible, utilizando una combinación de tecnologías de visión por computadora y herramientas de desarrollo de interfaces gráficas. La arquitectura se divide en varias capas, cada una responsable de diferentes aspectos del funcionamiento de la aplicación.

4.1 Visión general del sistema

El sistema se compone de varios componentes principales que trabajan en conjunto para ofrecer una solución integral para la detección y conteo de monedas. La interfaz de usuario (GUI), implementada con PyQt6, permite a los usuarios interactuar de manera intuitiva con la aplicación. El módulo de procesamiento de imágenes utiliza OpenCV para la captura y manipulación de imágenes, asegurando que los datos sean precisos y relevantes. A continuación, el módulo de detección y conteo aplica algoritmos específicos para identificar y contar las monedas mediante técnicas avanzadas de visión por computadora. Finalmente, el módulo de visualización presenta los

resultados del procesamiento, mostrando tanto las imágenes capturadas como las estadísticas del conteo de monedas en la GUI.

4.2 Interfaz de usuario (GUI)

La interfaz de usuario es fundamental para proporcionar una experiencia interactiva y amigable. Construida con PyQt6, está organizada en varias secciones clave. La sección de captura de imagen permite al usuario tomar imágenes directamente desde una cámara web o cargar archivos desde el sistema. A través de la sección de configuración, los usuarios pueden ajustar parámetros como la sensibilidad de detección y los umbrales de procesamiento. Por último, la sección de resultados muestra los resultados obtenidos, incluyendo las imágenes procesadas y las estadísticas correspondientes al conteo de monedas.

4.3 Módulo de procesamiento de imágenes

El módulo de procesamiento de imágenes, que utiliza OpenCV, realiza una serie de operaciones esenciales sobre las imágenes capturadas. Inicialmente, se lleva a cabo una conversión a escala de grises, lo que facilita la detección de bordes y formas. A continuación, se aplica un filtrado y suavizado que reduce el ruido en las imágenes, mejorando así la precisión del análisis posterior. En este contexto, se utiliza la detección de bordes mediante algoritmos como Canny para identificar los contornos de las monedas, complementado con la transformada de Hough, que permite detectar círculos correspondientes a las monedas.

4.4 Módulo de detección y conteo

Este módulo es el núcleo de la aplicación, encargado de la detección y conteo de monedas en las imágenes procesadas. Utiliza diversas técnicas para lograr su objetivo, comenzando con la detección de contornos, que identifica y extrae los bordes de las monedas presentes en la imagen. Posteriormente, se aplica un filtrado por tamaño, el cual elimina aquellos contornos que no cumplen con las dimensiones esperadas para una moneda. Finalmente, el módulo cuenta el número de contornos restantes, que corresponden a las monedas detectadas.

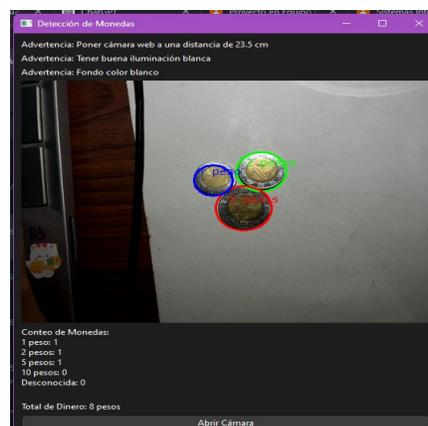


Figura 1. Detección de las monedas

Se muestra la detección de monedas de 1 peso (círculo azul), 5 pesos (círculo verde) y 2 pesos (círculo rojo), junto con el conteo total y el valor monetario calculado.

4.5 Módulo de visualización

El módulo de visualización se encarga de mostrar los resultados del procesamiento y el conteo en la interfaz de usuario [10]. Este módulo incluye la visualización de imágenes, que permite mostrar tanto la imagen original como la imagen procesada, destacando las monedas detectadas. Además, proporciona estadísticas de conteo, ofreciendo un resumen del número de monedas detectadas y su distribución.

4.6 Flujo de trabajo de la aplicación

El flujo de trabajo de la aplicación se puede resumir en varios pasos clave. Primero, el usuario puede capturar una imagen utilizando la cámara web o cargar una imagen desde el sistema de archivos. Una vez que la imagen es capturada, se procede al procesamiento de la imagen, donde se utilizan técnicas de visión por computadora para detectar los bordes y formas de las monedas. A continuación, el módulo de detección y conteo identifica y contabiliza las monedas en la imagen procesada. Por último, los resultados del conteo se muestran en la interfaz de usuario, junto con la imagen procesada, como se ilustra en la figura 2.

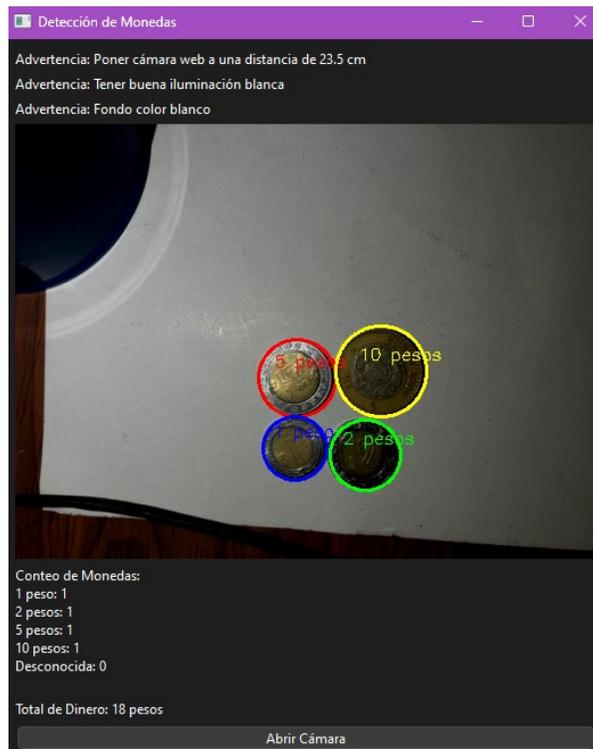


Figura 2. Detección y conteo de las monedas

Las monedas detectadas se destacan con diferentes colores según su valor: 1 peso (círculo azul), 2 pesos (círculo verde), 5 pesos (círculo rojo), y 10 pesos (círculo amarillo). El sistema también calcula y muestra el total de dinero detectado, que en este caso es 18 pesos.

5 Resultados

La aplicación desarrollada demostró un rendimiento robusto en la detección y análisis de monedas. Los resultados obtenidos incluyen:

5.1 Detección de monedas y valoración automática

La aplicación desarrollada demostró ser efectiva en la detección y clasificación de monedas de diferentes denominaciones (1 peso, 2 pesos, 5 pesos y 10 pesos en el contexto mexicano) en tiempo real, como se muestra en la figura 3. Los usuarios pueden abrir la cámara web, ajustar la distancia y la iluminación según las recomendaciones, y ver cómo la aplicación detecta y clasifica automáticamente monedas de diferentes denominaciones. Se incluyó un quinto elemento de detección, "Desconocido", para aquellos elementos detectados como monedas pero cuyo valor no puede ser identificado.

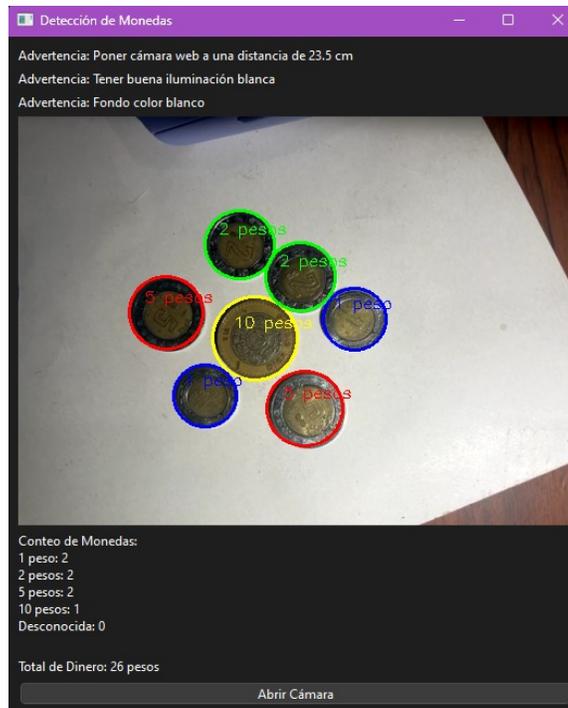


Figura 3. Detección de monedas y valoración

5.2 Interfaz gráfica intuitiva y responsiva

La interfaz gráfica fue desarrollada con PyQt6 proporciona una experiencia de usuario fluida y amigable. Los elementos de la interfaz se actualizan dinámicamente para reflejar el conteo de monedas y el valor de estas a medida que se procesan nuevos fotogramas de vídeo.

5.3 Pruebas de rendimiento

Para evaluar el rendimiento de nuestra aplicación, se comparó con "CountThings", una aplicación conocida por su capacidad de conteo automatizado de objetos, incluyendo monedas. A continuación, se presentan las principales diferencias y similitudes encontradas:

- **Precisión de detección:** Ambas aplicaciones mostraron una precisión comparable en la detección de monedas. Sin embargo, nuestra aplicación sobresale al tener la capacidad de detectar, contar y calcular el valor total de las monedas que aparecen en la imagen, cosa que "CountThings" no hace.
- **Velocidad de procesamiento:** A diferencia de "CountThings", que procesa imágenes estáticas y no en tiempo real, nuestra aplicación es capaz de hacerlo en tiempo real. Aunque es eficiente, podría beneficiarse de optimizaciones adicionales para mejorar la velocidad de detección en situaciones de alta demanda.
- **Interfaz de usuario:** La interfaz de nuestra aplicación se centra en la simplicidad y la claridad, mostrando directamente al usuario la imagen y la detección. En contraste, "CountThings" presentó una interfaz más compleja con características avanzadas de análisis y exportación de datos.

6 Conclusiones y trabajo a futuro

6.1 Conclusiones

Este trabajo ha presentado el desarrollo de una aplicación para la detección y conteo automático de monedas en tiempo real, utilizando las tecnologías OpenCV y PyQt6. A través de un análisis detallado y una implementación meticulosa, se lograron los siguientes aportes significativos:

1. **Eficiencia y precisión:** La aplicación demostró ser eficiente en la detección y conteo de monedas, mostrando resultados comparables a soluciones existentes, como CountThings, con una interfaz de usuario más intuitiva. Esto sugiere que la aplicación puede ser útil en entornos comerciales y educativos donde la rapidez y la facilidad de uso son prioritarias.
2. **Integración de tecnologías:** Al combinar OpenCV para el procesamiento de imágenes con PyQt6 para la creación de la interfaz gráfica, se estableció un modelo robusto que puede ser adaptado a otras aplicaciones de visión por computadora. Esta integración resalta la versatilidad de ambas herramientas en proyectos futuros.

6.2 Trabajo a futuro

Se sugiere que futuros desarrollos se centren en la mejora de la aplicación mediante la incorporación de algoritmos de aprendizaje automático para optimizar la detección y clasificación de monedas. Además, la implementación de un sistema de retroalimentación en tiempo real podría enriquecer la experiencia del usuario, permitiendo ajustes automáticos en el procesamiento según el contexto de uso.

También es recomendable explorar la posibilidad de ampliar la aplicación a otros objetos de interés en el ámbito de la visión por computadora, lo que podría abrir nuevas oportunidades en áreas como el inventario automatizado y la auditoría de monedas.

En conclusión, este trabajo ha proporcionado una base sólida para el desarrollo de aplicaciones prácticas en la detección de objetos mediante visión por computadora, y se espera que las futuras investigaciones continúen construyendo sobre estos hallazgos.

Referencias

[1] Gary Bradski, Adrian Kaehler et al. "OpenCV". En: Dr.Dobb's journal of software tools 3.2 (2000).

[2] Adrian Rosebrock. How-To: OpenCV and Python K-Means color Clustering. PyImageSearch. 2021. URL: <https://pyimagesearch.com/2014/05/26/opencv-python-k-means-color-clustering/>.

[3] Tomás Domínguez Mínguez. Visión artificial: aplicaciones prácticas con OpenCV-Python. Marcombo,2021.

[4] Enderson Ortega Guillina y Oscar Manuel Duque Suarez. "Diseño de interfaz usando visión artificial para fines de percepción de indicadores de interés para ambientes de enseñanza virtual en la población universitaria". En: Mundo FESC 8.15 (2018), págs. 31-40.

[5] OpenCV. Python: Displaying a video feed using OpenCV. https://docs.opencv.org/4.x/dd/d43/tutorial_py_video_display.html

[6] OpenCV. Gaussian, Median and Bilateral Filtering. https://docs.opencv.org/3.4/dc/dd3/tutorial_gaussian_median_blur_bilateral_filter.html

[7] OpenCV: Feature Detection. n.d. URL: https://docs.opencv.org/4.x/dd/d1a/group__imgproc__feature.html#ga47849c3be0d0406ad3ca45db65a25d2d

[8] Richard O. Duda y Peter E. Hart. "Use of the Hough Transformation to Detect Lines and Curves in Pictures". En: Communications of the ACM (1972).

[9] Teddy V Rojas, Wilmer Sanz y Francisco Arteaga. "Sistema de visión por computadora para la detección de objetos esféricos a través de la transformada de Hough". En: revista ingeniería UC 15.1 (2008), págs. 77-87.

[10] Reference Guide — PYQT Documentation v6.7.1. n.d. URL: <https://www.riverbankcomputing.com/static/Docs/PyQt6/>.

Detección y prevención de intrusos en sistemas operativos Linux con Suricata

Intrusion detection and prevention in Linux operating systems with Suricata

Vicente Zavaleta Sanchez¹, Judith Pérez Marcial¹, María del Carmen Santiago Díaz¹, Carina Toxqui Quitl², Gustavo Trinidad Rubín Linares¹, Ana Claudia Zenteno Vázquez¹

¹Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla
14 Sur y Avenida San Claudio, Cdad. Universitaria, 72592 Puebla, México

²Laboratorio de Visión por Computadora, Universidad Politécnica de Tulancingo
Calle Ingenierías 100, Huapalcalco, 43625, Hidalgo, México.

vicente.zavaletas@alumno.buap.mx, {judith.perez, marycarmen.santiago, gustavo.rubin, ana.zenteno}@correo.buap.mx, carina.toxqui@upt.edu.mx

Abstract

The cyberattacks are currently a constant and increasing threat, prompting the search for effective methods for their detection and mitigation. This research focuses on identifying and analyzing various types of cyberattacks, with a particular focus on Denial of Service (DoS) attacks and brute force attacks. To do so, Suricata, a freely available intrusion detection and prevention system (IDS/IPS) on the Linux platform, will be used. In addition to detection, the research seeks to develop effective methods to mitigate these attacks and reduce their impact on affected systems.

Resumen

En la actualidad, los ataques cibernéticos representan una amenaza constante y en aumento, lo que impulsa la búsqueda de métodos efectivos para su detección y mitigación. Esta investigación se centra en identificar y analizar diversos tipos de ciberataques, con un enfoque particular en los ataques de Denegación de Servicio (DoS) y los ataques de fuerza bruta. Para ello, se utilizará Suricata, un sistema de detección y prevención de intrusos (IDS/IPS, por sus siglas en inglés) disponible de forma gratuita en la plataforma Linux. Además de la detección, la investigación busca desarrollar métodos eficaces para mitigar estos ataques y reducir su impacto en los sistemas afectados.

Keywords and phrases: Ciberataque, DoS, Fuerza Bruta, IDS, IPS, Linux, Suricata.

1 Introducción

La mitigación de ciberataques es fundamental para que diversos servidores puedan seguir operando y ofreciendo sus servicios de manera segura y eficiente. Como bien se sabe, los ataques cibernéticos no solo pueden causar la interrupción de los servicios, sino que también pueden comprometer la integridad y confidencialidad de los datos. Por lo tanto, es necesario poder detectar y prevenir estos ataques en tiempo real.

En este contexto, el uso de herramientas como Suricata, que ofrece la posibilidad tanto de detectar como de prevenir intrusos es de suma importancia, permitiendo implementar medidas contra estos ataques antes de que el daño sea significativo.

2 Marco teórico y estado del arte

En este trabajo se proporciona una base sólida para comprender los fundamentos necesarios alrededor de la detección y prevención de intrusos en sistemas Linux utilizando Suricata.

2.1 Ataque cibernético

Los ataques cibernéticos son amenazas muy comunes en el entorno tecnológico y cada uno de estos puede tener diferentes objetivos, que van desde denegar un servicio hasta el acceso no autorizado. Según IBM, estos ataques se definen como esfuerzos para robar, exponer, modificar, deshabilitar o destruir datos, aplicaciones u otros activos mediante el acceso no autorizado a redes, sistemas informáticos o dispositivos digitales [1].

Los ataques de denegación de servicio (DoS), junto con otros tipos de ataques como los de fuerza bruta, suelen emplearse de manera combinada. Esto según Microsoft Threat Intelligence, ya que los ataques DoS pueden desviar la atención para facilitar la ejecución de otras acciones maliciosas. [2].

2.2 Ataque de denegación de servicio (DoS)

Un ataque de denegación de servicio (DoS) consiste en el intento malintencionado de saturar un servidor o servicio con una cantidad excesiva de tráfico, esto con el objetivo de interrumpir o suspender su operación normal [3].

Por mencionar algunos, los protocolos que suelen usarse para este tipo de ataques son los *ICMP* (usual en ataques como *ping flood*) y *TCP* (ataques que hacen uso del *HTTP GET*).

2.3 Ataque de fuerza bruta

Un ataque de fuerza bruta emplean un enfoque de prueba y error para lograr adivinar credenciales de acceso, claves de cifrado o para localizar recursos que se encuentran ocultos en un sistema o página web [4].

Estos ataques suelen ser eficaces debido a la debilidad de las contraseñas que suelen ser muy comunes.

2.4 Sistema de detección y prevención de intrusiones (IDS/IPS)

Los sistemas de detección y prevención de intrusiones (IDS/IPS o IDPS) son sistemas muy utilizados en el entorno de seguridad de redes. Un sistema de detección de intrusiones (IDS) es una herramienta se encarga de monitorizar el tráfico de red y dispositivos con el fin de lograr identificar actividad maliciosa, comportamientos sospechosos o violaciones a las políticas de seguridad [5].

En cambio, un sistema de prevención de intrusiones (IPS), no solo detecta posibles amenazas, sino que también toma medidas para bloquearlas, alertando al equipo de seguridad y finalizando conexiones o procesos peligrosos [6].

La principal diferencia entre estos dos radica en la capacidad de poder mitigar las amenazas.

2.5 Suricata

Suricata es un sistema de detección y prevención de intrusiones (IDS/IPS) desarrollado por la Open Information Security Foundation (OISF). Este permite la inspección profunda de paquetes y la detección de anomalías en el tráfico de red en tiempo real. Además de sus capacidades de IDS, Suricata también puede actuar en modo IPS, bloqueando amenazas potenciales antes de que afecten al sistema [7].

2.6 Reglas

Para poder detectar y actuar ante posibles amenazas Suricata, al igual que otros sistemas de detección y prevención de intrusiones, utiliza un lenguaje de reglas que permite la creación de reglas complejas para la detección de tráfico malicioso [8].

Estas reglas constan de tres partes: la acción, el encabezado y las opciones. La acción especifica qué ocurrirá si la regla coincide, el encabezado define el protocolo, las direcciones IP, los puertos, y la dirección del tráfico, mientras que las opciones definen los detalles de la regla [9].

Como ejemplo tomaremos la tabla 1 para entender mejor su estructura.

Tabla 1. Estructura de una regla en Suricata.

| Acción | Encabezado | | | | | | Opciones |
|--------|------------|--------------|--------|-------------------|-----------------|--------|--|
| | Protocolo | Dirección IP | Puerto | Dirección trafico | Dirección IP | Puerto | |
| alert | tcp | any | any | -> | 192.168.209.132 | 80 | (msg:"Posible Ataque de Ping Flood"; flow: from_client, stateless; threshold: type both, track by_src, count 10, seconds 600; classtype:attempted-PF ; sid:1000001; rev:1;) |

Esta regla genera una alerta cuando se detecta tráfico *TCP* desde cualquier dirección *IP* y puerto hacia la dirección 192.168.209.132 en el puerto 80. El mensaje que muestra la alerta es "Solicitud al servidor *HTTP*", además de que está clasificada como "not-suspicious" y tiene un ID único que en este caso viene siendo el 1000001 siendo esta la versión 1 de su revisión.

Las acciones que admite Suricata son:

- *alert*: Genera una alerta.
- *pass*: Detiene la inspección adicional del paquete.
- *drop*: Descarta el paquete y genera alerta.
- *reject*: Envía el error de no alcance al remitente del paquete.

Cabe recalcar que para usar las acciones *drop* y *reject*, Suricata debe operar en modo IPS. Este modo permite a Suricata interceptar el tráfico de red a través de una interfaz configurada en modo *inline* usando *NFQUEUE* e *IPTables*. De lo contrario, solo generará alertas sin tomar acciones. Todas las acciones, excepto *pass*, se registrarán en el archivo `/var/log/suricata/fast.log`.

En el apartado del encabezado, Suricata tiene soporte para varios protocolos, por mencionar algunos están los protocolos *TPC*, *UDP*, *ICMP* e *IP*.

En la regla podemos observar dos direcciones *IP* junto a su número de puerto, la primera dirección *IP* y número de puerto indica el origen del tráfico y la segunda dirección *IP* y número de puerto indican el destino del tráfico.

La dirección de tráfico, como su nombre lo menciona, indica la dirección en que se evaluará la regla, usualmente la mayoría de las reglas se evalúan con dirección a la derecha (*->*), pero también es posible que la regla coincida en ambas direcciones (*<->*).

El apartado de opciones se encuentra siempre entre paréntesis y están separadas por punto y coma.

```
(msg:"Posible Ataque de Ping Flood"; flow: from_client, stateless; threshold: type both, track by_src, count 10, seconds 600; classtype:attempted-PF; sid:1000001; rev:1;)
```

Las opciones mínimas recomendadas son el *msg*, *sid* y *rev*, donde *msg* define el mensaje que me mostrará la alerta, *sid* es el identificador único de la regla y *rev* indica el número de versión o revisión de la regla.

En esta regla podemos ver también otras opciones como:

- *classtype*: Especifica el tipo tráfico (sospechoso o normal) y su prioridad (usualmente van del 1 al 5, donde 1 tiene la prioridad más alta). Las reglas con prioridad 1 siempre se activará si sus condiciones se cumplen, omitiendo cualquier otra regla.
- *flow*: Este junto con las instrucciones *from_client* y *stateless*, indica el flujo de datos que se evalúa, en este caso los que se dirigen al servidor independientemente del estado en que se encuentre la conexión.

- *threshold*: Permite establecer un límite para activar la regla. Las instrucciones *type both*, *track by_src*, *count 10* y *seconds 600* nos indica su configuración, ya que la regla analizara tanto las conexiones entrantes como salientes, dándole seguimiento a la dirección *IP* de origen, además que, si se detectan 10 eventos o más provenientes de la misma dirección *IP* en un período de 600 segundos, se activará la regla.

2.8 IPTables

En los sistemas operativos Linux, el firewall que comúnmente forma parte del núcleo es IPTables, donde su función principal es la de filtrar y controlar el tráfico de red que entra y sale de un sistema [10].

Un aspecto importante de IPTables, es que su capacidad para poder realizar acciones específicas está basada en reglas definidas por el usuario. Las reglas de IPTables se organizan en tablas, cada una de las cuales está formada por un conjunto de condiciones (coincidencias) que, al cumplirse, desencadenan una acción específica (objetivo) sobre el tráfico de red [11].

2.9 NFQUEUE

NFQUEUE es un subsistema que transfiere la responsabilidad de evaluar y tomar decisiones sobre los paquetes a aplicaciones de usuario, como los sistemas de prevención de intrusiones (IPS) [12].

Este subsistema es crucial ya que es el intermediario entre IPTables y Suricata. IPTables redirigirá el tráfico de red a NFQUEUE, que a su vez lo envía a Suricata. Una vez Suricata analice y tome su decisión, envía su respuesta a NFQUEUE y este le indicara a IPTables si debe permitir o bloquear los paquetes de datos.

2.10 SlowHTTPTest

SlowHTTPTest es una herramienta que simula diferentes tipos de ataques DoS a nivel de aplicación. Uno de los ataques más destacados es el de tipo *Slow HTTP*, en el que se prolongan las conexiones *HTTP* para sobrecargar los servidores web [13].

2.11 Hydra

Hydra es una herramienta utilizada para realizar ataques de fuerza bruta capaz de probar credenciales en múltiples servicios de autenticación simultáneamente [14].

Para realizar los ataques Hydra hace uso de diccionarios de usuarios y contraseñas los cuales combina y prueba hasta que este logre su cometido o se acaben las combinaciones posibles. La duración del ataque es variable y depende en gran medida de la capacidad de procesamiento del equipo, la velocidad de la conexión a internet y la calidad de los diccionarios empleados [15].

3 Diseño de laboratorio de pruebas

En este apartado se encuentran los diagramas de laboratorio implementados y su respectiva explicación.

La Figura 1 muestra el diseño del laboratorio utilizado para las pruebas. Se empleó una Laptop Acer Aspire 5 con procesador AMD Ryzen 7 5700U, 20 GB de RAM y 475 GB de SSD, con Windows 11 Home 23H2 y VMware Workstation 17 Pro 17.6.0. Se crearon dos máquinas virtuales, una atacante y otro objetivo, ambas con 4 núcleos, 4 GB de RAM y 25 GB de almacenamiento. La máquina objetivo usa Ubuntu 24.04.1 y tiene instalado las herramientas y servicios como IPTables, NFQUEUE, Suricata, Apache2 y Vsftpd, mientras que la máquina atacante usa Kali Linux 2024.3 y cuenta con SlowHTTPTest y Hydra.

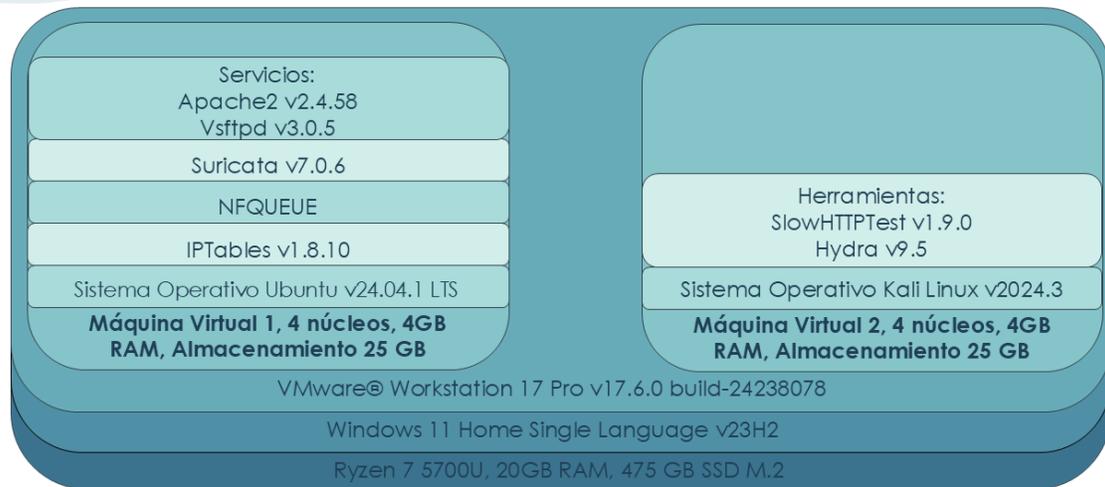


Figura 1. Software y hardware utilizado para llevar a cabo las pruebas de laboratorio.

En la figura 2, se puede observar un diagrama que representa la prueba de comunicación entre las dos máquinas virtuales mediante el uso de Ping. La máquina virtual 2 es quien envía un mensaje *ICMP Echo Request* hacia la máquina virtual 1, cuando la máquina virtual 1 recibe el mensaje este responde con un *ICMP Echo Reply*, este proceso verifica que haya comunicación entre las dos máquinas virtuales.

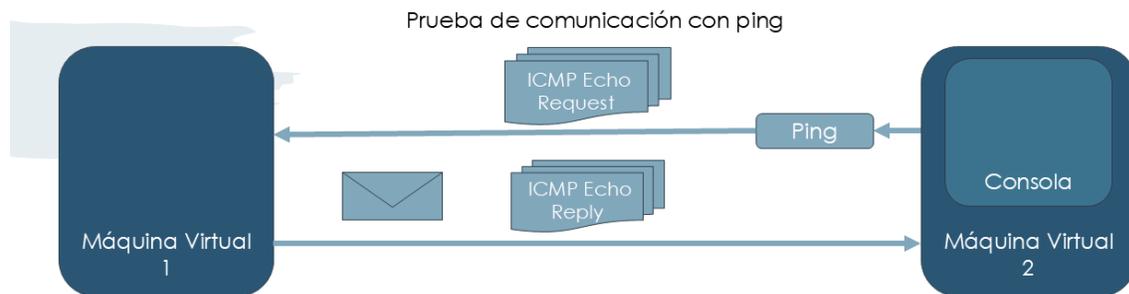


Figura 2. Prueba de comunicación con ping

En la figura 3, se puede ver la interacción de Suricata en modo IDS con tráfico de red. Suricata que se encuentra instalado en la máquina virtual 1, monitorea el tráfico de red que hay entre la máquina

virtual 1 y la máquina virtual 2, si alguna de las reglas es activada durante el monitoreo, Suricata registrara los eventos en un archivo log sin interrumpir o bloquear el tráfico de red.

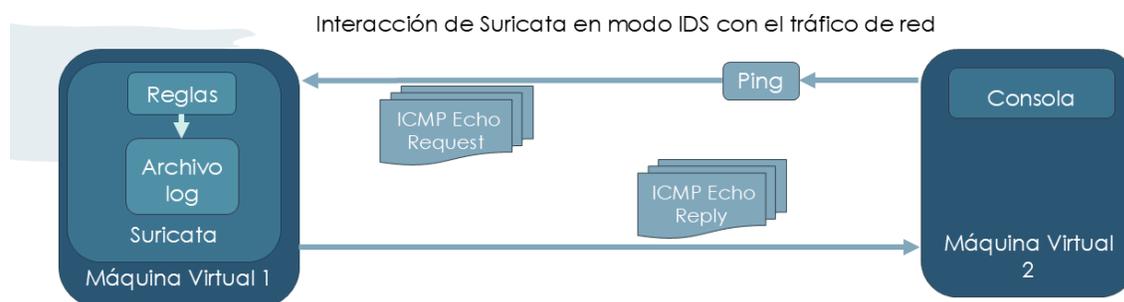


Figura 3. Interacción de Suricata en modo IDS con el tráfico de red.

La figura 4, muestra la interacción de Suricata en su modo IPS con el tráfico de red. En este modo, IPTables se encarga de redirigir el tráfico de red a NFQUEUE, donde Suricata intercepta y analiza los paquetes, y basado en sus reglas, decide si los paquetes continúan o si son registrados y bloqueados evitando que lleguen a su destino. Esta decisión tomada por Suricata es devuelta a NFQUEUE que a su vez le comunica a IPTables que hacer con el tráfico de red.



Figura 4. Interacción de Suricata en modo IPS con el tráfico de red.

En la figura 5, se puede observar una prueba de comunicación con ping y Suricata donde se hace el uso de la consola nos permitirá visualizar los mensajes generados y guardados en el archivo log, alertándonos sobre posibles ataques y mitigación de estos, al mismo tiempo en que las reglas configuradas en Suricatas son activadas.

En la figura 6, podemos observar un diagrama representando como sería un Ataque DoS dirigido hacia una maquina con Suricata. El ataque DoS es lanzado desde la máquina virtual 2 utilizando la herramienta SlowHTTPTest, que envía múltiples solicitudes *HTTP GET* basados en el protocolo *TCP* al servidor Apache2 de la máquina virtual 1.



Figura 5. Prueba de comunicación con ping y Suricata

Antes de que las solicitudes lleguen al servidor, estas son analizadas por Suricata, donde si las solicitudes son las normalmente esperadas durante un periodo de tiempo, este permitirá que lleguen al servidor Apache2, pero si Suricata detecta alguna anomalía, como el exceso de solicitudes de las que una persona realiza durante ese mismo periodo de tiempo, este no solo notificará el evento en consola y registrará los detalles en su archivo de log, sino que bloqueará todas las solicitudes pendientes y terminará las conexiones activas provenientes de esa misma dirección IP haciendo uso de IPTables.

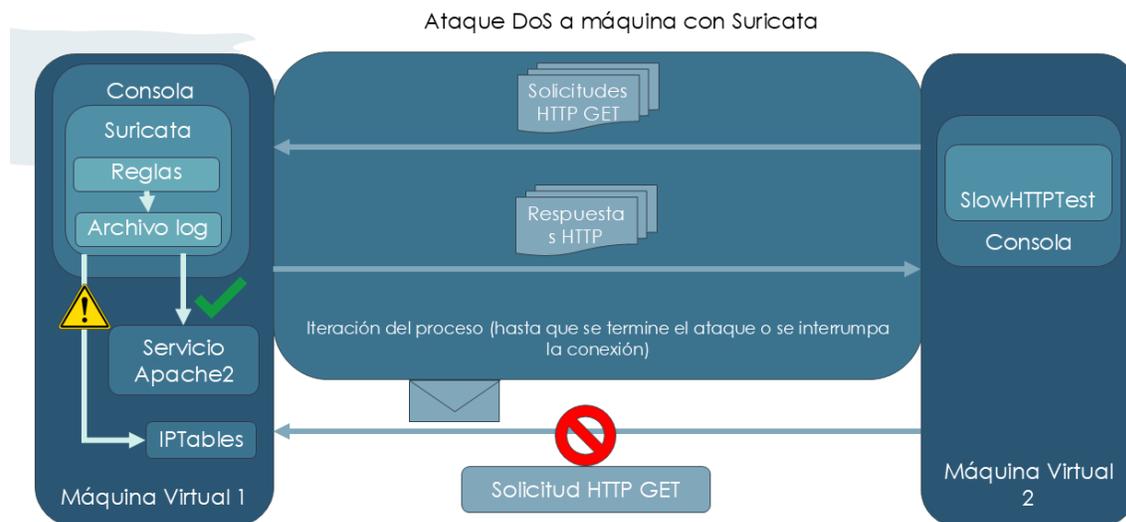


Figura 6. Ataque DoS a máquina con Suricata.

En la figura 7, se muestra un diagrama que representa un Ataque de fuerza bruta utilizando la herramienta Hydra. Este ataque emplea diccionarios para intentar obtener acceso no autorizado al servidor VSFTPD, enviando múltiples solicitudes de inicio de sesión a través del protocolo TCP/IP,

donde Hydra intentara diferentes combinaciones de usuario y contraseña hasta que se logre una autenticación exitosa o se agoten las contraseñas disponibles de los diccionarios. Suricata al detectar un numero anormal de intentos de inicio de sesión, notificara la actividad en consola y registrara los detalles en su archivo de log, al mismo tiempo que termina y bloquea tanto la conexión establecida como las futuras solicitudes del atacante mediante el uso de IPTables.



Figura 7. Ataque de fuerza bruta a máquina con Suricata.

4 Resultados

En este apartado se encuentran los resultados obtenidos con su respectiva explicación.

En la figura 8, se puede observar que Suricata que se encuentra instalada en la máquina virtual de la izquierda, logra registrar, alertar y mitigar la prueba de comunicación con ping de manera exitosa, mientras que la máquina virtual de la derecha es quien realiza la prueba de comunicación, donde después de enviar el paquete *ICMP* número 15, este ya no recibe respuesta teniendo como perdida 16 de los 30 paquetes que transmitió.

En la figura 9, se puede observar el funcionamiento de Suricata instalada en la máquina virtual de la izquierda, en donde se visualiza el registro que confirma la detección y mitigación de un ataque DoS dirigido al puerto 80, que generalmente es utilizado por el servidor web *HTTP* de Apache2, el cual podemos acceder sin problemas desde el navegador, mientras que, la máquina virtual de la derecha, que es la responsable de realizar el ataque con ayuda de la herramienta *SlowHTTPTest*, se muestra diferentes detalles, como las 968 conexiones que quedaron pendientes, las 30 conexiones que fueron cerradas y un mensaje de que el servicio no está disponible, además, de una ventana de Firefox que muestra un mensaje de error al no obtener respuesta del servidor *HTTP*.

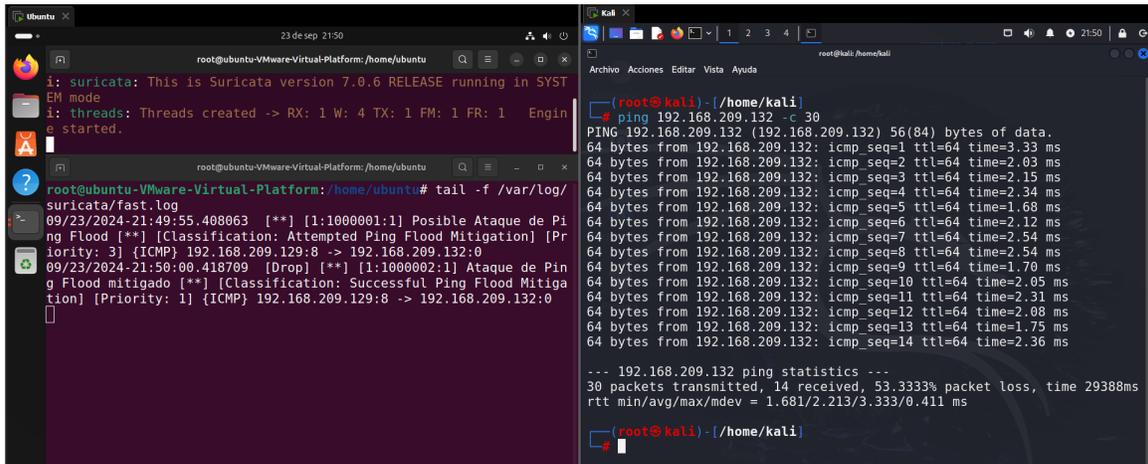


Figura 8. Alerta y mitigación de prueba de comunicación con Ping.

Este resultado ocurre porque Suricata, al detectar el tráfico de red anormal proveniente la misma dirección IP, bloquea nuevas peticiones y cierra las conexiones ya realizadas, asegurándonos así el correcto funcionamiento del servidor y la efectividad de Suricata.

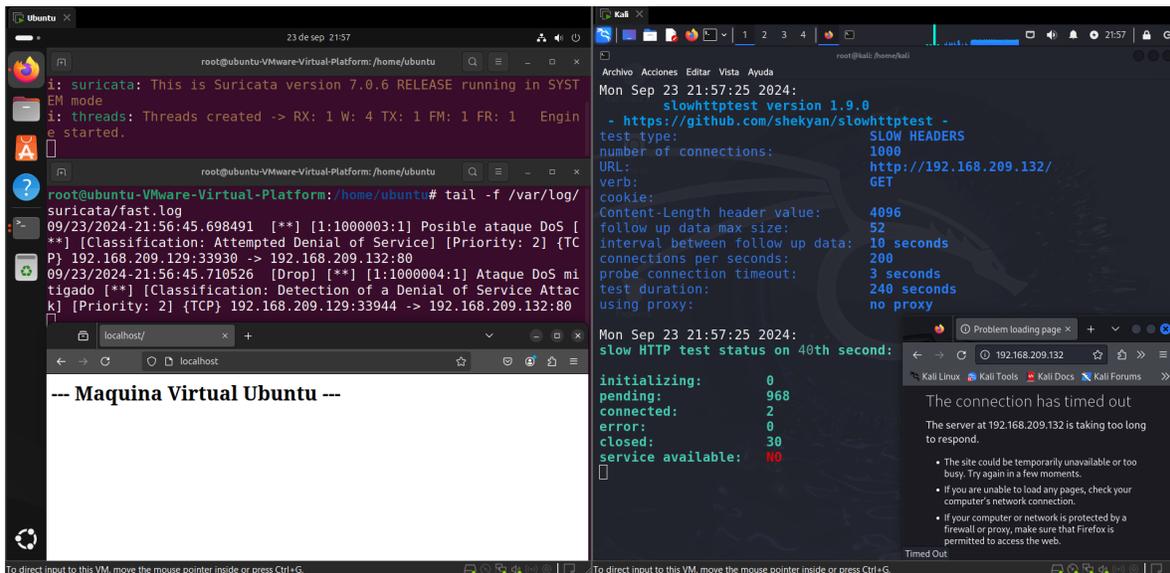


Figura 9. Alerta y mitigación de ataque DoS.

En la figura 10, se muestra la respuesta de Suricata que se encuentra en la máquina virtual de la izquierda, ante el ataque de fuerza bruta lanzado desde la máquina virtual de la derecha. El registro de Suricata podemos observar que se generó una alerta y que el ataque fue mitigado, ambos cuentan con información detallada, como el numero de la regla activada, un mensaje descriptivo de lo que este representa, la prioridad de la regla, el protocolo usado, la dirección IP y puerto de origen del ataque, así como la dirección IP y puerto de destino.

En la consola de la máquina virtual de la derecha, podemos observar que el ataque se realizó con la herramienta Hydra, y muestra la cantidad de intentos que realizó y que este ya no pudo continuar, finalizando su tarea sin éxito. Por lo tanto, Suricata logró mitigar este ataque exitosamente.

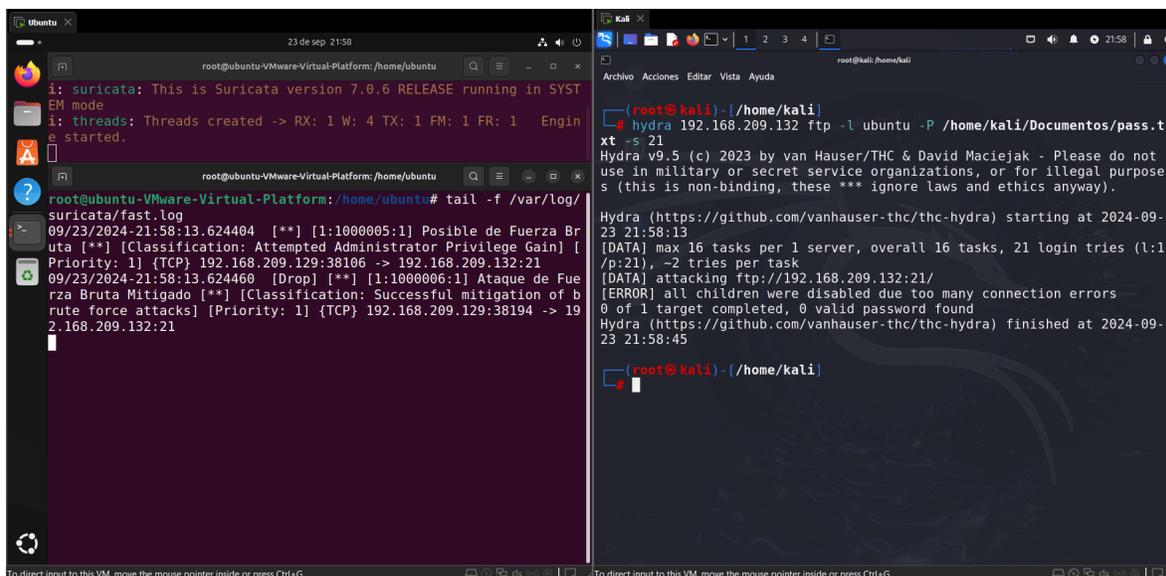


Figura 10. Alerta y mitigación de ataque de fuerza bruta.

5 Conclusiones y trabajo a futuro

En base a los resultados obtenidos, se puede decir que Suricata al estar configurado correctamente y tener las reglas necesarias con las condiciones adecuadas, puede ayudarnos a detectar, registrar, alertarnos y mitigar diferentes tipos de ataques que se puedan realizar hacia nuestra red, tales como ataques de denegación de servicio (DoS) o ataques de fuerza bruta. Sin embargo, cabe recalcar que la capacidad de detectar estos ataques depende en gran medida de las reglas utilizadas. Por lo tanto, es de suma importancia que estas reglas estén en constante actualización para que Suricata pueda detectar y actuar ante estas amenazas emergentes.

Agradecimientos

Quiero expresar mi más sincero agradecimiento a mi encargada, Ana Claudia Zenteno Vázquez, por su invaluable guía, paciencia y apoyo durante el desarrollo de este trabajo. Sus consejos y conocimiento han sido fundamentales para llevar a cabo este proyecto.

A mis padres quienes siempre me apoyaron y motivaron, lo que me han permitido llegar hasta este punto. Gracias por creer en mí y ser mis ejemplos a seguir.

Finalmente quiero agradecer a mis amigos, por su compañía, palabras de aliento y por brindarme la motivación necesaria para seguir adelante. Su amistad ha sido un gran apoyo durante todo este proceso.

Referencias

- [1] “¿Qué es un ataque cibernético?”. [Online]. Available: <https://www.ibm.com/mx-es/topics/cyber-attack>. [Accessed: 16/Septiembre/2024].
- [2] “Rise in XorDdos: A deeper look at the stealthy DDoS malware targeting Linux devices”. [Online]. Available: <https://www.microsoft.com/en-us/security/blog/2022/05/19/rise-in-xorddos-a-deeper-look-at-the-stealthy-ddos-malware-targeting-linux-devices/>. [Accessed: 16/Septiembre/2024].
- [3] “¿Qué es un ataque de denegación de servicio (DoS)?”. [Online]. Available: <https://www.cloudflare.com/es-es/learning/ddos/glossary/denial-of-service/>. [Accessed: 16/Septiembre/2024].
- [4] “Ataque de fuerza bruta: Definición y ejemplos”. [Online]. Available: <https://www.kaspersky.es/resource-center/definitions/brute-force-attack>. [Accessed: 16/Septiembre/2024].
- [5] “¿Qué es un sistema de detección de intrusiones (IDS)?”. [Online]. Available: <https://www.ibm.com/mx-es/topics/intrusion-detection-system>. [Accessed: 16/Septiembre/2024].
- [6] “¿Qué es un sistema de prevención de intrusiones (IPS)?”. [Online]. Available: <https://www.ibm.com/mx-es/topics/intrusion-prevention-system>. [Accessed: 16/Septiembre/2024].
- [7] “COMMUNITY DRIVEN OPEN SOURCE”. [Online]. Available: <https://oisf.net/>. [Accessed: 16/Septiembre/2024].
- [8] “Sistemas de Detección de Intrusos Suricata Open Source como mecanismo de seguridad corporativa en entornos libres, 2020”. [Online]. Available: <https://repositorio.unsch.edu.pe/server/api/core/bitstreams/0c07301a-8549-4a91-b5e5-2dc9649ed9ac/content>. [Accessed: 16/Septiembre/2024].
- [9] “Rules Format”. [Online]. Available: <https://docs.suricata.io/en/latest/rules/intro.html>. [Accessed: 16/Septiembre/2024].
- [10] “Seguridad en Sistemas Operativos Linux”. [Online]. Available: https://repositorio.unipiloto.edu.co/bitstream/handle/20.500.12277/11573/Seguridad_Sistemas_operativos_Linux.pdf?sequence=1&isAllowed=y. [Accessed: 16/Septiembre/2024].
- [11] “The netfilter.org project”. [Online]. Available: <https://www.netfilter.org/>. [Accessed: 16/Septiembre/2024].

[12] “IPS mode using NFQ”. [Online]. Available: <https://docs.mirantis.com/mcp/q4-18/mcp-security-best-practices/use-cases/idps-vnf/ips-mode/nfq.html>. [Accessed: 16/Septiembre/2024].

[13] “SlowHTTPTest”. [Online]. Available: <https://github.com/shekyan/slowhttpstest>. [Accessed: 16/Septiembre/2024].

[14] “Hydra”. [Online]. Available: <https://www.kali.org/tools/hydra/>. [Accessed: 16/Septiembre/2024].

[15] “Laboratorio Virtualizado de Seguridad Informática con Kali Linux”. [Online]. Available: <https://uvadoc.uva.es/bitstream/handle/10324/5792/TFG-B.511.pdf?sequence=1&isAllowed=y>. [Accessed: 16/Septiembre/2024].

Identificación de tendencias suicidas utilizando redes neuronales

Identification of suicidal tendencies employing neural networks

Ana Paola Rojas Martínez, Maya Carrillo Ruíz, Rogelio Gonzales Velázquez

Facultad de Ciencias de la Computación, Benemérita Universidad Autónoma de Puebla,
Av. San Claudio, Boulevard 14 Sur, Ciudad Universitaria, C.P. 72592. Heroica Puebla de
Zaragoza, Puebla, México.

rm223470501@alm.buap.mx, maya.carrilloruiz@viep.com.mx,
rogelio.gonzalez@correo.buap.mx

Abstract

The alteration of personal well-being results in consequences that can affect each person differently, one of the problems that can be triggered is suicide; therefore, many people who are thinking about taking their own lives or who have attempted it do not seek help, becoming a public health problem; according to the UN, few countries include suicide prevention in their health priorities and only 38 countries report having a national prevention strategy [2]. As an analysis strategy to prevent suicide, a method is proposed that develops the incorporation of a sequential neural network for the analysis of texts in Spanish, seeking the identification of suicidal tendencies to have a classification of those who show signs of suicidal tendencies and those who do not, finally it is tested with a corpus analyzing its effectiveness and results are compared with previous works.

Resumen

La alteración del bienestar personal resulta en consecuencias que pueden afectar de diferente manera a cada persona, uno de los problemas que puede desencadenar es el suicidio; por ello, muchas personas que están pensando en quitarse la vida o que lo han intentado no buscan ayuda convirtiéndose en un problema de salud pública; según la ONU, pocos países incluyen la prevención de suicidio en sus prioridades de salud y solo 38 países informan tener una estrategia nacional de prevención [2]. Como una estrategia de análisis para prevenir el suicidio, se propone un método que desarrolla la incorporación de una red neuronal secuencial para el análisis de textos en español, buscando la identificación de tendencias suicidas para tener una clasificación sobre aquellos que presentan indicios de tendencias suicidas y aquellos que no, finalmente es probada con un corpus analizando su efectividad y se comparan resultados frente a trabajos anteriores.

Keywords and phrases: Red Neuronal, Nuicidio, Clasificación, Efectividad.

1 Introducción

La salud mental es un estado de bienestar personal que permite afrontar el estrés de la vida, lleva al desarrollo de capacidades, aprendizajes y a la contribución a la comunidad; diversos factores pueden llevar a fortalecer o socavar la salud mental como la familia, la comunidad, factores individuales y estructurales [1]. Al existir una alteración en el estado de bienestar con sucesos como problemas financieros, ruptura de relaciones, enfermedad, conflictos, desastres, violencia, abuso, pérdida y aislamiento, están fuertemente asociados con el comportamiento suicida.

Actualmente, aún existe un estigma en torno a los trastornos mentales y al suicidio, muchas personas que están pensando en quitarse la vida o que lo han intentado no buscan ayuda; la prevención no es abordada adecuadamente debido a la falta de concientización sobre el suicidio como un importante problema de salud pública. Hasta la fecha, según la ONU, pocos países han incluido la prevención de suicidio en sus prioridades de salud y solo 38 países informan tener una estrategia nacional de prevención del suicidio [2].

Como parte de las acciones de prevención de suicidio que pueden llevarse a cabo se encuentra en la aplicación de la tecnología, más específicamente el método propuesto es la utilización de redes neuronales para realizar la clasificación de textos en español cuyo estudio no es tan amplio como sí lo son los textos en inglés. El método consiste en el análisis de los escritos respecto a las palabras empleadas por los autores, es conocido que las palabras que empleamos al escribir y los temas de los que hablamos dependen mucho de nuestras emociones y pensamientos y no solo de la amplitud de nuestro léxico; por ello, el análisis de los textos nos lleva a encontrar ciertos patrones o características recurrentes que nos permitan realizar una clasificación que dependa de nuestro tema de interés. En este caso las tendencias suicidas detectadas en textos, con el propósito de contribuir a futuras aplicaciones cuyo objetivo sea la prevención del suicidio.

El presente documento, está formado por cuatro secciones entre las cuales tenemos la sección dos, siendo el marco teórico y estado del arte; la sección tres, con el método propuesto, la sección cuatro con los resultados y finalmente la sección cinco que contiene las conclusiones y trabajo futuro.

2 Marco teórico y estado del arte

El suicidio es un problema importante de salud pública cuya prevención debe ser llevada a cabo por diferentes frentes, si bien la intervención de las personas que contribuyen a formar el estado de bienestar personal es importante, así como también la inclusión de especialistas como médicos, psicólogos y psiquiatras, hoy en día existe la opción de incluir otras herramientas que sean de ayuda en la detección de esos casos.

Fuchiang R. et al. [4] aplican análisis de textos de registro médico con cuatro algoritmos y compararon sus resultados, siendo estos Naive Bayes (NB), regresión del operador de selección y contradicción mínima absoluta (LASSO por sus siglas en inglés), bosque aleatorio (Random Forest, RF), y el conjunto de aumento de gradiente extremo (Ensemble of extreme gradient boosting,

EXGB). Todos los modelos estimaron la probabilidad posterior de intento de suicidio basándose en datos históricos extraídos dentro de los 2 años anteriores a la visita índice.

Coppersmith G. et al., [3] demuestran la viabilidad del uso de datos de redes sociales para detectar conductas de riesgo suicida, mediante técnicas de procesamiento de lenguaje natural y aprendizaje profundo. De un conjunto de datos combinado de usuarios de redes sociales, realizaron un análisis de los textos otorgando probabilidades según el algoritmo, la cual es más alta según el riesgo de suicidio. El modelo primero utiliza una capa de incrustación de palabras para proyectar cada palabra en un espacio vectorial denso diseñado de tal manera que las palabras semánticamente similares permanecen a poca distancia entre sí en el espacio euclidiano. El modelo se inicia con incrustaciones GloVe previamente entrenadas y se perfeccionan durante el proceso de entrenamiento del modelo para cambiar a matices del lenguaje relacionados con la salud mental. Las secuencias de vectores de palabras se procesan a través de una capa bidireccional de memoria a corto plazo (LSTM por sus siglas en inglés) para capturar información contextual entre palabras. La salida se combina con un vector mediante conexiones de salto en una capa de autoatención, la cual se utiliza para generar ponderaciones al paso del tiempo de la secuencia y finalmente una capa lineal con salida softmax predice la probabilidad.

Continuando con el análisis de datos obtenidos de redes sociales, en el documento [10] los autores analizan una base de datos obtenida del sitio web Reddit, específicamente de SuicideWatch diseñado para que los usuarios compartan pensamientos, ideas y formas relacionadas con el suicidio. Realizando un preprocesamiento de textos y entrenando una red neuronal programada en Python obtuvieron resultados de un 92% de exactitud en su caso 1 de prueba, en el caso 2 un 89% y un 87% para su último caso de prueba.

En un estudio del 2023 de igual manera se usaron datos de redes sociales (Twitter) [11] los autores estudiaron el Leet Speaking para identificar palabras ofuscadas cuyo significado indicara tendencias suicidas. Utilizando el modelo pre-entrenado BERT (Bidirectional Encoder Representations from Transformers) para clasificar tuits con etiquetas “suicidal” y “NotSuicidal” alcanzaron una precisión del 95.58%.

Un trabajo más, relacionado con la red social Twitter es del 2023 de [12] donde los autores clasifican datos de redes sociales referentes al suicidio como: pensamiento suicida activo, pensamiento suicida pasivo, sarcasmo relacionado con el pensamiento suicida y tuits relacionados con el suicidio (conciencia sobre el suicidio, noticias, charlas). Con un total de 81519 tuits aplicando preprocesamiento, realizaron el entrenamiento del modelo BERT (por sus siglas en inglés) y Roberta obteniendo resultados de precisión 96.68% y 95.8% respectivamente.

Finalmente, un estudio más con datos obtenidos de Twitter es [13] del 2023 en donde empleando el procesamiento de lenguaje natural identifican patrones o indicadores de riesgos de suicidio. Empleando *machine learning* y una variedad de técnicas de regresión su sistema detecto ideas suicidas con un 94.02% de precisión para métricas de SVM, de 93.7% para regresión logística y finalmente para un modelo de CNN obtuvieron un 92.09%.

Debido a la importancia del tema, en [14] se analizan las ideas suicidas en personal militar. Usando una base de datos pública de la plataforma Reddit y utilizando técnicas de procesamiento de lenguaje natural, los autores con una red de aprendizaje profundo llamada perceptrón multicapa obtuvieron resultados de identificación de tendencias suicidas con una precisión del 92.22%.

3 Método propuesto

Para el análisis de texto se necesita tener una representación que facilite su interpretación según las características de nuestro interés, el idioma, dependiendo del que se hable está compuesto de cierta gramática, expresiones, grafía y signifiante que permiten una estructura coherente que facilita la comunicación entre quienes lo utilizan; algunas de estas características son tomadas en cuenta en el análisis de textos mediante un computador y por ello es necesario transformar un texto a una representación que pueda ser fácilmente analizada pero que no pierda la esencia del texto original. En este trabajo se optó por una representación vectorial para el *corpus*. No obstante, es necesario realizar un conjunto de operaciones conocidas como preprocesamiento.

Debido a que la interpretación de texto por una máquina conlleva un análisis previo y su transformación para una representación que puede ser vectorial, la cual consiste en las siguientes operaciones:

3.1 Lowercase

Debido a que el idioma español tiene una doble representación que son las mayúsculas y minúsculas de las letras que conforman su alfabeto, es recomendable cambiar las letras mayúsculas a minúsculas antes de cualquier proceso. [5]

3.2 Tokenización

Este paso consiste en seccionar una secuencia de texto en partes más pequeñas, esto puede ser por caracteres, palabras, oraciones, o a conveniencia del texto que se analice. Para este trabajo se decidió hacer una Tokenización por palabras para obtener un vocabulario del corpus.

3.3 Remoción de signos de puntuación y de elementos no alfabéticos

Consiste en la remoción de los siguientes caracteres: '!', '#', '\$', '%', '&', '(', ')', '*', '+', ',', '-', '.', ':', ';', '<', '=', '>', '?', '@', '[', '\\', ']', '^', '{', '|', '}', '~', '1', '2', '3', '4', '5', '6', '7', '8', '9'.

3.4 Remoción de palabras vacías

Las palabras vacías son aquellas que tienen una alta frecuencia en un texto y generalmente tienen un menor contenido léxico, por ello, no nos son de gran ayuda en el análisis de textos. En español las palabras vacías son los artículos, pronombres, preposiciones, etc. Por ejemplo: un, sobre, todo, otro, ser, estamos, atrás, pero, fue, fin, va, el, la, ultimo, dos, aquel, etc.

3.5 Stemming

Es el proceso de llevar una palabra a su palabra base o raíz, eliminando sufijos y prefijos. Algunos ejemplos de palabras raíz son: cantar, jugar, viajar, comer, entre otros.

3.6 Construcción del vocabulario.

El vocabulario puede ser definido como el conjunto de palabras de un idioma [8], para este trabajo, el vocabulario es el conjunto de palabras obtenidas del corpus a partir del preprocesamiento resultando así en un vocabulario de 4190 palabras que serán un punto de referencia para la representación vectorial de los textos.

3.7 TF-IDF (term frequency – inverse document frequency)

Es una representación vectorial de los documentos, que representa que tan importante es una palabra para una colección de documentos cuyo cálculo es el producto de dos estadísticos TF (Term Frequency) e IDF (Inverse Document Frequency). El primer término es el número de veces que una palabra se encuentra en un documento mientras que el segundo término es que tanto una palabra es común o rara a lo largo de los documentos [9]. El TF de una palabra es baja si se encuentra pocas veces en un documento, asimismo si es alto se encuentra muchas veces en un documento. Por otra parte, el IDF de una palabra es bajo si se encuentra en muchos documentos, y es alto si se encuentra pocas veces. Para obtener los valores TF-IDF del vocabulario, se emplean las siguientes fórmulas [9]:

$$TF(x_{ij}) = x_{ij} * (1/\log_{10}(N(t_i))) \quad (1)$$

$$IDF(x_{ij}) = \log_{10}(M/(1 + DF(s_j))) \quad (2)$$

$$TF_IDF(x_{ij}) = TF(x_{ij}) * IDF(x_{ij}) \quad (3)$$

En donde:

$TF(x_{ij})$ es el número de ocurrencias del término i en el documento j

$N(t_i)$ es el número de palabras en el documento t_i

M es el número de documentos en una colección (corpus)

$DF(s_j)$ indica el número de documentos que contienen i

Con ayuda de estas fórmulas (1) (2) (3) y el vocabulario formado se realizó una representación vectorial de los textos para llevarlos a la siguiente etapa de clasificación.

3.8 Clasificación

La tarea de clasificación consiste en la acción y efecto de relacionar los clasificados en una determinada prueba según la RAE [6]. Esta actividad es realizada por una red neuronal, la cual es un modelo de aprendizaje automático que toma decisiones de manera similar al cerebro humano imitando el proceso de funcionamiento de las neuronas biológicas; por ende, análogamente a la neurona biológica constan de una capa de entrada, capas ocultas y una capa de salida; cada neurona se comunica y conecta con otras mediante una ponderación y umbral asociados.

La red neuronal usada para este método es una secuencial de 50 capas de densidad internas implementada con la librería de Keras. Esta red, fue entrenada con el 80% de los textos que conforman el corpus, tomando así un total de 40 textos por autor y se realizaron pruebas con el 20% de textos restantes para analizar la efectividad en la identificación de tendencias suicidas presentes en los textos, siendo esta etapa la que finalmente nos proporciona los resultados que se muestran en la figura 1.

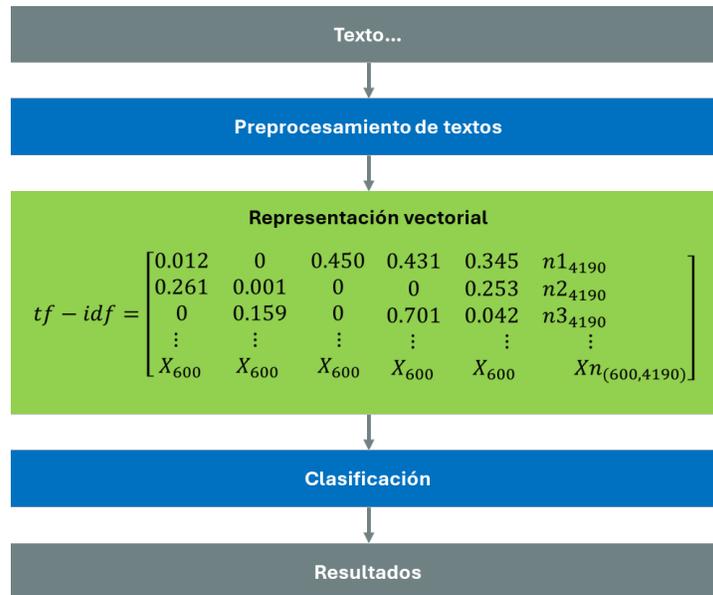


Figura 1. Diagrama del método propuesto

4 Resultados

Para comprobar el método propuesto, se recopiló un *corpus* de poemas de 12 autores distintos, de los cuales 6 cometieron suicidio y 6 murieron de causas naturales. El *corpus* consta de 600 poemas, con 50 poemas por autor. Se realizó una clasificación utilizando la red neuronal planteada en el método. Los resultados obtenidos se compararon con estudios previos realizados con *corpus* similares, obteniendo mejores resultados, como se muestra en la Tabla 1, en donde se aprecia la efectividad de métodos como Naive Bayes, SVM y la red neuronal secuencial [7].

Indudablemente, la efectividad que se obtuvo de la red neuronal secuencial fue mayor en comparación con los otros dos métodos empleando el mismo *corpus*, logrando una clasificación correcta del 94% de los datos de prueba como se puede observar en la figura 2.

5 Conclusiones y trabajo a futuro

Hoy en día los temas relacionados con la salud mental continúan siendo temas tabú y especialmente el suicidio entre ciertos sectores de la sociedad, a pesar de que se trata de concientizar a la población es aún difícil para personas que sufren una alteración en su estado de bienestar personal pedir ayuda a su comunidad cercana. Por ello, el uso de modelos computacionales como las redes

neuronales para el análisis de tendencias suicidas puede ser beneficioso para ayudar a este sector de la población, siendo un método preventivo interesante. Indiscutiblemente, el método propuesto para el corpus de poemas y la identificación de tendencias suicidas otorgo resultados buenos en la clasificación, mejorando la efectividad en un 20% y 30% frente a técnicas como Naive Bayes y SVM respectivamente. Como trabajo a futuro se pretende crear un *corpus* de tuits en español y experimentar con el mismo, así como, utilizar nuevos modelos como BERT y Roberta.

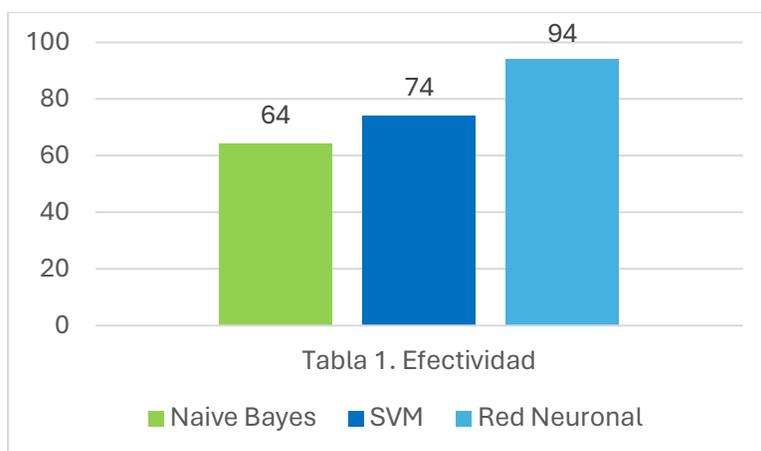


Figura 2. Comparación de efectividad en clasificación de diferentes métodos aplicados al *corpus* de poemas en español.

Referencias

- [1] World Health Organization: WHO. (2019, 19 diciembre). Mental health. https://www.who.int/health-topics/mental-health#tab=tab_1
- [2] World Health Organization: WHO. (2023, 28 agosto). Suicide. <https://www.who.int/news-room/fact-sheets/detail/suicide>
- [3] Coppersmith G, Leary R, Crutchley P, Fine A. Procesamiento del lenguaje natural de las redes sociales como detección del riesgo de suicidio. *Perspectivas de la informática biomédica*. 2018;10. doi: 10.1177/1178222618792860
- [4] Fuchiang R Tsui, Lingyun Shi, Victor Ruiz, Neal D Ryan, Candice Biernesser, Satish Iyengar, Colin G Walsh, David A Brent, Natural language processing and machine learning of electronic health records for prediction of first-time suicide attempts, *JAMIA Open*, Volume 4, Issue 1, January 2021, oaab011, <https://doi.org/10.1093/jamiaopen/oaab011>
- [5] Waheed A, Imran S. B, Shabana R. Design and Implementation of a Machine Learning-Based Authorship Identification Model. *Volumen 2019*. doi.org/10.1155/2019/9431073.

- [6] Real Academia Española (Consultado: 09 de junio 2024) RAE. <https://dle.rae.es/vocabulario>
- [7] Sánchez, J. Detecta Tendencias Suicidas Utilizando Procesamiento de Lenguaje Natural. [Tesis para obtener grado de Licenciatura]. Benemérita Universidad Autónoma de Puebla.
- [8] Real Academia Española (Consultado: 09 de junio 2024) RAE. <https://dle.rae.es/clasificaci%C3%B3n>
- [9] Zhuo Z. et al. TFIDF-FL: Localizing Faults Using Term Frequency-Inverse Document Frequency and Deep Learning. DOI: 10.1587/transinf.2018EDL8237
- [10] Saldarriaga Gómez, P. (2021). *Algoritmo para la detección del suicidio mediante el análisis de datos en las redes*. Tecnológico de Antioquia, Institución Universitaria.
- [11] Segura, C. C. Impacto del Leet Speaking en la detección de tendencias suicidas en tweets. [Tesis para obtener el grado de maestría en ciencias de la computación]. Instituto Politécnico Nacional.
- [12] Berrios Butron, J. et al. (2023). *Detección de ideas suicidas en Twitter*. Vol. 4, No.2, Febrero 2023. doi.10.48169/innosoft.s12.a97.
- [13] Shanskar, R. et al. (2023). *Analyzing Socia Media texts for Suicidal Risk Identification using Natural Language Processing*. 9th International Conference on Signal Processing and Communication. IEEE.
- [14] Sapna, K. et al. (2023). *Deep learning and natural language processiong-based model for the prediction of suicidal Ideation in Military Personnel*.

Dispositivos reconfigurables de interconexión en redes en chip

Reconfigurable on-chip network interconnect devices

Laura García Luciano¹, Remberto Sandoval Aréchiga², Guillermo González Martínez¹, Luis Octavio Solís Sánchez¹, Juvenal Villanueva Maldonado², Salvador Ibarra Delgado², José Ricardo Gómez Rodríguez², Viktor Iván Rodríguez Abdalá², Oscar Osvaldo Ordaz García².

¹Posgrado en Ingeniería y Tecnología Aplicada (PITec),
Universidad Autónoma de Zacatecas,
Av. Ramón López Velarde, 801, Col. Centro, Zacatecas 98000, México.

²Posgrado en Ingeniería para la Innovación Tecnológica (PIIT),
Universidad Autónoma de Zacatecas,
Campus UAZ Siglo XXI. Carr. Zacatecas-Guadalajara Km. 6,
Ejido "La Escondida", C.P. 98160 Zacatecas, Zac., México.

lgl.hardware@uaz.edu.mx, rsandoval@uaz.edu.mx, ing.memogm@uaz.edu.mx
lsolis@uaz.edu.mx, juvenal.villanueva@uaz.edu.mx, sibarra@uaz.edu.mx,
jrgrodri@uaz.edu.mx, abdala@uaz.edu.mx, oscarordazg@uaz.edu.mx

Abstract

The increase in the number of interconnect components in SoC presents the challenge of efficient communication, scalability, reliability, fault tolerance, power consumption and reconfigurability. This article identifies the types of reconfigurable interconnects used and describes the importance of reconfigurability in them. It is of great importance to know the context of on-chip reconfigurable interconnects, and to know in which areas there is little research and thus have a breakthrough for the scientific community.

Resumen

El incremento en los componentes que se interconectan en los SoC, presenta el reto de tener una comunicación eficiente, de ser escalable, de fiabilidad, de tolerancia a fallos, de consumo de energía y además de que se puedan reconfigurar. En este artículo se identificaron los tipos de interconexiones reconfigurables usados y se describe la importancia de la reconfigurabilidad en ellos. Es de gran importancia saber el contexto de las interconexiones reconfigurables en chip, y conocer en qué áreas existe poca investigación y así tener un parteaguas para la comunidad científica.

Keywords and phrases: Interconexión en Chip, Reconfigurable, SoC.

1 Introducción

La arquitectura On-Chip ha incrementado sus componentes a lo largo de los años debido a la mejora continua en la tecnología de fabricación que permite tener más transistores en una misma área de un chip. Dicha arquitectura provee de soluciones que satisfacen las necesidades de sistemas embebidos de alto rendimiento como por ejemplo en la industria de telecomunicaciones [1]. Además, los sistemas en chip albergan varios núcleos de procesamiento que realizan distintas funciones y que tienen la necesidad de comunicarse entre sí. Existen sistemas establecidos que interconectan los núcleos dentro del chip y que dependiendo de las restricciones del sistema en chip se elige alguno o una combinación de ellos.

Los sistemas de interconexión que nos encontramos en distintos trabajos son: la NoC, de conmutación (Crossbar) y el bus de datos. Cuando se aumenta el número de elementos a interconectar, la fiabilidad de las señales, el rendimiento y el consumo de energía hace que el diseño de los sistemas de interconexión se vuelva un tema crítico. En general, la comunicación entre núcleos busca ser escalable sin tener que configurar la arquitectura de interconexión y también la adaptabilidad de la interconexión para mejorar el rendimiento de algoritmos concurrentes. En una gran mayoría de artículos consultados implementan arquitecturas de redes en chip en sus distintas topologías o con una combinación de ellas por proveer escalabilidad a los sistemas en chip complejos. [2]

La reconfigurabilidad en los sistemas de interconexión en chip permite adaptaciones porque su entorno así lo requiere, que si bien, implica una mayor complejidad en diseño, los beneficios de tener un sólo sistema para distintos requerimientos hacen que sea flexible que optimizan en rendimiento, la eficiencia energética, etc.

En este trabajo se identifican a las distintas interconexiones reconfigurables en chip y se clasifican según su elemento reconfigurable y según la métrica en la que se beneficia con la arquitectura de comunicación usada. Con ello se quiere tener un panorama de las interconexiones reconfigurables en chip y las mejoras que se han obtenido de ellas.

2 Marco teórico y estado del arte

A continuación, se definen conceptos importantes para la comprensión de este artículo.

2.1 Básicos en Interconexiones en Chip

Los sistemas en chip pueden ser de tipo homogéneos o heterogéneos, dónde el heterogéneo se constituye de diferentes tipos de elementos de procesamiento (PE) y de núcleos de propiedad intelectual (IP core) que se comunican unos con otros [3].

A medida que aumenta el número de elementos en los sistemas en chip, la comunicación entre ellos se vuelve más compleja. Por tanto, los mecanismos de comunicación son clave en el rendimiento del sistema en chip (ver Figura 1).

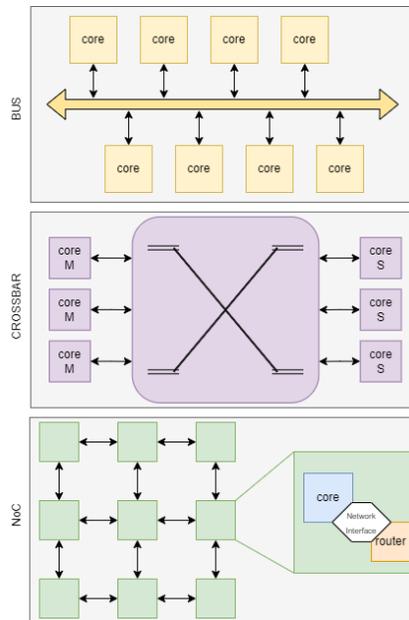


Figura 1. Arquitecturas básicas de interconexión en chip.

2.1.1 Elementos de interconexión

- **Enrutador:** El enrutador es un elemento que se compone de registros, switches, lógica de control y que se encarga de implementar el enrutamiento y el control de flujo requerido para almacenar y reenviar datos a sus destinos. El arbitraje es un tema crítico para el rendimiento de éste.
- **Bus:** La arquitectura de interconexión más común, donde todas las unidades de comunicación comparten el mismo medio de transmisión. Cada elemento interconectado agrega una capacitancia parásita que eventualmente incrementará el tiempo de llegada de los datos a su destino. Esto último implica que su escalabilidad sea limitada.
- **Red en chip:** En esta arquitectura de interconexión se conectan varios elementos donde la combinación de enrutadores, interfaces de red y enlaces, determinarán la topología con la cual se intercambiarán datos entre PE's o IP's. Una red en chip cuenta con un alto grado de modularidad y flexibilidad [4]; y es el algoritmo de enrutamiento el que se encarga de encaminar los datos desde el origen hasta el destino a través de los enrutadores.

- Enlaces dedicados: Un enlace dedicado se crea a través de una conexión directa entre dos componentes sin intermediarios, por lo que se puede utilizar en pequeños sistemas donde la escalabilidad y la flexibilidad no sea un factor necesario.
- Crossbar: Esta arquitectura interconecta cualquier entrada con cualquier salida, permitiendo que se realicen transmisiones de forma paralela sin interrupción, lo que la hace flexible [5].

2.1.2 Métricas en interconexiones en chip

El espacio de diseño de un SoC crece por la variedad de recursos disponibles de hardware y software. Por ello, el diseño se torna complejo cuando se piensa en definir una arquitectura eficiente y flexible para los sistemas actuales y futuros. Para definir una arquitectura adecuada para una aplicación, es necesario un análisis exhaustivo de la aplicación. En este artículo, definimos varias métricas que permiten analizar las características del sistema a un alto nivel de abstracción [6]. Las métricas permiten evaluar el impacto en el rendimiento del sistema, de modo que se consideran durante los primeros pasos del flujo de diseño antes de cualquier definición de arquitectura de hardware y software.

- Potencia: Consumo de energía necesaria durante la operación.
- Velocidad: Rapidez con la que se pueden transferir datos a través de los diferentes componentes del chip.
- Ancho de banda: Indica cuantos bits de datos pueden ser transmitidos por segundo a través de las interconexiones en chip.
- Escalabilidad: Capacidad del sistema de mantener el rendimiento al aumentar el número de elementos en la red.
- Latencia y Throughput: La latencia y el throughput son métricas de rendimiento, la primera indica el tiempo que le toma a un paquete atravesar la red y el segundo es el número de bits por segundo que la red puede transportar desde la entrada a la salida de manera exitosa.
- Tolerancia a fallos: Se refiere a la robustez de la red y a los mecanismos de detección y corrección de errores, manteniendo el funcionamiento y rendimiento de la red cuando ocurren fallos.
- Área de diseño: Espacio físico que ocupan tanto los componentes como las interconexiones en el chip. Es importante minimizar el área para reducir costos y mejorar la eficiencia del diseño en chip.

2.2 Reconfigurabilidad en chip

La reconfigurabilidad puede hacerse en tiempo de síntesis y en tiempo de ejecución [7]. Cuando se hace en tiempo de síntesis (Reconfiguración estática), sólo se realiza una vez en la etapa de diseño cuando los requerimientos de comunicación son bien conocidos y no varían en el tiempo. Mientras que la reconfiguración en tiempo de ejecución comprende la siguiente clasificación:

- Dinámica: Una configuración dinámica se realiza en tiempo de ejecución por que los requerimientos se conocen solamente en tiempo de ejecución [8].

- **Dinámica parcial:** La configuración se realiza únicamente en una parte del sistema, de modo que no debe afectar la funcionalidad del resto.

Por ejemplo, un adaptador de red contiene los protocolos que pueden ser configurados para adaptar los elementos de una red (el tamaño de los paquetes, el tipo de control de flujo, QoS y las técnicas de multiplexado). Así mismo, se puede conocer el número de paquetes transmitidos y su estado final. Estas configuraciones pueden hacerse de manera dinámica o estática, según los requerimientos del sistema.

2.2.1 Mecanismos de reconfigurabilidad

Los mecanismos de reconfigurabilidad nos permiten saber en qué métrica se está enfocando el sistema de interconexión en chip. Por ejemplo, en una red se definen técnicas de enrutamiento y topología, por tanto, se podrá configurar desde las políticas de arbitraje, el algoritmo de enrutamiento y hasta la técnica de conmutación.

Otro ejemplo, en una NoC reconfigurable se pueden agregar o quitar enrutadores, interfaces de red o enlaces por reconfiguración dinámica parcial, o utilizando switches para cambiar la topología de red, cambiar dinámicamente el tamaño del buffer y ajustar el ancho de bit del enlace siempre de acuerdo con los requerimientos del sistema [9].

En este trabajo se identifican los siguientes mecanismos de reconfigurabilidad:

- **Topología:** La topología es el primer aspecto fundamental del diseño de NoC y tiene un profundo efecto en el costo y el rendimiento general de la red. La topología determina el diseño físico y las conexiones entre nodos y canales. La topología influye significativamente en la latencia y el consumo de energía. Además, dado que la topología determina el número de caminos alternativos entre nodos, afecta la distribución del tráfico de la red y, por lo tanto, el ancho de banda y el rendimiento de la red logrados.
- **Ancho de banda:** Un componente crítico de los procesadores de alto rendimiento como los GPU es la red en chip (NoC) que interconecta los núcleos y las particiones de memoria. Se han propuesto diferentes arquitecturas de NoC para procesadores de rendimiento, pero a menudo se han basado en principios similares a los NoC multinúcleo (o CPU), incluido el énfasis en el ancho de banda de bisección y el patrón de tráfico [10].
- **Arbitraje:** En un sistema de interconexión en chip es necesario que el acceso a los recursos compartidos sea asignado por un árbitro [6]. El requerimiento en un arbitraje es la equidad y según los diversos criterios y necesidades se escoge entre varias políticas de arbitraje. Los elementos que pidan un recurso van a tener cierta probabilidad de que les sea asignado, evitando así la situación llamada “inanición” [11]. Y con respecto al QoS, es necesario que la justa asignación de recursos concuerde con los requerimientos del sistema de red para que se pueda hablar de una política de servicio [12].
- **Enrutamiento:** Se refiere a la transferencia de datos entre los diferentes componentes dentro del chip. Es un aspecto crucial del diseño de NoC, ya que afecta significativamente el rendimiento, la latencia, el consumo de energía y la fiabilidad del sistema.

- **Tamaño de buffer:** El enfoque de cambiar el tamaño del buffer resulta de querer reducir el área de un enrutador y/o mejorar el rendimiento de una arquitectura de interconexión, así como también, poder disminuir el bloqueo de paquetes [13]. En este último, existen buffers flotantes que son asignados a los buffers fijos cuando están congestionados.
Existe un enfoque no considerado en que se pudiera escribir de manera simultánea en un buffer flotante y también a su buffer fijo asignado, desde dos puertos de entrada distintos, o a cualquier buffer mejorando el rendimiento de la red con el costo de más lógica de control.
- **Calendarización:** Gestión y programación de las tareas o flujos de datos que se transmiten a través del sistema. En una red en chip, varios núcleos o módulos están interconectados y pueden comunicarse entre sí para ejecutar tareas y procesar datos. Por ejemplo, en [14] un programador de tareas asigna dinámicamente las tareas a los núcleos de una NoC en función de los requisitos de tiempo real y consideraciones de eficiencia energética.

3 Reconfigurabilidad en sistemas de interconexión en chip. Criterios y necesidades.

La reconfigurabilidad en interconexiones SoC ha permitido una mayor flexibilidad, eficiencia y capacidad de adaptación en el diseño de circuitos integrados, impulsando avances significativos en diversas aplicaciones. Por lo que la necesidad de reconfigurar las interconexiones en chip se ha vuelto cada vez más importante en el diseño de sistemas modernos.

Existen aplicaciones que demandan una creciente adaptabilidad y la optimización de recursos, especialmente en aplicaciones complejas como redes definidas por software, procesamiento de inteligencia artificial, entre otras. Según estudios, las arquitecturas reconfigurables permiten ajustar dinámicamente las interconexiones internas del chip para optimizar el rendimiento y minimizar el consumo energético. Un ejemplo de esto es la investigación sobre una herramienta que diseña NoCs dinámicamente reconfigurables [15], donde a partir de un rendimiento necesario mínimo y la carga del tráfico esperado nos da como resultado un diseño con una configuración que satisface dichos requerimientos en un área mínima.

La reconfigurabilidad en hardware tiene múltiples aplicaciones y varios requerimientos de rendimiento, y esta flexibilidad no es posible en las arquitecturas de interconexión fijas. En [16] se desarrolló una CNN basado en una NoC, es decir, el diseño busca adaptarse dinámicamente a las distintas CNN mejorando su rendimiento en términos de velocidad y consumo de recursos sin necesitar de más diseños para cada modelo de red neuronal, por lo que soporta una gran variedad de aplicaciones.

Existen dispositivos que diseñados de cierta forma, determinan el elemento de la interconexión que permita varias configuraciones y diversos autores la escogen dependiendo de la aplicación. Por ejemplo, en [16], el elemento reconfigurable es el enrutador que ajusta la comunicación entre los nodos de procesamiento, cambiando las rutas y la distribución de los datos según los requisitos de la red neuronal en tiempo real. Así, se optimiza el uso de recursos y mejora la eficiencia del sistema sin necesidad de un hardware fijo.

En [17] nos muestran métodos en arquitecturas reconfigurables donde crean una configuración óptima de interconexiones que pueden adaptarse dinámicamente a las cargas de trabajo y condiciones del sistema y en [18] se verifica la factibilidad del sistema dependiendo de distintos escenarios de reconfigurabilidad en varios procesadores.

En ocasiones es necesario decidir que métrica se puede sacrificar para cumplir con los requerimientos del sistema, por ejemplo en una NoC reconfigurable [2] se mejora el rendimiento del SoC a expensas de la potencia y de hardware adicional, para diferentes aplicaciones SoC. Una de las características principales de esta NoC es que cada IP core tiene un acceso exclusivo a un puerto de un enrutador a través de uno o más switches reconfigurables. Dentro de las métricas que mejoran en comparación con NoCs convencionales está la reducción significativa de latencia de 27% para la aplicación MPEG-4 en la red reconfigurable AHRNoC1, lo que significa que el tiempo en que el sistema tarda en procesar y transferir datos es más corto, y ocupa menos espacio en el chip consumiendo un 0,7% menos de área y consume 11% menos de potencia y puede ejecutarse 3 veces más rápido que las NoC tradicionales también.

Como en el ejemplo anterior, enfocarse en la aplicación es efectivo para cerrar la brecha entre las métricas de potencia y rendimiento de la NoC actual y lo ideal para los SoC multinúcleo de aplicación específica[19].

Las configuraciones son diversas cuando se diseñan las arquitecturas de interconexión reconfigurable, por ejemplo [20] modifica el sentido de los enlaces que conectan los nodos de una red mesh haciéndolos bidireccionales, aprovechando ancho de banda y además tiene rutas más cortas entre nodos remotos. Y también las configuraciones nacen de necesidades específicas, como en [13] donde se estudia una técnica de redimensionamiento dinámico del búfer, creándose la necesidad de enrutadores dinámicos.

Es importante mencionar que en los últimos años se han desarrollado redes inalámbricas que responden a distintas necesidades de las NoCs convencionales, donde la finalidad de la reconfigurabilidad sigue siendo la misma. Por ejemplo, en [21], se enfoca en la implementación de OWiNoCs donde se combinan las ventajas de las comunicaciones ópticas e inalámbricas para mejorar la eficiencia de las redes en chip. En este contexto, las OWiNoCs cuentan con una menor latencia por la naturaleza de la transmisión de luz, pueden alcanzar los terabits densidades de ancho de banda en el orden de los terabits por segundo, escalan mejor en entornos densos sin cuellos de botella y pueden tener mejor consumo de energía debido a que la señal óptica no se degrada tanto como la eléctrica. Y de igual manera, el dispositivo reconfigurable es un enrutador que ajusta la dirección del haz de radiación mediante el cambio de la señal de entrada a las antenas del OPA, permitiendo que la red se adapte a varias topologías y requisitos de comunicación.

En los últimos años, los trabajos nos introducen en el diseño de redes en chip reconfigurables. [22] nos presenta el modelo llamado StocNoC que trata sobre la simulación estocástica de interacciones en redes complejas, por lo que finalmente tiene la capacidad de configurar el sistema para simular diferentes topologías de red y dinámicas de propagación. Y también hacen uso de la combinación

de tecnologías, como en [23] que se centra en una red híbrida optoelectrónica para mejorar la comunicación en chips reconfigurables, que contiene un enrutador dinámico utilizado para configurar la comunicación entre los nodos de procesamiento del chip, adaptándose a diferentes condiciones de tráfico mejorando la eficiencia de la red. Entonces, independientemente de la tecnología usada, siempre existirá un elemento reconfigurable perteneciente a la interconexión en chip que permitirá cumplir los requerimientos de la aplicación.

Una manera de clasificar los trabajos en torno a la reconfigurabilidad en interconexiones en chip, se puede observar en la Tabla 1, donde se relacionan los elementos base de un sistema de interconexión con el criterio reconfigurable en el que se enfocó el trabajo. Por ejemplo, [24], [25] y [26] abordaron la reconfiguración en enrutamiento y topología, y utilizando NoCs donde el elemento base es un enrutador.

Se encontró que, utilizando las tres distintas interconexiones en chip, muchos trabajos se enfocan en una topología, buffer, árbitros, ancho de banda reconfigurables, pero en criterios como la calendarización y calidad de servicio hay mucho que desarrollar aún. Por ejemplo, en [24] se observa que activan y desactivan enrutadores en una Malla de 4x4 por medio de conmutadores para cambiar la topología y con ello el enrutamiento también debe ser adaptado según los enrutadores que se encuentren activos. Y en [20] vemos una Malla de 8x8 que a su vez es subdividida en 4 regiones de 4x4 que, para una mejor gestión del tráfico, utilizan “enrutadores hub” que facilitan el acceso a los enrutadores de cada región, canales bidireccionales que permiten ajustar dinámicamente la topología según el tráfico en tiempo de ejecución, así como también un algoritmo de enrutamiento que se adapta al tráfico de la red.

En cuanto a reconfigurabilidad de arbitraje, surge la necesidad cuando un esquema de arbitraje no satisface los requerimientos de varias aplicaciones que pudieran cambiar en tiempo de ejecución, y es entonces cuando con un árbitro reconfigurable se selecciona la política más apropiada para los diversos escenarios. Por ejemplo en [27], se cuenta con tres políticas de arbitraje a seleccionar: de prioridad fija, TDMA y Round-Robin; y en conjunto con su enfoque hardware/software, permite un arbitraje de bus reconfigurable en tiempo de ejecución para lograr los objetivos de rendimiento de las aplicaciones.

En la Tabla 1, observamos que los trabajos se concentran en NoCs, en tanto que el Bus y el Crossbar tienen bastante área de oportunidad en el ámbito de la reconfiguración en cualquiera de los criterios reconfigurables.

Por otro lado, (ver Tabla 2) las métricas en las cuales se enfocan los trabajos son: menor consumo de energía, menor área, menor latencia y un mayor Throughput, y el costo, refiriéndose este último al corto tiempo de simulación que implicaría reducir costos. Se tienen dos áreas de oportunidad, enfocarse también en el ancho de banda y en la calidad de servicio, así como también en la tolerancia a fallos de las NoCs eléctricas, porque ya se comentó que las redes inalámbricas revisadas cumplen con ello.

Tabla 1. Criterios reconfigurables vs elemento base.

| | | Criterio Reconfigurable | | | | | | |
|---------------|----------|--|-----------|----------------|---|------------------|-----------------|------|
| | | Topología | Arbitraje | Ancho de Banda | Enrutamiento | Tamaño de Buffer | Calendarización | QoS |
| Elemento base | Bus | | [27] | | [8] | | | |
| | Crossbar | | | | | | | |
| | Router | [16][15][17][18][19][21] [22][23][24][25][26][29] | [30] | | [4][8][10][18][19][20][22] 23][24][25][26][28][29] | [19] | | [14] |

Tabla 2. Métricas vs Sistema de interconexión.

| METRICS | Bandwidth | Through put | Energy consumption | Area | Latency | Fault tolerance | QoS | Costo |
|---------|-----------|---|---|-----------------------------|--|----------------------|-----|-------|
| NoC | [21] | [2][23][21] [4][5][13][15][16] [18][19][20][21] [22][23][24][25][26][28][29][30] | [14][16][17][18][19][23][24][25]][28] | [8][13][15][16][19][24][28] | [16][23][4][5][18][20][21] [22][23][24][28] | [19][21][23][25][26] | | [22] |
| Bus | | [27] | | | | | | |

Y finalmente en otros trabajos, se desarrolla de manera aislada un dispositivo reconfigurable que mientras no se encuentre en una arquitectura de comunicación, no se puede evaluar su desempeño en un sistema. Sin embargo, es importante mencionar dichos trabajos que dan pauta a los sistemas de interconexión reconfigurables. Aquellos trabajos que se enfocan, no en la arquitectura de interconexión, si no en un enrutador reconfigurable son [28] y [29] , mientras que [30] lo hace en un árbitro reconfigurable.

3.1 Desafíos de la reconfigurabilidad en interconexiones en chip

Existen algunas coincidencias entre las redes eléctricas y las redes inalámbricas de desafíos importantes a considerar como:

Complejidad del diseño: La creación de un sistema reconfigurable implica un diseño más complejo. Esto puede llevar a un mayor tiempo de desarrollo y a la necesidad de herramientas de diseño más sofisticadas.

Sobrecarga de recursos: La reconfigurabilidad puede requerir recursos adicionales, como memoria y capacidad de procesamiento, para gestionar la lógica de reconfiguración y el estado del sistema. Esto puede resultar en un mayor consumo de energía y en un uso ineficiente de los recursos disponibles.

Desafíos de fiabilidad: La reconfigurabilidad puede complicar la gestión de errores y la fiabilidad del sistema. Los cambios frecuentes en la configuración pueden aumentar la probabilidad de fallos o comportamientos inesperados, lo que puede ser problemático en aplicaciones críticas.

Integración de componentes: La integración de componentes ópticos, como los arreglos de fase óptica y los dispositivos de conversión electro-ópticos, puede ser tecnológicamente desafiante [21].

Requerimientos de energía: Aunque se espera que la reconfigurabilidad mejore la eficiencia energética, el proceso de ajuste constante de las configuraciones puede requerir un consumo adicional de energía, lo que podría contrarrestar algunos de los beneficios esperados.

Gestión de contención: La necesidad de gestionar la contención en la red, donde múltiples transmisores intentan enviar datos al mismo receptor, añade un nivel adicional de complejidad en el diseño del control lógico necesario para la red.

Y aunque existan retos en el tema de reconfigurabilidad, se debe tener en cuenta la relación costo-beneficio dependiendo de la aplicación y los requerimientos del sistema.

4 Arquitecturas para la reconfigurabilidad

Cuando ciertas características no fueron definidas en tiempo de diseño, es necesario poder cambiar la arquitectura del sistema para agregarlas. En este sentido, la adaptabilidad es una solución que permite elementos de control que manipulan parámetros tanto de hardware como de software para crear nuevas condiciones de ejecución con un impacto mínimo. Cuando se tiene una arquitectura reconfigurable, una infraestructura de monitoreo (observa el rendimiento, la eficiencia de energía y la tolerancia a fallos) y un proceso que decida la nueva configuración dependiendo del estado del sistema (centralizada o distribuida), se habla entonces de una arquitectura adaptable.

La observabilidad del sistema permite entender su estado interno basado únicamente en el conocimiento de sus señales externas. Por tanto, entre más observable sea un sistema, más rápido se encuentra la fuente de un problema sin pruebas adicionales. La observabilidad permite un mejor monitoreo, debido a que recoge información de forma continua describiéndola (describiendo la relación entre datos) y dando un contexto del entorno.

- REDNoC (Reconfigurable Dynamic Network-on-Chip): Es una arquitectura NoC diseñada para ser altamente reconfigurable. Permite la modificación dinámica de la topología y la configuración de la red para adaptarse a diferentes cargas de trabajo y condiciones del sistema. En [8] y en [31] se introdujeron las ideas clave de las redes dinámicamente reconfigurables.
- SDNoC (Software-Defined Network-on-Chip): Aplica los principios de las SDN a las NoC, permitiendo una gestión y control centralizados de la red a través del software. SDNoC facilita la reconfiguración dinámica y la optimización de la red según las necesidades de las aplicaciones y el sistema [32].

5 Conclusiones y trabajo a futuro

Es un hecho de que con el aumento de componentes en SoC, es necesario realizar diseños que permitan adaptarse a su entorno y que las métricas se vean lo menos afectadas, es decir, según los requerimientos del sistema se puede decidir el tipo de reconfigurabilidad y el criterio reconfigurable que tendrá el sistema de interconexión. Con este trabajo se observan que los sistemas de interconexión con posibilidades de desarrollo en el tema de reconfigurabilidad son Crossbar y Bus en todos los criterios reconfigurables, a su vez, las NoCs tienen huecos de desarrollo en el tema de

Calendarización y QoS. Y dada una arquitectura de interconexión reconfigurable, es posible que desde una capa abstracta como puede ser SDN, se pueda monitorear el sistema y hacer la reconfiguración en tiempo real donde las interconexiones deben ajustarse a los flujos de datos cambiantes, de ahí la importancia de tener dichas arquitecturas de interconexión y saber los campos que tienen área de oportunidad para su desarrollo.

Agradecimientos

Los autores de este artículo agradecen al Consejo Nacional de Humanidades, Ciencia y Tecnología (CONAHCYT) y al Posgrado en Ingeniería y Tecnología Aplicada de la Universidad Autónoma de Zacatecas por el apoyo otorgado en la realización de este trabajo.

Abreviaturas y Símbolos

SoC: System on Chip.

NoC: Network on Chip.

PE: Process Element.

IP: Intellectual Property.

QoS: Quality of Service.

GPU: Graphics Processing Unit.

CPU: Central Processing Unit.

REDNoC: Reconfigurable and Extensible Dynamic Network-on-Chip.

SDN: Software-Defined Network.

SDNoC: Software-Defined Network-on-Chip.

OWiNoCs: Optical Wireless Networks on-Chip.

OPA: Optical Phased Arrays.

Referencias

[1] A. S. Kumar and T. V. K. H. Rao, "Scalable benchmark synthesis for performance evaluation of NoC core mapping," *Microprocess Microsyst*, vol. 79, p. 103272, Nov. 2020, doi: 10.1016/j.micpro.2020.103272.

[2] M. Oveis-Gharan and G. N. Khan, "Reconfigurable on-chip interconnection networks for high performance embedded SoC design," *Journal of Systems Architecture*, vol. 106, no. March 2019, 2020, doi: 10.1016/j.sysarc.2020.101711.

[3] D. Göhringer, L. Meder, S. Werner, O. Oey, J. Becker, and M. Hübner, "Adaptive multiclient network-on-chip memory core: Hardware architecture, software abstraction layer, and application exploration," *International Journal of Reconfigurable Computing*, vol. 2012, 2012, doi: 10.1155/2012/298561.

- [4] C. Killian, C. Tanougast, F. Monteiro, and A. Dandache, "A new efficient and reliable dynamically reconfigurable network-on-chip," *Journal of Electrical and Computer Engineering*, vol. 2012, no. c, 2012, doi: 10.1155/2012/843239.
- [5] M. Jahanshahi and F. Bistouni, "Scalable Crossbar Network," pp. 137–162, 2018, doi: 10.1007/978-3-319-78473-1_5.
- [6] W. J. Dally and B. P. Towles, *Principles and practices of interconnection networks*. Elsevier, 2004. doi: 10.1109/ICSICT49897.2020.9278309.
- [7] D. Koch, *Partial reconfiguration on {FPGA}s : architectures, tools and applications*, vol. 153. 2013. [Online]. Available: <http://link.springer.com/10.1007/978-1-4614-1225-0%5Cnhttp://link.springer.com/book/10.1007/978-1-4614-1225-0/page/1>
- [8] C. Bobda and A. Ahmadinia, "Dynamic interconnection of reconfigurable modules on reconfigurable devices," *IEEE Design and Test of Computers*, vol. 22, no. 5, pp. 443–451, 2005, doi: 10.1109/MDT.2005.109.
- [9] J. P. Diguët, "Self-adaptive network on chips," *SBCCI 2014: Proceedings of the 27th Symposium on Integrated Circuits and Systems Design*, 2014, doi: 10.1145/2660540.2660992.
- [10] J. Kim, S. Cho, M. Rhu, A. Bakhoda, T. M. Aamodt, and J. Kim, "Bandwidth Bottleneck in Network-on-Chip for High-Throughput Processors," in *Proceedings of the ACM International Conference on Parallel Architectures and Compilation Techniques*, New York, NY, USA: ACM, Sep. 2020, pp. 157–158. doi: 10.1145/3410463.3414673.
- [11] S. Ibarra-Delgado, "Diseño de Sistemas Embebidos Reconfigurables Empleando Elementos de Interconexión en Chip Definidos por Software," Universidad de Córdoba, 2021.
- [12] W. J. Dally and B. P. Towles, *Principles and practices of interconnection networks*. Elsevier, 2004. doi: 10.1109/ICSICT49897.2020.9278309.
- [13] M. P. Vestias and H. C. Neto, "A dynamic buffer resize technique for networks-on-chip on FPGA," in *2011 VII Southern Conference on Programmable Logic (SPL)*, IEEE, Apr. 2011, pp. 227–232. doi: 10.1109/SPL.2011.5782653.
- [14] A. Gammoudi, A. BenZina, M. Khalgui, and D. Chillet, "Energy-Efficient Scheduling of Real-Time Tasks in Reconfigurable Homogeneous Multicore Platforms," *IEEE Trans Syst Man Cybern Syst*, vol. 50, no. 12, pp. 5092–5105, Dec. 2020, doi: 10.1109/TSMC.2018.2865965.
- [15] R. Ahmed, H. Mostafa, y A. H. Khalil, "Design of a reconfigurable network-on-chip for next

generation FPGAs using Dynamic Partial Reconfiguration”, *Microelectronics J.*, vol. 108, p. 104964, feb. 2021, doi: 10.1016/j.mejo.2020.104964.

[16] A. Firuzan, M. Modarressi, M. Reshadi, y A. Khademzadeh, “Reconfigurable Network-on-Chip based Convolutional Neural Network Accelerator”, *J. Syst. Archit.*, vol. 129, p. 102567, A. Farhadi Beldachi, M. Hosseinabady, and J. L. Nunez-Yanez, “Configurable Router Design for Dynamically Reconfigurable Systems based on the SoCWire NoC,” *International Journal of Reconfigurable and Embedded Systems (IJRES)*, vol. 2, no. 1, p. 4864, 2013, doi: 10.11591/ijres.v2i1.2144.

[17] J. Huang, X. Xu, N. Wang, and S. Chen, “Reconfigurable topology synthesis for application-specific NoC on partially dynamically reconfigurable systems,” *Integration*, vol. 65, pp. 331–343, Mar. 2019, doi: 10.1016/j.vlsi.2018.02.012.

[18] I. Khemaissia, O. Mosbahi, M. Khalgui, and Z. Li, “New Methodology for Feasible Reconfigurable Real-Time Network-on-Chip NoC,” in *Proceedings of the 11th International Joint Conference on Software Technologies*, SCITEPRESS - Science and Technology Publications, 2016, pp. 249–257. doi: 10.5220/0005992002490257.

[19] M. VenkataTheertha, R. Nandi, and B. V. S. L. Bharathi, “Topology Re-Configuration for On-Chip Networks with Back-Tracking,” *Int J Comput Appl*, vol. 79, no. 17, pp. 40–45, 2013, doi: 10.5120/13965-1939.

[20] Z. Qian, P. Bogdan, and G. Wei, “A traffic-aware adaptive routing[] algorithm on a highly reconfigurable network-on-chip architecture,” *Proceedings of the eighth ...*, pp. 161–170, 2012, doi: 10.1145/2380445.2380475.

[21] G. Calo et al., “Reconfigurable Optical Wireless Switches for On-Chip Interconnection”, *IEEE J. Quantum Electron.*, vol. 59, núm. 3, pp. 1–10, 2023, doi: 10.1109/JQE.2022.3224778.

[22] A. Zhanbolatov, K. Vipin, A. Dadlani, y D. Fedorov, “StocNoC: Accelerating Stochastic Models Through Reconfigurable Network on Chip Architectures”, *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12083 LNCS, núm. March, pp. 361–375, 2020, doi: 10.1007/978-3-030-44534-8_27.

[23] B. Yang, Y. Li, C. Xi, R. Shan, Y. Feng, y J. Luo, “Design and Implementation of Reconfigurable Array Adaptive Optoelectronic Hybrid Interconnect Shunting Network”, *Electron.*, vol. 13, núm. 9, 2024, doi: 10.3390/electronics13091668.

- [24] A. Ramy, H. Mostafa, and A. H. Khalil, "Design of a reconfigurable network-on-chip for next generation FPGAs using Dynamic Partial Reconfiguration," *Microelectronics J*, vol. 108, no. January, p. 104964, 2021, doi: 10.1016/j.mejo.2020.104964.
- [25] Y. Chen, K. Ren, and N. Gu, "Router-shared-pair mesh : a reconfigurable fault-tolerant network-on-chip architecture," vol. 10, no. 6, pp. 526–536, 2018.
- [26] Y. Ren, L. Liu, S. Yin, J. Han, Q. Wu, and S. Wei, "A fault tolerant NoC architecture using quad-spare mesh topology and dynamic reconfiguration," *Journal of Systems Architecture*, vol. 59, no. 7, pp. 482–491, 2013, doi: 10.1016/j.sysarc.2013.03.010.
- [27] É. Sousa, D. Gangadharan, F. Hannig, y J. Teich, "Runtime reconfigurable bus arbitration for concurrent applications on heterogeneous MPSoC architectures", en *Proceedings - 2014 17th Euromicro Conference on Digital System Design, DSD 2014*, oct. 2014, pp. 74–81, doi: 10.1109/DSD.2014.105.
- [28] A. Farhadi Beldachi, M. Hosseinabady, and J. L. Nunez-Yanez, "Configurable Router Design for Dynamically Reconfigurable Systems based on the SoCWire NoC," *International Journal of Reconfigurable and Embedded Systems (IJRES)*, vol. 2, no. 1, p. 4864, 2013, doi: 10.11591/ijres.v2i1.2144.
- [29] J. E. Rosales, F. Tobajas, V. de Armas, J. A. Mori, and R. Sarmiento, "Dynamically reconfigurable router for NoC congestion reduction," *VLSI Circuits and Systems V*, vol. 8067, p. 80670H, 2011, doi: 10.1117/12.887473.
- [30] Ching-Chien Yuan, Yu-Jung Huang, Shih-Jhe Lin, and Kai-hsiang Huang, "A reconfigurable arbiter for SOC applications," in *APCCAS 2008 - 2008 IEEE Asia Pacific Conference on Circuits and Systems*, IEEE, Nov. 2008, pp. 713–716. doi: 10.1109/APCCAS.2008.4746123.
- [31] T. Pionteck, C. Albrecht, y R. Koch, "A dynamically reconfigurable packet-switched network-on-chip", en *Proceedings -Design, Automation and Test in Europe, DATE*, 2006, vol. 1, p. 8 pp., doi: 10.1109/date.2006.244035.
- [32] J.R. Gomez-Rodríguez, R.Sandoval-Arechiga,S. Ibarra-Delgado,V.I. Rodriguez-Abdala,J.L. Vazquez-Avila,and R.Parra-Michel, "A Survey of Software-Defined Networks-on-chip: Motivations,Challenges,and Oportunities",*Micromachines(Basel)*,pp.1-27,2021.
- Haga clic aquí para escribir texto.

Inteligencia artificial para la detección temprana de trastornos alimenticios en datos textuales

Artificial intelligence for early detection of eating disorders in textual data

Evelin del Carmen Ruiz Preza, María del Carmen Santiago Díaz, Judith Pérez Marcial, Abraham Sánchez López, Ana Claudia Zenteno Vázquez, Gustavo Trinidad Rubín Linares

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.

evelin.ruizp@alumno.buap.mx, {marycarmen.santiago, judith.perez, abraham.sanchez, ana.zenteno, gustavo.rubin}@correo.buap.mx

Resumen.

Los trastornos alimentarios, como la anorexia y la bulimia nerviosas, se caracterizan por alteraciones en la alimentación y preocupación por los alimentos, así como por problemas notables de peso corporal y forma. Los síntomas o comportamientos dan lugar a un riesgo o daño considerables para la salud, una angustia notable o una discapacidad funcional importante. En este trabajo presentamos un modelo de IA aplicando NEA para el diagnóstico de trastornos alimenticios que apoya a la toma de decisiones del especialista de la salud e intervención oportuna.

Abstract.

Eating disorders, such as anorexia and bulimia nervosa, are characterized by eating disturbances and preoccupation with food, as well as notable problems with body weight and shape. Symptoms or behaviors result in significant health risk or harm, significant distress, or significant functional disability. In this work, we present an AI model applying NEA for the diagnosis of eating disorders to support health care specialist decision-making and timely intervention.

Keywords and phrases: Trastornos Alimenticios, Inteligencia Artificial, National Eating Assessment, NEA.

1 Introducción

Los trastornos alimenticios son enfermedades graves que afectan tanto la salud física como mental de quienes las padecen. Entre los más comunes se encuentran los trastornos de la anorexia, la bulimia nerviosa y el trastorno por atracón, los cuales se caracterizan por una preocupación obsesiva por el peso, la comida y la

aparición, lo que lleva a comportamientos alimentarios muy dañinos. De acuerdo con datos de la Organización Mundial de la Salud (OMS), estos trastornos afectan aproximadamente al 9% de la población mundial, siendo más prevalentes en mujeres jóvenes, aunque también se ha visto un aumento significativo en hombres y en personas de todas las edades. La anorexia nerviosa tiene una tasa de mortalidad del 10%, una de las más altas entre los trastornos mentales. Es crucial una intervención temprana para prevenir complicaciones a largo plazo.[1]

1.1 Trastornos alimenticios

Los trastornos alimenticios son enfermedades médicas graves con una influencia biológica que se caracterizan por alteraciones severas de conductas alimentarias. Estas enfermedades son progresivas y se manifiestan a través de la conducta alimentaria y consisten en una serie de síntomas, entre los que predomina una variación de la autoimagen corporal. En la tabla 1 se muestra una estadística sobre el impacto de los padecimientos por sexo. La falta de tratamiento adecuado puede llevar a la cronificación del trastorno y a complicaciones severas que impactan en su calidad de vida de quien lo padece. [2]

Tabla 1. Prevalencia de los trastornos de la conducta alimentaria en adolescentes. [8]

| | Mujeres % (n) | Hombres % (n) | Totales %(n) |
|--------------------------|---------------|---------------|--------------|
| Anorexia | 0,36% (2) | 0% (0) | 0,19% (2) |
| Bulimia | 1,07% (6) | 0% (0) | 0,57% (6) |
| TCANE | 4% (21) | 2,55% (8) | 3,34% (26) |
| Prevalencia total | 5,47% (29) | 2,55% (8) | 3,8% (34) |
| Prevalencia riesgo EAT + | 33% (243) | 20,62 (125) | 27,42% (368) |

1.1.1 Anorexia nerviosa

La anorexia nerviosa se caracteriza por una restricción extrema de la ingesta de alimentos, un miedo intenso a ganar peso y una distorsión de la imagen corporal. Las personas con anorexia nerviosa se ven a sí mismas con sobrepeso, aunque estén peligrosamente delgadas, en donde entre los signos físicos se incluyen una pérdida significativa de peso, atrofia mamaria y en los hombres disminución de los niveles de testosterona con problemas testiculares en hombres. Los niveles de ciertas hormonas, como las gonadotrofinas, disminuyen, mientras que los niveles de cortisol y de la hormona del crecimiento aumentan, estas alteraciones hormonales pueden llevar a problemas reproductivos, como la ausencia de menstruación en mujeres y la disminución del deseo sexual en hombres. Además, el aumento del cortisol, una hormona del estrés, puede provocar una serie de problemas de salud, como debilidad muscular y problemas del sistema inmunológico. De acuerdo con estadísticas a nivel mundial, se estima que el 0.9% de las mujeres y el 0.3% de los hombres desarrollan anorexia nerviosa en algún momento de sus vidas. [4]. Figura 1.

1.1.2 Bulimia nerviosa

La bulimia nerviosa se caracteriza por episodios de atracones seguidos de conductas compensatorias, como el vómito autoinducido, el uso excesivo de laxantes, el ayuno o el ejercicio excesivo, estos ciclos pueden causar daños significativos en el cuerpo, incluyendo erosión del esmalte dental y callosidades en las manos, conocidas como el signo de Russell. Las personas con bulimia nerviosa a menudo mantienen un peso normal, lo que puede dificultar su detección. Sin embargo, los vómitos repetidos pueden llevar a problemas médicos graves, como esofagitis, pancreatitis y desequilibrios de electrolitos que pueden causar problemas cardíacos y, en casos extremos, la muerte. De acuerdo con la NEDA los trastornos de conducta de bulimia nerviosa se tienen que el 1.5% de las mujeres lo han sufrido y el 0.5% de los hombres lo han tenido.[4]. Vea figura 1.

1.1.3 Trastorno por atracón

El trastorno por atracón implica episodios recurrentes de comer en exceso, acompañados de una sensación de pérdida de control. A diferencia de la bulimia, no se llevan a cabo comportamientos compensatorios después de los atracones, lo que a menudo conduce a la obesidad y problemas de salud relacionados. Las complicaciones del trastorno por atracón incluyen un mayor riesgo de enfermedades como diabetes tipo 2, hipertensión, problemas cardíacos y apnea del sueño. La obesidad resultante del trastorno por atracón también puede llevar a problemas psicológicos, como baja autoestima, depresión y ansiedad. Actualmente las que sufren más de este trastorno son las mujeres teniendo un porcentaje de 3.5% y los hombres del 2%. [3] Vea figura 1.

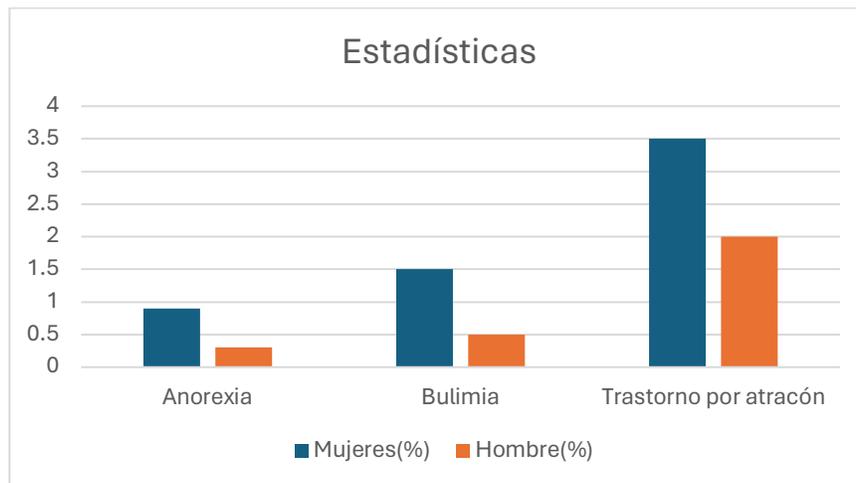


Fig. 1. Estadísticas de trastornos en mujeres y hombres. [5]

1.2 Complicaciones asociadas

Las complicaciones asociadas a los trastornos alimenticios incluyen problemas físicos como desequilibrios electrolíticos, daño cardíaco, osteoporosis, problemas gastrointestinales y daño renal. Psicológicamente, pueden causar depresión, ansiedad, distorsión de la imagen corporal y

conductas autolesivas. Socialmente, llevan al aislamiento, conflictos interpersonales y bajo rendimiento académico o laboral, en casos graves, pueden resultar en insuficiencia orgánica o muerte.

Dentro de las complicaciones endocrinas, los trastornos alimenticios pueden causar alteraciones hormonales significativas. En las mujeres, se observa atrofia mamaria y en los hombres se presentan niveles bajos de testosterona y problemas testiculares. Además, hay un aumento en los niveles de cortisol y de la hormona del crecimiento, estas alteraciones pueden afectar la función reproductiva, el metabolismo y el sistema inmunológico, provocando una serie de problemas de salud que pueden ser difíciles de manejar y que requieren atención médica continua. Por otro lado, dentro de las complicaciones óseas, la osteopenia y la osteoporosis son comunes entre las personas con trastornos alimenticios debido a la desnutrición crónica. En los niños y adolescentes, puede haber un estancamiento en el crecimiento y un retraso en la maduración ósea, lo que aumenta el riesgo de fracturas.

La falta de nutrientes esenciales, como el calcio y la vitamina D, contribuye a la pérdida de densidad ósea, aumentando el riesgo de fracturas. En los adolescentes, la desnutrición puede interferir con el crecimiento y el desarrollo normal, resultando en baja estatura y problemas óseos a largo plazo. En caso de las complicaciones dermatológicas y dentales, la piel seca, agrietada o escamosa, así como el lanugo (vello fino en la piel) son comunes debido al aumento de cortisol. Las equimosis (moretones) pueden aparecer por la trombocitopenia. La bulimia nerviosa puede causar erosión del esmalte dental debido al ácido del vómito recurrente y callosidades en las manos por el vómito autoinducido, estos problemas de la piel y los dientes pueden ser dolorosos y afectar la autoestima de las personas, empeorando su situación emocional y su salud general.[5] En figura 2, pueden observarse las tendencias en las complicaciones.

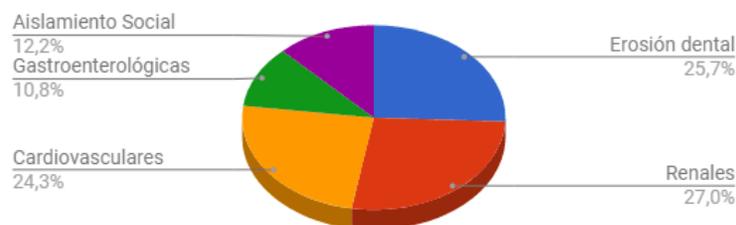


Figura 2. Porcentaje de complicaciones asociadas a trastornos alimenticios.

2 Diagnóstico y tratamiento de los trastornos alimenticios

El diagnóstico de los trastornos alimenticios es un proceso complejo que requiere una evaluación exhaustiva y detallada. Los profesionales de la salud deben realizar una serie de exámenes clínicos y entrevistas psicológicas para confirmar el diagnóstico y medir el alcance de los efectos de la enfermedad.

2.1 Exámenes clínicos

Los exámenes de laboratorio son esenciales para identificar desequilibrios hormonales y nutricionales, así como para descartar otras condiciones médicas que puedan presentar síntomas similares, como la diabetes, tumores Endocraneales, colagenopatías, problemas de tiroides y embarazo. Los análisis de sangre pueden revelar anemia, deficiencias de vitaminas y minerales, y alteraciones en los electrolitos. También se pueden realizar estudios de imagen, como radiografías o densitometría ósea, para evaluar la salud ósea y descartar fracturas osteoporosis.[6]

2.2 Evaluación psicológica

La entrevista psicológica es crucial para confirmar el diagnóstico de un trastorno alimenticio, evaluar la situación de riesgo e identificar los factores que desencadenan las conductas sintomáticas. Esta evaluación permite conocer el estado emocional del paciente, explorar su funcionamiento mental y rasgos de carácter, así como sus conflictos básicos, dificultades actuales y relaciones interpersonales. La evaluación debe incluir preguntas sobre la percepción del cuerpo, las conductas alimentarias, los patrones de ejercicio, el uso de laxantes o diuréticos, y la presencia de atracones o purgas, es importante evaluar el estado emocional del paciente, identificando posibles problemas como la depresión, la ansiedad o los trastornos de personalidad.

El diagnóstico diferencial también es crucial para descartar otras condiciones médicas que puedan presentar síntomas similares a los de los trastornos alimenticios. Es necesario realizar exámenes pertinentes para diferenciar entre trastornos alimenticios y condiciones como diabetes, tumores endocraneales, colagenopatías, problemas de tiroides y embarazo. En el caso de los pacientes que presentan vómitos, es esencial detectar signos físicos como pancreatitis, esofagitis, dilatación gástrica, estreñimiento o ruptura del tubo digestivo.[6]

2.3 Tratamiento

El tratamiento de los trastornos alimenticios requiere un enfoque multidisciplinario que involucra a médicos, nutricionistas, psicólogos y psiquiatras. El objetivo del tratamiento es restaurar el peso saludable del paciente, corregir los desequilibrios nutricionales, tratar las complicaciones médicas y abordar los problemas psicológicos subyacentes. El tratamiento médico puede incluir la hospitalización en casos graves, especialmente cuando hay complicaciones médicas serias o cuando el peso del paciente está críticamente bajo. La intervención nutricional es fundamental para ayudar al paciente a desarrollar hábitos alimentarios saludables y recuperar el peso perdido, esto puede incluir la supervisión de la ingesta calórica, la suplementación con vitaminas y minerales, y la educación nutricional. Por otro lado, la terapia cognitivo-conductual (TCC) es uno de los enfoques más efectivos para tratar los trastornos alimenticios, ayuda a los pacientes a identificar y cambiar los patrones de pensamiento y comportamiento que contribuyen a su trastorno alimenticio. La terapia familiar también puede ser beneficiosa, especialmente para adolescentes y jóvenes, al involucrar a los miembros de la familia en el proceso de recuperación y mejorar la comunicación y el apoyo familiar. Además, en algunos casos, los medicamentos pueden ser útiles para tratar los

síntomas de los trastornos alimenticios y las comorbilidades asociadas, como la depresión y la ansiedad. Los antidepresivos, como los inhibidores selectivos de la recaptación de serotonina (ISRS), pueden ser efectivos para reducir los síntomas de la bulimia nerviosa y el trastorno por atracón. Sin embargo, el uso de medicamentos debe ser supervisado cuidadosamente por un profesional de la salud.[7]

2.4 Herramientas tecnológicas de apoyo

La tecnología ha jugado un papel importante en el diagnóstico y apoyo al especialista de la salud para el tratamiento de pacientes. Enseguida presentamos una revisión del estado del arte.

- MindSurf
- MindSurf, creada en 2016 en México por Grecia Reynoso debido a su enfoque en ayudar a las personas a explorar su pensamiento, como una terapia cognitiva transdiagnóstica, que promueve el bienestar y la resolución de conflictos internos.[13] Inicialmente lanzada en Telegram, permitía a los usuarios interactuar con un chatbot que utilizaba un procesamiento de lenguaje natural (NLP) para detectar estrés, ansiedad, depresión y trastornos alimentarios (TCA). Además, incluía preguntas de estilo MOL para ayudar al público general y complementar terapias clínicas con monitoreo en tiempo real, incluyendo acceso a terapeutas en línea. Mindsurf emplea la API de GPT-4 supervisada y controlada para garantizar la seguridad de la información proporcionada al usuario. [14]
- JENNY
- Jenny es una solución mexicana diseñada para empresas, enfocada en el bienestar personal y la salud mental de los colaboradores. Desarrollada por Grecia Reynoso y un equipo de expertos, combina ciencia, psicología e inteligencia artificial para ayudar a las personas a conocerse mejor y manejar sus emociones. Utiliza NLP para procesar texto interesado por los usuarios y generar respuestas empáticas, así como aprendizaje automático (Machine Learning) para personalizar interacciones basadas en el historial y preferencias del usuario. [15]
- TCApp: Aplicación informática para trastornos alimenticios

HealthApp, conocida como TCApp, es una aplicación diseñada para asistir en el tratamiento de trastornos de la conducta alimentaria como bulimia y anorexia. La terapia cognitivo-conductual, empleada en estos casos, requiere que los pacientes registren sus estados de ánimo, pensamientos y sentimientos durante todo el día. TCApp digitaliza este proceso, permitiendo a los pacientes acumular información de manera fácil y divertida mediante técnicas de ludificación, como actividades y juegos.

Los datos recopilados están disponibles para la terapeuta, facilitando el seguimiento del paciente, la elaboración de informe, el envío de mensajes de refuerzo y la programación de alarmas para prevenir crisis. Ensayos realizados con TCApp, supervisados con la participación de médicos especialistas y la supervisión de los comités de ética de los hospitales colaboradores, muestran que

con esta aplicación se consigue una mayor adherencia a la terapia y un mejor seguimiento del paciente en su evolución día a día.[8]

TCApp utiliza un tipo de inteligencia artificial (IA), basada en algoritmos de aprendizaje automático con machine learning y procesamiento de lenguaje natural (NLP), permitiendo personalizar las recomendaciones y respuestas para los usuarios en cuestionarios o interacciones.

- LEA: Chatbot de apoyo psicológico

Lea es un chatbot con inteligencia artificial diseñado para brindar apoyo emocional y actuar como un psicólogo virtual. Se basa en tecnologías avanzadas y algoritmos de procesamiento de lenguaje natural, lo que le permite entender tus preguntas, emociones y problemas. Lea está disponible las 24 horas del día, los 7 días de la semana y está lista para ayudarte en cualquier momento. Lea garantiza confidencialidad y comodidad, convirtiéndose en una herramienta útil para gestionar ansiedad, depresión, trastornos de la conducta alimentaria y otros problemas de salud mental. Creada en México por Leanscout, Inc, bajo la dirección de Andrea Campos, Lea ofrece conversaciones terapéuticas simuladas, proporcionando recursos y herramientas personalizadas. Su tecnología IA la posiciona como un referente en el manejo de la salud emocional. Tiene una combinación de tecnologías de IA que son el procesamiento de lenguaje Natural (NLP), Análisis de sentimiento y Redes Neuronales están entrenadas con datos psicológicos y patrones de interacción para mejorar continuamente en la detección de señales de riesgos emocionales.[12]

- Chatbot Wysa: Apoyo basado en evidencia

Wysa es un chatbot que ofrece apoyo emocional y utiliza técnicas basadas en la evidencia para ayudar a los usuarios a manejar estrés, problemas de ansiedad y trastornos alimenticios. Ofreciendo un programa de recuperación utilizando agente conversacional digital, es factible y aceptable. Utiliza un lenguaje de aprendizaje automático para adaptar las respuestas y recomendaciones a las necesidades individuales del usuario. Creado en una parte de la India [9].

3 Metodología

Se realizó una investigación amplia sobre las evaluaciones para el diagnóstico de trastornos alimenticios en pacientes. Los especialistas en el área, reportan que para evaluar los trastornos alimenticios y en donde han tenido resultados más efectivos en diferentes contextos clínicos y de investigación es mediante el uso de cuestionarios de auto administrativo. Se trata de una prueba estructurada que evalúa cuatro áreas principales como lo son la preocupación por la alimentación, por el peso, por la figura y conducta de desorden alimenticio. El instrumento está compuesto por 20 reactivos que evalúan la frecuencia de síntomas en las últimas semanas. Es preferible por su amplia confiabilidad y validez en la evaluación que hace, ya que la prueba les permite a los pacientes responder de una manera más honesta y abierta. Lo realizan en privado y sin presión. Adicionalmente, el tratamiento de la información proporcionada permite una evaluación rápida y estandarizada de los síntomas, lo que ahorra tiempo en el proceso de diagnóstico. Haciendo que los psicólogos y nutriólogos puedan utilizar los resultados para priorizar y planificar las

intervenciones terapéuticas ya que brindan datos específicos sobre los patrones de alimentación y las preocupaciones de la imagen corporal, (NEA).[9]

El NEA es ampliamente reconocido por su confiabilidad y validez, ya que permite una evaluación rápida y estandarizada de los síntomas. Este instrumento ofrece a los pacientes un espacio seguro para responder de manera honesta y abierta, lo que mejora la calidad de los datos obtenidos. En el chatbot, NEA fue adaptado a un formato interactivo que clasifica automáticamente a los usuarios según el riesgo identificado, facilitando una intervención temprana y eficaz. [16]

3.1 Requerimientos para el chatbot

El diseño de un chatbot para el tema de salud mental y emocional implica la creación de un sistema que pueda interactuar de manera eficiente y empática con los usuarios que buscan apoyo en su salud mental y nutricional. Los requerimientos abarcan tanto aspectos técnicos como funcionalidades orientadas a las necesidades específicas de los pacientes.

Aspectos generales:

- Registro de usuarios y acceso a historiales.
- Chat en tiempo real para crisis.
- Respuestas personalizadas según sus necesidades.
- Registro de conversaciones.
- Programar citas con especialistas.
- Interfaz intuitiva y fácil de usar.
- Accesibilidad y estabilidad.

Requerimientos específicos para trastornos alimenticios

- Evaluación de riesgos para planificar intervenciones.
- Registro de sus alimentos y bebidas.
- Monitoreo de progreso: Los nutriólogos y psicólogos podrán monitorear el progreso del paciente y de acuerdo con eso ajustar su plan de alimentación según sea necesario.
- Apartado de evaluación de estado de ánimo: Al entrar al chatbot se colocarán estados de ánimo para seleccionar de cómo se siente el día de hoy, para poder notificar a los especialistas.

3.2 IA empleada en chatbots

El uso de la inteligencia artificial generativa y las redes neuronales han transformado la asistencia emocional digital. Estas tecnologías permiten interacciones naturales y empáticas, cruciales en la salud mental. Los modelos de IA generativa generan texto coherente y contextual, mientras que las redes neuronales procesan datos complejos y detectan patrones emocionales. Con controles adecuados para evitar respuestas inapropiadas o riesgosas, garantizando así un comportamiento ético y seguro. Los chatbots basados en IA son herramientas seguras y efectivas para brindar apoyo emocional. [17] Por un lado, la IA generativa son los modelos basados en redes neuronales profundas (deep learning), tiene la capacidad de aprender y generar texto de manera coherente y contextual. Esto es crucial en situaciones donde los usuarios buscan comprensión y apoyo, ya que

una respuesta emocionalmente adecuada puede influir en la experiencia del usuario y su bienestar. Por otro lado, las redes neuronales imitan el funcionamiento del cerebro humano y son particularmente efectivas en la detección de patrones complejos dentro de los datos, su estructura en capas permite procesar información de forma más profunda, lo cual la hace muy esencial para entender las emociones humanas. Las redes pueden identificar señales sutiles en la comunicación del usuario, facilitando así las respuestas que ayuden en su estado emocional.[18]

3.3 Tecnologías empleadas

Para desarrollar de nuestro chatbot de salud mental se usaron las siguientes tecnologías:

- Visual Studio code: Es un entorno integrado (IDE – Integrated Development Environment) creado por Microsoft para desarrollar aplicaciones de software en diversos lenguajes. Ofrece herramientas para escribir, depurar, compilar y mantener código de manera eficiente.
- Java Script: Es un lenguaje de programación utilizado hoy en día, sobre todo en aplicaciones web, y aprendizaje automático.
- WebSockets: para la comunicación en tiempo real. Discord mantiene conexiones persistentes entre los usuarios y los servidores, permitiendo la transmisión instantánea de mensajes, notificaciones y actualizaciones.
- APIs y Bots: Discord proporciona una API robusta que permite a los desarrolladores crear Bots personalizados para automatizar tareas, interactuar con los usuarios y enriquecer la experiencia de los servidores. Estos bots pueden integrarse con inteligencia artificial (IA).

3.4 Diagrama de secuencia

Con el diagrama de secuencia se tiene que el actor principal en este caso un usuario quiere informarse y saber si tiene algún trastorno alimenticio, el diagrama muestra el flujo de interacción entre el usuario, el chatbot y las entidades relacionadas con el cuestionario de trastornos alimenticios.

El usuario interactúa directamente con el chatbot, *consultar* representa las etapas donde el chatbot recoge y procesa información, el chatbot realiza las tareas de generación de respuestas, evaluación de datos y entrega de resultados. (Figura 3)

- El usuario inicia la interacción con el chatbot.
- Preguntar: usuario responde una pregunta de NEA.
- Enviar pregunta la información proporcionada por el usuario se envía el chatbot para su procesamiento.
- Generando respuestas el chatbot procesa la pregunta y genera una respuesta en función de los algoritmos y datos almacenados en este caso en los puntajes.
- Mostrar Resultados chatbot evalúa las respuestas del usuario acumuladas hasta el momento y presenta los resultados en forma de mensaje.
- Resultados se entregan y en caso de ser niveles medio o alto sugiere la consulta con un especialista (psicólogo).

3.5 Diagrama de interacción

En el diagrama de interacción se representa el funcionamiento general, en donde se analiza el flujo de interacción de datos. (Figura 4)

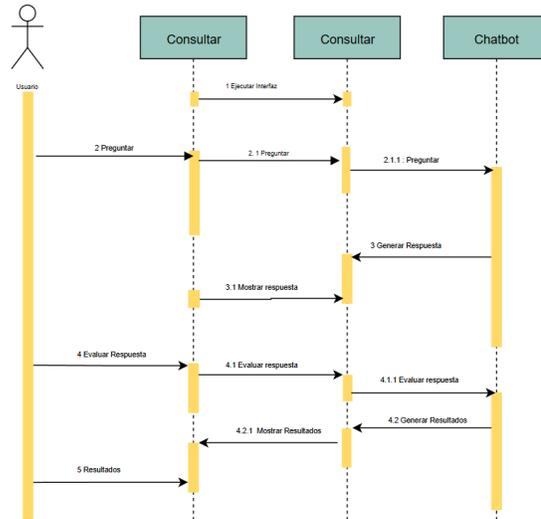


Figura 3. Diagrama de secuencia de chatbot.

- Un paciente completa el cuestionario en el chatbot.
- Los resultados se almacenan en “Evaluación NEA” y se vinculan con las preguntas respondidas.
- Si el resultado indica un nivel de riesgo alto, el sistema genera una alerta para el médico.
- Se programa cita entre el paciente y médico.
- Tras la evaluación, se genera un expediente con los resultados y recomendaciones.

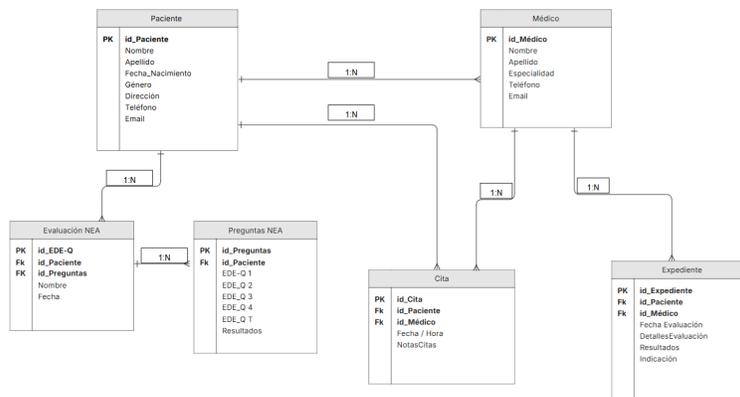


Fig. 4. Diagrama de cómo interactúa el modelo.

3.5 Diagrama de IA

El diagrama detalla cómo la inteligencia artificial (IA) interactúa en el flujo de recepción de datos, su procesamiento y cómo es que la red neuronal identifica patrones para responder de una manera eficiente al usuario se puede observar en la figura 5.

- Recepción de datos en donde el usuario envía un mensaje desde la plataforma Discord y Discord actúa como intermediario reenviando el mensaje al servidor.
- El servidor inicia el preprocesamiento del mensaje en donde se realiza una limpieza eliminando caracteres no deseados, errores o ruido.
- Posteriormente el servidor analiza el mensaje y determina su intención ya sea en pregunta si requiere información o afirmación si podría aportar contexto adicional.
- El servidor envía los datos procesados a la red neuronal.
- La red neuronal realiza su procesamiento en donde IA convierte los datos en vectores numéricos, facilitando que la red procese información textual, analice los patrones y comprenda las preguntas o respuestas enviadas.
- Se genera la respuesta la IA envía la respuesta al servidor que ya esté bien y el servidor la reenvía a Discord.

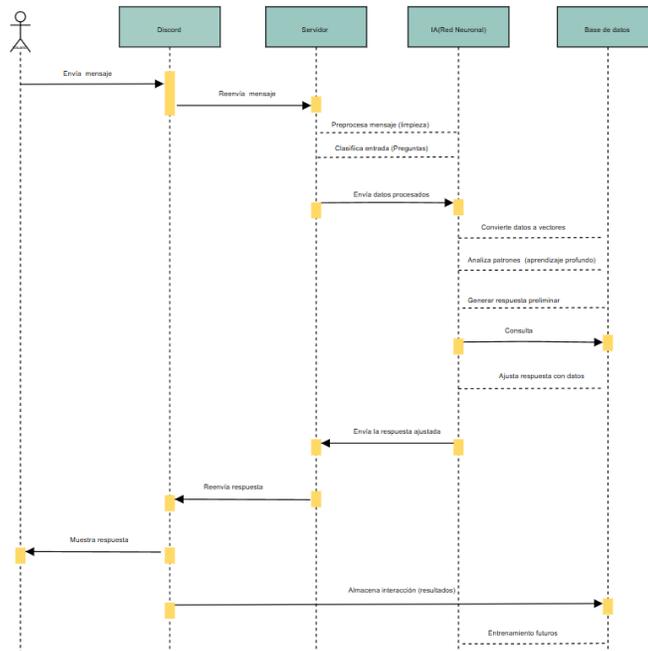


Fig. 5. Diagrama de cómo interactúa la IA.

4 Resultados

La interfaz del chatbot diseñada para Discord presenta una experiencia conversacional sencilla y efectiva. El chatbot está orientado para evaluar el riesgo de trastornos alimenticios mediante el cuestionario basado en NEA (National Eating Assessment). Su estructura de la interfaz principal opera dentro de un canal de texto de Discord, etiquetado como general, los usuarios interactúan en tiempo real con el Bot mediante mensajes en este canal, se formulan las preguntas y se reciben respuestas.

Los usuarios, identificados por su nombre, envían las respuestas al chatbot. El bot aparece con el nombre de “Mental APP” y formula las preguntas o confirma la recepción de las respuestas. Bot presenta preguntas del cuestionario de NEA utilizando un lenguaje claro y de opciones específicas de respuestas, los usuarios responden, seleccionado entre las opciones sugeridas por el bot o escribiendo sus propias respuestas si es necesario, como lo puede observar en las figuras 6 y 7.



Fig. 6. Interfaz del chatbot.

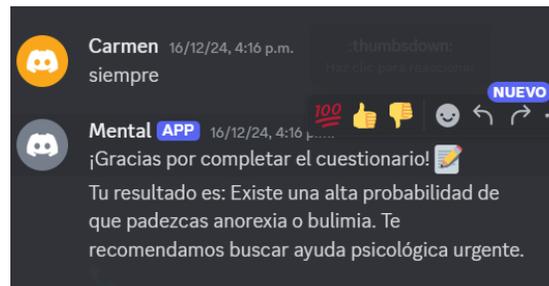


Fig. 7. Resultados Chatbot.

4.1 Resultados parciales

El chatbot fue probado con un grupo de alumnos universitarios con edades comprendidas entre los 18 y 25 años. Los resultados preliminares mostraron los siguiente: (figura 8)

- El 20% de los usuarios fueron identificados con un riesgo alto de trastornos alimenticios.
- El 30% mostró comportamientos que probablemente se estuviera desarrollando un trastorno.
- El 50% no sufre de ningún trastorno alimenticio.

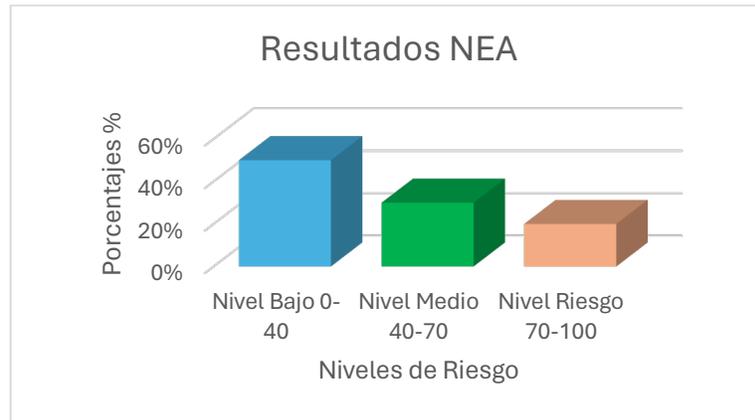


Fig. 8. Estadística de resultados de chatbot NEA.

De los resultados obtenidos destacamos que los chatbots son herramientas elaboradas para la detección temprana de trastornos alimenticios.

5 Conclusiones y trabajos futuros

5.1 Conclusiones

Los trastornos alimenticios son enfermedades complejas que requieren un enfoque diagnóstico y terapéutico integral y personalizado, la identificación temprana y el tratamiento adecuado son esenciales para prevenir complicaciones graves y mejorar la calidad de vida de las personas afectadas. La educación y la concientización sobre estos trastornos son cruciales para reducir la estigmatización y promover una intervención temprana y eficaz; y es aquí donde la tecnología juega un papel crucial proveyendo de herramientas inteligentes que permiten en tiempo real y de manera interactiva conocer sobre nuestros padecimientos y el especialista de la salud tome decisiones oportunas.

El diagnóstico y tratamiento de los trastornos alimenticios pueden ser apoyados por la IA de la mano con el especialista de la salud.

El chatbot obtenido es una herramienta eficaz para identificar trastornos alimenticios en etapas iniciales, gracias al uso del instrumento NEA. Su accesibilidad, diseño intuitivo y capacidad para proporcionar recursos inmediatos lo convierte en un recurso valioso en el ámbito de la salud mental. Además, de ser un instrumento accesible, el chatbot fomenta la educación sobre los trastornos alimenticios, lo que contribuye a disminuir la estigmatización.

5.2 Trabajos futuros

Sin duda ampliar el diagnóstico a otros padecimientos como anorexia y bulimia y enfocarse en ellos para combatirlos.

Seguimiento a largo plazo, implementar funciones que permitan monitorear el progreso de los usuarios identificados en riesgo y evaluar la efectividad de las recomendaciones proporcionadas.

Agradecimientos.

Agradezco a mi asesora María del Carmen Santiago Díaz por la orientación valiosa y exitosa que me dio, su apoyo constante y los comentarios de retroalimentación enriquecieron nuestro trabajo.

Referencias

[1] Bautista, I., & Serra, L. (2015). Prevalencia de los trastornos de la conducta alimentaria en adolescentes de Gran Canaria. Nutrición <https://scielo.isciii.es/pdf/nh/v31n5/50originalotros04.pdf>

[2] Núñez-Chaves, A. (2016). Trastornos de la alimentación: un verdadero problema en adolescentes. Ciencias <https://www.uhsalud.com/index.php/revhispano/article/view/125/79>

[3] Sánchez, I. (2021). Impacto psicológico de la COVID-19 en niños y adolescentes. Medisan, 25(1), 123-141. <http://scielo.sld.cu/pdf/san/v25n1/1029-3019-san-25-01-123.pdf>

[4] Guía de Trastornos Alimenticios. (2004). printed and made in Mexico. (Obra original publicada en 2004)

[5] GÓMEZ-PERESMITRE G, ALVARADO G, MORENO L, SALOMA S, PINEDA G. Trastornos de la alimentación. Factores de riesgo en tres diferentes grupos de edad: pre-púberes, púberes y adolescentes. Revista Mexicana de Psicología, 2001;18(3): 313-324.

[6] Tratamiento de la imagen corporal en los trastornos alimenticios cambio clínicamente significativo. (2014). anales de psicología, 30(2), Artículo 2.

[7] Trastornos mentales. (s.f.). World Health Organization (WHO). <https://www.who.int/es/news-room/fact-sheets/detail/mental-disorders>.

[8] (s.f.). SciELO España - Scientific Electronic Library Online. <https://scielo.isciii.es/pdf/nh/v31n5/50originalotros04.pdf>

[9] Evaluating a Digital Mental Health Intervention (Wysa) for Workers' Compensation Claimants: Pilot Feasibility Study. (s.f.). PubMed Central (PMC). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9897276/>

[10] Effectiveness of a Web-based and Mobile Therapy Chatbot on Anxiety and Depressive Symptoms in Subclinical Young Adults: Randomized Controlled Trial - PubMed. (s.f.).

PubMed. <https://pubmed.ncbi.nlm.nih.gov/38506892/>

[11] Effectiveness of a Web-based and Mobile Therapy Chatbot on Anxiety and Depressive Symptoms in Subclinical Young Adults: Randomized Controlled Trial. (s.f.). PubMed Central (PMC). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10993129/>

[12] Towards an Artificially Empathic Conversational Agent for Mental Health Applications: System Design and User Perceptions. (s.f.). Journal of Medical Internet Research. <https://www.jmir.org/2018/6/e10148>

[13] MindSurf: a pilot study to assess the usability and acceptability of a smartphone app designed to promote contentment, wellbeing, and goal achievement. (s.f.). PubMed Central (PMC). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5153894/>

[14] Guevara, C. F. (2024, 16 de enero). Mindsurf, psicología aplicada. REFORMA. <https://www.reforma.com/mindsurf-psicologia-aplicada/ar2741160>

[15] Jenny. (s.f.). Jenny. <https://jenny.com.ai/>

[16] Unikel-Santoncini, C., Bojórquez-Chapela, I., & Carreño-García, S. (s.f.). Validación de un cuestionario breve para medir conductas alimentarias de riesgo. https://www.scielo.org.mx/scielo.php?pid=S003636342004000600005&script=sci_arttext.

[17] El papel de la inteligencia artificial en la atención de salud mental | RMHP . (sf). Dove Medical Press - Editorial de acceso abierto de revistas médicas. <https://www.dovepress.com/exploring-the-role-of-artificial-intelligence-in-mental-healthcare-cur-peer-reviewed-fulltext-article-RMHP>

[18] Descripción general de las aplicaciones móviles de salud mental basadas en chatbots: perspectivas extraídas de la descripción de la aplicación y las reseñas de los usuarios . (sf). PubMed Central (PMC). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10242473/>

***Avances en Computación:
Perspectivas Académicas y de Investigación en
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

***Avances en Computación:
Perspectivas Académicas y de Investigación en
Inteligencia Artificial, Ciencia de Datos y Seguridad***
Coordinado por Gustavo Trinidad Rubín Linares

