

APORTACIONES DE LAS CIENCIAS COMPUTACIONALES DURANTE LA PANDEMIA COVID19



Gustavo Trinidad Rubín Linares

APORTACIONES DE LAS CIENCIAS COMPUTACIONALES DURANTE LA PANDEMIA COVID19

APORTACIONES DE LAS CIENCIAS COMPUTACIONALES DURANTE LA PANDEMIA COVID19

María del Carmen Santiago Díaz
Gustavo Trinidad Rubín Linares
Ana Claudia Zenteno Vázquez
Yeiny Romero Hernández
Judith Pérez Marcial
(Editores)

Gustavo Trinidad Rubín Linares
(Coordinador)

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

Gustavo Trinidad Rubín Linares
(coordinador BUAP)

María del Carmen Santiago Díaz, Gustavo Trinidad Rubín Linares, Ana Claudia Zenteno Vázquez, Yeiny Romero Hernández, Judith Pérez Marcial, María Blanca del Carmen Bermúdez Juárez, María de Lourdes Sandoval Solís, Luis Carlos Altamirano Robles, Irene Olaya Ayaquica Martínez, Abraham Sánchez López, María Josefa Somodevilla García, Maya Carrillo Ruíz, Hilda Castillo Zacatelco, Luis Enrique Colmenares Guillén, Josefina Guerrero García, Meliza Contreras González, Gabriel Juárez Díaz, José de Jesús Lavalle Martínez, Miguel Ángel León Chávez, Ivo Humberto Pineda Torres, Roberto Contreras Juárez, Graciano Cruz Almanza, Andrés Vázquez Flores, Darnes Vilariño Ayala, Héctor David Ramírez Hernández, Nelva Betzabé Espinoza Hernández, José Martín Estrada Analco, Rogelio González Velázquez, Pedro García Juárez, Beatriz Beltrán Martínez, Ana Luisa Ballinas Hernández, Nicolás Quiroz Hernández, Luz del Carmen Reyes Garcés, Alba Maribel Sánchez Gálvez, Alfredo Toriz Palacios, Rogelio Alfredo Campos Serapio.
(revisores locales)

Eden Belouadah, Jessica Nayeli López Espejel, Paola Eunice Rivera Salas, María Concepción Landa Arnaiz, Francisco Marroquín González, Oleg Starostenko Basarab, Raúl Antonio Aguilar Vera, Juan Pablo Ucán Pech, Carina Toxqui Quitl, Juan Alberto Guevara Jaramillo, Heidy Marisol Marín Castro, Julio Cesar Díaz Mendoza, Hermes Moreno Álvarez
(revisores externos)

Primera edición: 2021
ISBN: 978-607-8728-80-0

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:
Mtro. 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	7
Chatbot acerca de Alimentación Saludable <i>Luis René Marcial Castillo</i> <i>Cecilia América Barranco Herrera</i> <i>Marcela Rivera Martínez</i> <i>María de Lourdes Sandoval Solís</i>	8
Sistema de Adquisición de Datos Vehiculares por medio del OBD-II con Interfaz Gráfica en Labview para un Posterior Uso en Aplicaciones de IoV (Internet de los Vehículos). <i>Nicolás Quiroz Hernández</i> <i>Saul Luna Minor,</i> <i>Roberto Carlos Ambrosío Lázaro</i>	17
Comparativa de Herramientas para Realizar un Análisis a través del Cómputo Forense en CentOS 8.0 <i>Ana Claudia Zenteno Vázquez</i> <i>María del Carmen Santiago Díaz</i> <i>Yeiny Romero Hernández</i> <i>Judith Pérez Marcial</i> <i>Gustavo Trinidad Rubín Linares,</i> <i>Miguel Ángel Pérez Xilo</i> <i>María de los Ángeles Pájaro Cozatl</i>	29
Sobre Sistemas de Lógica Difusa para la Predicción de Riesgo de Padecer de Diabetes <i>Luis Enrique Morales Márquez</i> <i>Maya Carrillo Ruiz</i> <i>Pedro García Juárez</i>	37
Análisis de Herramientas para el Testeo de Vulnerabilidades de Aplicaciones Web <i>Ricardo Emmanuel García Cruz</i> <i>Ana Claudia Zenteno Vázquez</i> <i>María del Carmen Santiago Díaz</i> <i>Yeiny Romero Hernández</i> <i>Judith Pérez Marcial</i> <i>Gustavo Trinidad Rubín Linares</i>	46
Análisis de Comentarios de Cambio Climático en Twitter bajo Técnicas de Procesamiento de Lenguaje Natural <i>María Beatriz Bernábe Loranca</i> <i>Gerardo Martínez Guzmán</i>	

<i>Carmen Cerón Garnica</i>	
<i>Alberto José Luis Carrillo Canán</i>	52

Fundamentos del Álgebra Lineal en la Computación Cuántica

<i>Héctor David Ramírez Hernández</i>	
<i>Roberto Contreras Juárez</i>	
<i>Nelva Betzabel Espinoza Hernández</i>	
<i>Yusdivia Huerta Hernández</i>	65

Diseño de un Robot Sanitizante anti COVID-19 con Navegación Autónoma

<i>María del Carmen Santiago Díaz</i>	
<i>Ana Claudia Zenteno Vázquez</i>	
<i>Judith Pérez Marcial</i>	
<i>Yeiny Romero Hernández</i>	
<i>Gustavo Trinidad Rubín Linares</i>	
<i>Hermes Moreno Álvarez</i>	
<i>Michel Yadira Montelongo Flores</i>	
<i>Alejandro Vargas Pérez</i>	73

Prefacio

A más de dos años del inicio de la pandemia de COVID-19 el mundo no ha logrado contener sustancialmente la pandemia, muchos países han sufrido las variantes Delta, Omicron y se encuentran apenas saliendo de la cuarta ola de contagios y con la incertidumbre de cuántas olas más surgirán, cuántas variantes y mutaciones, y más aún cuántas vacunas o refuerzos de las existentes vendrán, incluso el tratamiento específico en los hospitales no es caso cerrado, hay una diversidad de estos. Aunque ha disminuido el número de defunciones no ha sido de forma tan radical y no hay garantía de que así sea, en conclusión estamos inmersos en un caos que nosotros mismos generamos, todos los estudios realizados no han podido ser concluyentes y no se ha parado de hacer esfuerzos tanto en generar herramientas de mitigación de la pandemia como para disminuirla, sin embargo, somos los herederos del método científico, la generación que debe resolver los problemas actuales y de generar más y mejor conocimiento así como su aplicación, por ello en este ejemplar se presentan algunos trabajos de Inteligencia Artificial, Ciencia de Datos, Internet de las Cosas, Ciberseguridad y otros, que buscan mostrar aplicaciones directas o indirectas en la actual contingencia del COVID-19.

Actualmente la premisa para generar y aplicar el conocimiento no es solo para brindar mejores condiciones de vida para los seres humanos, sino ha cambiado al punto de generar condiciones de supervivencia, nunca hubiéramos imaginado estar en éste punto, sin embargo, la pelea contra este enemigo invisible debe continuar, y por ello aunque se diga y suene bonito que “nunca hubo tanta información disponible, ni hubo tantos esfuerzos en torno a un mismo problema...”, si debemos continuar avanzando en el desarrollo tecnológico y científico, y sobre todo brindando a las siguientes generaciones las armas del método científico que aunque puede parecer de poca utilidad para algunos, sabemos que es el único capaz de permitirnos cambiar el destino que muchos creen que ya está escrito.

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

Chatbot acerca de Alimentación Saludable

Luis René Marcial Castillo, Cecilia América Barraco Herrera,
Marcela Rivera Martínez, Lourdes Sandoval Solís
Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Computación, Puebla, Pue. México
{luis.marcial, marcela.rivera, lourdes.sandoval}@correo.buap.mx
{cecilia.barranco}@alumno.buap.mx

Resumen. La importancia de llevar una alimentación saludable se está volviendo cada vez más preocupante, tal es el impacto que se está recurriendo a la inteligencia artificial para conseguir una alimentación mucho más sana. Cada vez vemos más el desarrollado de chatbots a través de los cuales se hacen recomendaciones nutricionales a diferentes tipos de familias que junto con el compromiso y la voluntad han empezado a mejorar su educación alimentaria. En este trabajo se desarrolla un chatbot sobre educación en comida saludable basado en las recomendaciones del IMSS, obteniendo una aplicación que puede ejecutarse en un navegador y en la red social de telegram.

Palabras clave: Chatbot, inteligencia artificial, Alimentación saludable, Telegram.

1 Introducción

Un chatbot es un programa de cómputo cuyo propósito fundamental es que la comunicación con las personas sea personalizada, próspera, útil, eficiente y con cierto grado de fidelidad. Los chatbots ha evolucionado desde que el profesor del MIT Joseph Weizenbaum creara ELIZA en los años 60 del siglo pasado, la empresa china Wechat fue la primera que contó con una plataforma para el desarrollo de chatbots desde el año 2013, en el año 2001 se inicia el desarrollo del chatbot Eugene Goostman que es un bot conversacional desarrollado por tres programadores; el ruso Vladimir Veselov, el ucraniano Eugene Demchenko, y el ruso Sergey Ulasen, para el 7 de Junio de 2014, en una prueba para medir la inteligencia artificial organizado por la Royal Society de Londres, Reino Unido, Eugene Goostman logró convencer a un 33% de los 150 jueces humanos que tenían que decidir si estaban conversando con una persona o no, por lo que se consideró que es el primer chatbot en pasar la prueba de Turing. En los últimos años el uso de chatbots evolucionó rápidamente en numerosos campos incluyendo marketing, sistemas de apoyo, educación, atención de la salud, patrimonio cultural y entretenimiento [1, 6]. Hay varios trabajos recientes sobre el desarrollo de chatbots: Nichifor, Trifan y Nechifor en el año 2021 [8] realizan un estudio que tiene como objetivo cubrir empíricamente el impacto del uso de la inteligencia artificial a través de chatbots en el comercio minorista en línea, en términos de contenido implementado en el proceso de comunicación. La investigación presentada aporta una contribución a la literatura especializada al analizar la utilidad percibida y demostración de la instalación, conceptos clave de la aceptación de tecnología modelo en la creación de chatbots; Mohamed, Zhuopeng en el año 2020 [7], a partir de la revisión de diferentes revistas concluyen que el uso de chatbots debe ser fácil de usar y pueden ser utilizados por cualquier persona que

sepa escribir en su propio idioma en la aplicación móvil o en la versión de escritorio, proponen un esquema de chatbot médico que proporciona diagnósticos personalizados basados en los síntomas en donde el reconocimiento de síntomas y el rendimiento del diagnóstico del bot podrían mejorarse mucho al agregar soporte para más características médicas, por ejemplo, ubicación, duración e intensidad de los síntomas, y una descripción más detallada de los síntomas. La implementación de un asistente médico personalizado se basa en gran medida en algoritmos de inteligencia artificial, así como en los datos de entrenamiento. La era del futuro es la era de la aplicación de mensajería porque la gente pasará más tiempo en la aplicación de mensajería que en cualquier otra aplicación. Por lo tanto, los chatbots médicos tienen un alcance futuro enorme y amplio. No importa qué tan lejos estén, la gente puede tener esta conversación médica; Kedar, Dahale, Patel, Lahamange y Chordiya en el 2020 [5] presentan el principio general de funcionamiento y los conceptos básicos de los chatbots basados en inteligencia artificial y conceptos relacionados, así como sus aplicaciones en varios sectores como telecomunicaciones, banca, salud, centros de atención telefónica para clientes y comercio electrónico. Además, muestran los resultados de un chabbot de ejemplo para el servicio de donación desarrollado para el proveedor de servicios de telecomunicaciones; Athota, Shukla, Pandey, Rana en el 2020 [2] presentan un chatbot que ayuda a los usuarios, sobre temas de salud. Inicialmente, cuando el usuario visita el sitio web, se registran ellos mismos y luego pueden hacer sus consultas al bot. El sistema utiliza un sistema experto para responder a las consultas. Los datos del chatbot son almacenados en la base de datos en forma de patrón-plantilla. Aquí se usa SQL para manejar la base de datos; Bulla, Parushetti, Teli, Aski y Koppad en el 2020 [3] presentan un chatbot médico que es fácil de usar y puede ser utilizado por cualquier persona, se presenta una aplicación móvil y una versión de escritorio, el chatbot proporciona diagnósticos personalizados basados en los síntomas del paciente.

Como se puede notar el desarrollo de chatbots es un área en franco desarrollo, de ahí que este trabajo, propone la realización de un chatbot de tipo declarativo que tiene definido un propósito específico y requiere de las interacciones con el usuario. Este tiene un funcionamiento basado en una serie de patrones, es decir, para cada palabra, frase o pregunta que el usuario escribe, el chatbot tiene programada una respuesta determinada. Esto implica que solo puede responder a las palabras o frases que tenga registradas.

La estructura del documento es la siguiente: en la sección dos se define el problema, en la sección tres se proporciona el desarrollo del chatbot, en la sección cuatro se muestran las pruebas y resultados obtenidos, las conclusiones y el trabajo futuro están plasmados en el capítulo 5, finalmente se listan las referencias bibliográficas utilizadas en el desarrollo de este trabajo.

2 Definición del Problema

Existen dos tipos de chatbots [4][9]:

Chatbots orientados a tareas también conocidos como declarativos. Tienen un funcionamiento sencillo basado en una serie de patrones. Son para un solo propósito que se enfocan en realizar una sola función. Generan respuestas automáticas pero conversacionales, usualmente de temas con pocas variables. Son altamente específicos y estructurados. Los chatbots orientados a tareas están diseñados para una tarea en particular y están configurados para tener conversaciones breves, generalmente dentro de un dominio cerrado, se basan en una serie de patrones y generan respuestas automáticas.

Chatbots no orientados a tareas. Este tipo de chatbots son más generales, pueden simular una conversación con una persona y parecen realizar una charla con fines de entretenimiento en dominios abiertos.

La propuesta fue crear un chatbot orientado a tareas o de tipo declarativo, ya que se tiene definido un propósito específico y requiere aprender de las interacciones con el usuario. Este tiene un funcionamiento basado en una serie de patrones, es decir, para cada palabra, frase o pregunta que el usuario escribe, el chatbot tiene programada una respuesta determinada. Esto implica que solo puede responder a las palabras o frases que tenga registradas.

3 Desarrollo del Chatbot para la Educación sobre un Estilo de Vida Saludable

El funcionamiento del chatbot se basa en las recomendaciones del Instituto Mexicano del Seguro Social (IMSS) en lo que se refiere a comida saludable y se basa en una serie de patrones. Es decir, para que el chatbot nos dé una respuesta necesita tener almacenado en un diccionario palabras clave para así “entender” una palabra, frase o pregunta que el usuario escribe, para esto el chatbot tiene programada una respuesta determinada. Esto implica que solo puede responder a las palabras o frases que tenga registradas, en este caso una vez que el chatbot encuentre una palabra clave en una frase que haya introducido el usuario éste nos dará una respuesta.

El chatbot implementado tiene como objetivo que las personas puedan conocer información sobre cómo llevar una vida saludable, tocando temas como: actividad física, comida saludable, sobrepeso y obesidad, bajar de peso, plan de alimentación saludable, tipos de colesterol, vitaminas, grasas, saber en qué nos ayudan las frutas y verduras, así como conocer acerca de los tipos de nutrientes, derivándose en micronutrientes y macronutrientes, entre otras de una forma muy sencilla.

Para llevar esto a cabo se ocupó información de “El decálogo para un buen estilo de vida” de la página oficial del IMSS; en la cual se abordan los siguientes temas: ingesta diaria recomendada de diferentes grupos de alimentos, información nutricional, dieta saludable y estilo de vida. El chatbot está disponible en el sitio web “<https://salubot.aplicacionesdemo.com/>” y también se puede encontrar en la red social Telegram con el nombre de “Salubot”. La versión web fue realizada con HTML5, con la lógica programada en Javascript. Las herramientas utilizadas para el chatbot desarrollado en la red social Telegram son Dialogflow de Google Cloud para la lógica del chatbot (backend) y Botfather de Telegram para la interacción con el usuario (frontend).

El chatbot contiene lo siguiente: **Saludo inicial del bot.**

```
imprimir_linea_salubot("Hola, yo soy Salubot un asistente que te hablará de la importancia de tener comidas saludables, bienvenido!");
```

Para contestar el saludo del usuario:

```
registrar_en_diccionario("Hola",["hola", "lindo día", "lindo día"],[ "Hola, me da gusto que te interese tener una vida saludable" , "Hola, espero que te encuentres muy bien, ¿Con qué te gustaría empezar para tener una vida saludable?"], [""], false);
```

A continuación, se listan algunas preguntas y respuestas que incluye el menú del chatbot y que se guardan en el diccionario.

¿Qué es la comida saludable?

Una comida saludable resalta la importancia de las frutas, las verduras, los cereales integrales, la leche y los productos lácteos sin grasa o bajos en grasa. Incluye una variedad de alimentos con proteínas como mariscos, carnes magras, huevos, legumbres (frijoles y chícharos), productos de soya, nueces y semillas. Contiene poca cantidad de grasas saturadas, grasas transaturadas, colesterol, sal (sodio) y azúcares agregados. Se mantiene dentro de sus necesidades calóricas diarias.

¿Por qué es importante la actividad física?

Claro, tomemos en cuenta que la actividad física regular es importante para tener buena salud, en especial si está intentando bajar de peso o mantener un peso saludable. La actividad física también ayuda a: mantener el peso, reducir la presión arterial alta, reducir el riesgo de diabetes tipo 2, ataque cardíaco, accidente cerebrovascular y varios tipos de cáncer. Reducir el dolor de la artritis y la discapacidad asociada con esta afección. Reducir el riesgo de osteoporosis y las caídas. Reducir los síntomas de depresión y ansiedad.

¿Qué es el sobrepeso?

Los términos sobrepeso y obesidad se refieren a un peso corporal que es mayor de lo que se considera normal o saludable para cierta estatura. El sobrepeso generalmente se debe a la grasa corporal adicional. Sin embargo, el sobrepeso también puede ocurrir debido a exceso de músculo, hueso o agua.

¿Qué es la Obesidad?

La obesidad es una enfermedad compleja que consiste en tener una cantidad excesiva de grasa corporal. La obesidad no es solo un problema estético, es un problema médico que aumenta el riesgo de enfermedades y problemas de salud, tales como enfermedad cardíaca, diabetes, presión arterial alta y ciertos tipos de cáncer. Por ello recomiendo llevar una vida un poco más saludable.

¿Qué son los Micronutrientes?

Los micronutrientes, generalmente derivados de la ingesta de alimentos, son pequeñas cantidades de vitaminas y minerales requeridos por el cuerpo para la mayoría de las funciones celulares. Las deficiencias más comunes de micronutrientes incluyen vitamina A, vitamina D, vitamina B12, hierro, yodo y zinc.

¿Qué son los Macronutrientes?

Grasas: son una de las principales fuentes de energía para el organismo, que al consumirlas en las cantidades adecuadas se convierten en un elemento primordial que ayuda al crecimiento, al desarrollo y a mantener una buena salud. Se debe tener en cuenta que las grasas se pueden encontrar en tres presentaciones: sólida (manteca), semisólidas (mantequillas o margarinas) o líquidas (aceites); cada una de ellas tiene efectos diferentes en el metabolismo. No todas las grasas son favorables; uno de estos casos es el de las grasas transaturadas, un tipo específico que se forma cuando los aceites líquidos se convierten en grasas sólidas como la manteca o la margarina en barra. Carbohidratos: constituyen la principal fuente de energía de nuestra alimentación, son el combustible para desarrollar las actividades diarias, mantener la temperatura corporal y el buen funcionamiento de órganos vitales del cuerpo. Proteínas: son las responsables de la formación de células, los tejidos y órganos, así como de construir los músculos, parte de las hormonas, de las enzimas que transportan ciertas moléculas (como la grasa) y transmitir señales. También están presentes en una gran cantidad de funciones del organismo. Durante los periodos de crecimiento, embarazo, lactancia y recuperación de enfermedades es necesario el consumo de alimentos fuentes de proteína en mayor cantidad. Las fuentes de proteína pueden ser de origen animal como huevos, pescados, leche, carnes magras, pavo, y pollo; o de origen vegetal como las leguminosas (frijol, lenteja, garbanzo, chícharo), nueces y frutos secos, quinoa, entre otros.

¿Cómo bajar de peso de forma saludable?

Bajar de peso en forma saludable no se trata solamente de seguir una dieta o programa, es llevar un estilo de vida estable que incluya cambios a largo plazo, en la alimentación diaria y los hábitos de ejercicio. Para bajar de peso, usted debe gastar más calorías de las que consume. Como una libra equivale a 3,500 calorías, necesitará reducir su consumo calórico en 500-1000 calorías por día para perder entre 1 y 2 libras por semana. Una vez que logra llegar a un peso saludable, usted tendrá menos probabilidades de recuperar el peso perdido a largo plazo si sigue una alimentación saludable y realiza actividad física la mayoría de los días de la semana (entre 60 a 90 minutos de intensidad moderada). Bajar de peso no es sencillo, así que ánimo, puedes lograrlo con una buena alimentación, y no olvides ejercitarte.

¿Que son los nutrientes?

Los nutrientes son aportados por los alimentos y se clasifican en dos grupos: los macronutrientes, del cual hacen parte las grasas, los carbohidratos, y las proteínas; y los micronutrientes compuestos por los minerales y las vitaminas.

¿Qué es el Colesterol?

El 70% del colesterol que circula en nuestro cuerpo es producido por el hígado, el otro 30% debe ser aportado por la alimentación, pero aportes mayores producen daños a la salud. Existen 2 tipos de colesterol: Colesterol HDL: llamado comúnmente colesterol bueno, tiene un efecto protector contra enfermedades cardiacas y evita el aumento del colesterol LDL en sangre. Colesterol LDL: conocido como colesterol malo, tiene un efecto nocivo en el organismo ya que aumenta el riesgo de obstrucción de las arterias del corazón, predisponiendo a las personas a padecer enfermedades cardiovasculares como el infarto o el accidente cerebrovascular.

¿Para qué sirve un plan alimenticio?

Un plan de alimentación que ayuda a controlar el peso incluye una variedad de alimentos saludables, agregando una variedad de colores al plato. Las verduras de hoja verde oscura, las naranjas y los tomates —incluso las hierbas frescas— están llenos de vitaminas, fibra y minerales. Agregar pimientos, brócoli o cebollas congelados a guisos y tortillas aumenta el color y los nutrientes. Un plan de alimentación saludable: resalta la importancia de las frutas, las verduras, los cereales integrales, la leche y los productos lácteos sin grasa o bajos en grasa.

Incluye una variedad de alimentos con proteínas como mariscos, carnes magras, huevos, legumbres (frijoles y chícharos), productos de soya, nueces y semillas. Contiene poca cantidad de grasas saturadas, grasas transaturadas, colesterol, sal (sodio) y azúcares agregados. Se mantiene dentro de sus necesidades calóricas diarias.

¿Para qué consumir fruta?

Las frutas frescas, congeladas o enlatadas son excelentes opciones. Pruebe frutas más allá de las manzanas y los plátanos como el mango, la piña o el kiwi. Cuando la fruta fresca no esté en temporada, pruebe una variedad congelada, enlatada o seca. Tenga en cuenta que las frutas secas y enlatadas pueden contener azúcares o jarabes añadidos. Elija variedades de frutas enlatadas en agua o en su propio jugo.

¿Para qué consumir vegetales?

Agregue variedad a las verduras asadas o al vapor con una hierba como el romero. También puede sofreír verduras en un sartén con un poco de aceite. Puede probar verduras congeladas o enlatadas para un acompañamiento rápido, simplemente cocínelos en el microondas y sirva. Busque verduras enlatadas sin sal, mantequilla o salsas cremosas agregadas. Para variar, pruebe una verdura nueva cada semana.

¿Para qué consumir Carne?

Si su receta favorita requiere freír pescado o pollo empanizado, pruebe variaciones más saludables horneando o asado a la parrilla. Tal vez incluso pruebe frijoles secos en lugar de carnes. Pregúntales a sus amigos, busca en internet y revistas recetas con menos calorías; ¡Te sorprenderá descubrir que tienes un nuevo plato favorito!

¿Qué son los Carbohidratos?

Bien, te cuento, los carbohidratos constituyen la principal fuente de energía de nuestra alimentación, son el combustible para desarrollar las actividades diarias, mantener la temperatura corporal y el buen funcionamiento de órganos vitales del cuerpo. Pueden ser simples y complejos. Los primeros se encuentran de manera natural en caña de azúcar, remolacha y miel o son añadidos en forma de azúcares, como por ejemplo a productos industrializados o de pastelería, mientras que los carbohidratos complejos están en alimentos como arroz, papa, yuca y plátanos entre otros. Se deben consumir de manera moderada, pues en exceso pueden ocasionar sobrepeso, obesidad, diabetes y otras

enfermedades crónicas; mientras que el bajo consumo produciría desnutrición o bajo peso.

¿Qué son y para qué sirven las Proteínas?

Claro, las proteínas son las responsables de la formación de células, tejidos y órganos. También están presentes en una gran cantidad de funciones del organismo. Durante los periodos de crecimiento, embarazo, lactancia y recuperación de enfermedades es necesario el consumo de alimentos ricos en proteína. Las fuentes de proteína pueden ser de origen animal como huevos, pescados, leche, carnes magras, pavo, y pollo; o de origen vegetal como las leguminosas (frijol, lenteja, garbanzo, chícharo), nueces y frutos secos, quinoa, entre otros.

¿Qué son y para qué sirven las Vitaminas?

Las vitaminas son un grupo de sustancias necesarias para el funcionamiento celular, el crecimiento y el desarrollo normal. Existen 13 vitaminas esenciales. Esto significa que estas vitaminas se requieren para que el cuerpo funcione apropiadamente. La Vitamina A: mantiene el funcionamiento del sistema inmune, la piel, los ojos y su visión, contribuye a la reparación de las vellosidades intestinales y el tejido pulmonar. Su consumo está relacionado con la prevención de infecciones, se encuentra en alimentos de origen animal como carne, hígado y vísceras, en frutas y verduras de color naranja y amarillo como calabaza, papaya, melón y tomate, entre otros. Vitaminas del complejo B: participan en la digestión y la absorción de carbohidratos. Son parte de este grupo la tiamina, la riboflavina, el niacina y el ácido fólico. Cuando hay bajos niveles en el organismo de esta vitamina, se afectan los glóbulos rojos, las células intestinales y se desarrolla anemia. Esta vitamina debe ser consumida en mayor cantidad durante el embarazo para evitar problemas de malformación del sistema nervioso en el feto. Las hortalizas de hojas verdes y oscuras, las leguminosas como frijol, lenteja y chícharo, así como frutas son fuente de ácido fólico. Vitamina C: ayuda en la producción y mantenimiento del tejido conectivo del cuerpo (huesos, dientes, piel, y tendones), en la cicatrización de heridas y tiene funciones antioxidantes. Todas las frutas y verduras son buena fuente de vitamina C.

Despedida

Me dió gusto conocerte, cuídate, espero haberte ayudado, que estés muy bien.

Palabra no encontrada

Lo siento, no puedo procesar eso puedes intentarlo de nuevo.

Para cada palabra, frase o pregunta que el usuario escribe, el chatbot tiene programada una respuesta determinada, esto implica que solo puede responder a las palabras o frases que tenga registradas.

4 Pruebas y Resultados

El funcionamiento es sencillo ya que se basa en una serie de patrones. Es decir, para que el chatbot nos dé una respuesta necesita tener almacenado en un diccionario palabras clave para así “entender” una palabra, frase o pregunta que el usuario escribe, para esto el chatbot tiene programada una respuesta determinada. Esto implica que solo puede responder a las palabras o frases que tenga registradas, en este caso, si el chatbot encuentra una palabra clave en una frase que el usuario introduzca, nos dará una respuesta.

El chatbot implementado tiene como objetivo que las personas puedan conocer más información sobre cómo llevar una vida saludable, tocando temas como: actividad física, comida saludable, sobrepeso y obesidad, bajar de peso, plan de alimentación saludable, tipos de colesterol, vitaminas, grasas, saber en qué nos ayudan las frutas y verduras, así como conocer acerca de los tipos de nutrientes, derivándose en micronutrientes y macronutrientes, entre otras de una forma muy sencilla.

Para llevar esto a cabo se ocupó información de “El decálogo para un buen estilo de vida” de la página oficial del IMSS; en la cual se abordan los siguientes temas: ingesta

diaria recomendada de diferentes grupos de alimentos, información nutricional, dieta saludable y estilo de vida [11] y también se utiliza información disponible del ministerio de salud de Colombia [12]. El chatbot está disponible en un sitio web cuya interfaz se muestra en la figura 1, y también se puede encontrar en telegram mostrada en la figura 2.

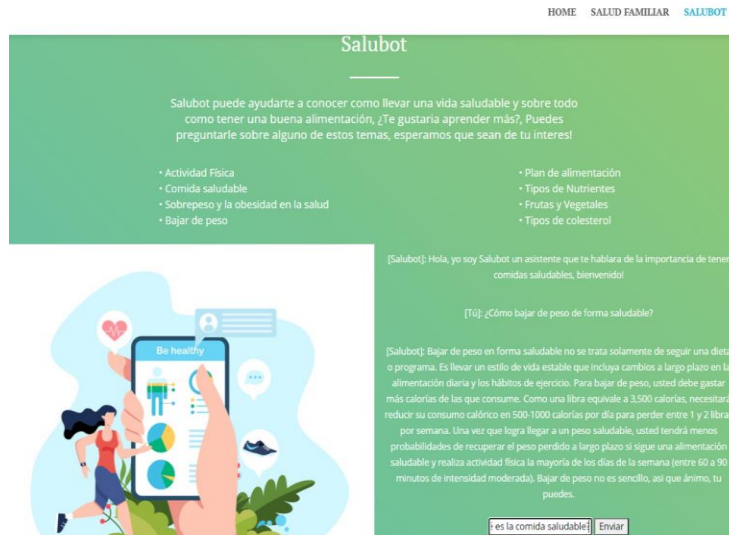


Fig. 1. Chatbot implementado en un sitio web.



Fig. 2. Chatbot implementado en telegram.

Se realizó una prueba de usabilidad, en donde se pide contestar lo siguiente:

Sexo:

Edad:

Conocimiento del uso de chatbots: Nada, bajo, medio, bueno, excelente.

¿Cuáles son las redes sociales que usas con mayor frecuencia?:

Tiempo empleado en minutos en el uso del chatbot para completar una tarea:

Cantidad de mensajes realizados:

Mensajes realizados con resultado satisfactorio:

Mensajes realizados con resultado no satisfactorio:

Mensajes realizados no comprendidos por el chatbot:

Precisión del chatbot: Mala, deficiente, regular, buena, excelente.

Satisfacción del chatbot: Mala, deficiente, regular, buena, excelente.

Facilidad de uso del chatbot: Mala, deficiente, regular, buena, excelente.

La prueba de usabilidad la contestaron 16 personas: 12 hombres y 4 mujeres, con una edad promedio de 23 años, 4 tienen un conocimiento excelente en el uso de chatbots y 12 un conocimiento medio, las redes sociales que usan son: Instagram, Facebook, WhatsApp y Facebook Messenger, el tiempo promedio en minutos para completar la tarea fue de 20/16 igual a 1.25 minutos, el promedio de mensajes realizados por cada usuario fue de 102/16 igual a 6.375, los promedios en el caso de mensajes con resultado satisfactorio fue de 72/16 igual a 4.5, el de resultados no satisfactorios fue de 32/16 igual a 2.0, el de mensajes no comprendidos por el chatbot fue de 28/16 igual a 1.75. En lo que se refiere a precisión y satisfacción 12 lo evalúan como bien y 4 como regular, finalmente en lo que se refiere a facilidad 8 usuarios lo evalúan como bueno y 8 como regular. Tomando en cuenta los resultados de la prueba de usabilidad, y mapeando malo a 1, deficiente a 3, regular a 6, bueno a 8 y excelente a 10, el chatbot tiene una calificación de 7.5 en lo que se refiere a precisión y satisfacción, en facilidad obtiene una calificación de 6.375.

5 Conclusiones y Trabajo Futuro

El trabajo desarrollado es funcional, permite llevar a cabo conversaciones sobre hábitos saludables, se tienen dos versiones una que corre en sitio web y otra en la red social telegram. Como trabajo futuro, se busca implementar un desarrollo profesional con ayuda de nutriólogos para poder tener mejor información, retroalimentando al chatbot sobre comida saludable y poder dar un asesoramiento para diferentes tipos de persona de acuerdo a sus necesidades, por ejemplo, hacer plan alimenticio de acuerdo a rangos de edad sabiendo si busca aumentar o disminuir de peso de una forma saludable. Los resultados de la prueba de usabilidad sugieren que el chatbot en cuanto a la interacción con los usuarios requiere ser más amigable.

Referencias

1. Adamopoulou E. and Moussiades L., An Overview of Chatbot Technology, Springer Nature Switzerland, pp. 373–383, (2020)
2. Athota L., Shukla V., Pandey N. and Rana A., Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), pp. 619-622, (2020)
3. Bulla, Ch., Parushetti, Ch., Teli, A., Aski, S., Koppad, S.: Review of AI Based Medical Assistant Chatbot, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 9, Issue 2, pp. 103-106 (2020)
4. Chou, T. L., Hsueh, Y. L.: A Task-oriented Chatbot Based on LSTM and Reinforcement Learning, Proceedings of the 2019 3rd International Conference on Natural Language Processing and Information, pp. 87-91 (2019)
5. Kedar A., Dahale J., Patel K., Lahamange S. and Chordiya S. G., Chatbot System for Healthcare using Artificial Intelligence, IJSDR | Volume 5 Issue 9, pp. 495-497, (2020)

6. Mehrjardi, M. S., Trablesi, A., Zaiane, O. Z.: Self-Attentional Models Application in Task-Oriented Dialogue Generation Systems, *Proceedings of Recent Advances in Natural Language Processing*, pp. 1031–1040 (2019)
7. Mohamed, M.M., Zhuopeng, W.: Artificial Intelligence Healthcare Chatbot System, *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 9, Issue 2, pp. 103-106, (2020)
8. Nichifor, E., Trifan, A. and Nechifor, E.M., Artificial Intelligence in Electronic Commerce: Basic, Chatbots and the Consumer Journey. *Amfiteatru Economic*, 23(56), pp. 87-101 (2021)
9. Yan, Z., Duan, N., Chen, P., Zhou, M., Zhou, J., Li, Z.: Building Task-Oriented Dialogue Systems for Online Shopping. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, pp 4616-4625 (2017)
10. IMSS, Nutrición
http://www.imss.gob.mx/sites/all/statics/salud/guias_salud/adolescentes/guiaadolesc_nutricion.pdf Fecha de acceso: Nov. 6, 2021
11. Ministerio de Salud y Protección Social , ¿Qué es una alimentación saludable?
<https://www.minsalud.gov.co/salud/publica/HS/Paginas/que-es-alimentacion-saludable.aspx> Fecha de acceso: Nov. 6, 2021

Sistema de Adquisición de Datos Vehiculares por medio del OBD-II con Interfaz Gráfica en Labview para un Posterior Uso en Aplicaciones de IoV (Internet de los Vehículos)

Nicolas Quiroz Hernández, Saul Luna Minor, Roberto Carlos Ambrosio Lázaro.
Benemérita Universidad Autónoma de Puebla,
Laboratorio SLED. Puebla, Pue. México
nicolas.quirozh@correo.buap.mx, saul.lunam@alumno.buap.mx,
roberto.ambrosio@correo.buap.mx

Resumen. Los avances tecnológicos en el ámbito automotriz referentes al internet de las cosas, ha permitido hacer un diagnóstico en línea, sobre las posibles fallas y desgaste de los sensores dentro del vehículo. Sin embargo un 60% del parque vehicular de México carece de lo anterior, por lo que se diseñó un sistema de adquisición de datos vehiculares, con interfaz gráfica en *Labview*, que permitirá en una primera etapa visualizar datos erróneos, interactuar con el automóvil por medio del *OBD-II* y posteriormente al integrarle al sistema presentado un módulo *4G/LTE*, enviar datos a la nube para que realice aplicaciones del *IoV*(Internet de los vehículos). Este sistema está especializado para implementar una comunicación CAN bus, integrando un microcontrolador ARM-Cortex M0, el cual posee: un controlador CAN (Controller Area Network), un transceptor CAN MCP2551 que permite acoplarse a los niveles físicos del bus, además de prevenir daños por cortocircuitos, un sensor de temperatura, un potenciómetro, pines de entrada-salida para conectar diversos sensores, actuadores o señales que se requieran trabajar, adicionalmente un regulador de voltaje para alimentar al microcontrolador. Esto fue desarrollado en el laboratorio SLED, ubicado en la Facultad de Ciencias de la Electrónica. Para la verificación del correcto funcionamiento, se realizaron dos pruebas, implementando una comunicación entre dos microcontroladores (ARM-Cortex M y PIC), y la otra comunicándose a un automóvil JETTA Hybrid, a través del sistema OBD-II, con el fin de obtener estos parámetros: RPMS, velocidad, presión y temperatura del automóvil, visualizándolos en una interfaz gráfica. Cuando se tenga implementada la comunicación con la nube, será de gran ayuda este interfaz porque ayudará a detectar fallas en la comunicación con el automóvil, errores en la transmisión de datos, complicaciones en la conexión al servidor, etc.

Palabras Clave. *CAN bus*, Redes Automotrices, *ARM Cortex-M*, *OBD-II*.

1 Introducción

El desarrollo acelerado de las Tecnologías de la Información y Comunicación (TIC), permite la conexión de diversos dispositivos a internet, a esto se le conoce como el Internet de las cosas [1,2] (*IoT*). En 2020 se estimaron alrededor de 20 mil millones de

dispositivos con *IoT* [2], de los cuales una gran porción pertenece a automóviles. Bajo este marco surge el concepto de Internet de los Vehículos *IoV* [2] por sus siglas en inglés), enfocándose en la recolección de datos vehiculares y su interacción con la inteligencia humana, las cosas, el entorno y las redes, para proporcionar múltiples servicios para las grandes ciudades [4].

Entre algunos ejemplos del uso de *IoV*, se encuentran las *Apple CarPlay* [2], que ofrecen una completa integración de los mapas de *Apple* dando seguimiento de voz para llegar al destino. Similarmente, *Google Android Auto* [2] provee una interfaz libre de distracciones que permite a los conductores control de funciones como GPS mapeo/navegación, reproducción de música, mensajes (*SMS*), telefonía y búsqueda en la web.

Actualmente los vehículos de alta gama tienen cientos de sensores, actuadores y Unidades de Control Electrónico (*ECUs*) en las redes automotrices [4]. Cada *ECU* realiza una función específica en el automóvil, como controlar los frenos, vigilar el nivel de combustible y la temperatura del motor, entre otros más, así el vehículo es seguro y fiable [5].

Para el 2030 se pronostica que al menos un 30 % de los vehículos serán autónomos en las ‘Mega ciudades’ [4], donde el conductor observa el comportamiento del automóvil detrás del volante o remotamente, reduciendo el tráfico con rutas Óptimas, mediante algoritmos que seleccionen la trayectoria más favorable en colaboración con mapas altamente precisos [5]. Es fundamental la integración de múltiples cámaras, sensores y actuadores que envíen datos con una baja latencia, además de recibir avisos sobre incidentes y situaciones de peligro que se pudiesen presentar en tiempo real [5]. Así que es indispensable una red inalámbrica que transmita y reciba grandes cantidades de datos en cortos periodos de tiempo [2].

Actualmente la red móvil *4G/LTE* [6,7], alcanza 100 Mbps. Las redes de quinta generación *5G* [8,9], tienen como objetivo llegar a velocidades superiores a 1 Gbps, sin embargo, esta carece de suficiente infraestructura [9]. A finales de 2018 se estimaron 3.16 billones de usuarios [6] *4G/LTE*, siendo preferible para aplicaciones de *IoV*.

1.1 Propósito y Enfoque del Proyecto en el Parque Vehicular en México

En el 2019 se registró un parque vehicular de 32.5 millones de unidades en México [13], de las cuales 20 millones tienen entre 6 y 25 años de antigüedad. Representando más del 60 % de los vehículos del país, siendo estos los que carecen de las nuevas tecnologías actuales. El *IoV* está enfocado principalmente en los automóviles híbridos y eléctricos, debido a que contaminan menos y predominarán en el 2030 [4], sin embargo, en el país se tiene un 2 % de unidades. Una ventaja del *IoV* es la notificación al conductor de posibles averías y desgastes al disponer del estado parcial del automóvil en tiempo real, permitiendo realizar un mantenimiento preventivo.

Sin embargo, en México al tener un parque vehicular en su mayoría carente de tecnologías para el diagnóstico en línea, el conductor debe llevar su unidad con personal especializado en reparación automotriz, donde las averías son detectadas por algún dispositivo *OBD-II*, realizando así un mantenimiento correctivo, que tiene desventajas, tales como: el vehículo no se supervisa después de la compra, por lo que los fallos son muy imprevisibles, desconociendo que autoparte se necesita para la reparación y en caso de no estar disponible en el paso se tendría que importar, incrementando costos y tiempo. Además, no se protege ni cuida la unidad, lo que reduce la vida útil de este.

Bajo la problemática antes mencionada este trabajo desarrollará un sistema empotrado que integre el *OBD-II*, proporcionando datos vehiculares para su visualización en una interfaz gráfica en *Labview*, siendo esta la primera etapa para posteriormente establecer

comunicación con la nube y siendo el enfoque principal que se abordará a lo largo del documento.

Una vez que se compruebe que los datos mostrados en la interfaz corresponden a los visualizados en el cuadro de instrumentación del vehículo, se procederá en una segunda etapa donde se integre un módulo *4G/LTE* para el envío de los datos hacia la nube, teniendo a disposición el estado parcial del vehículo para un diagnóstico en línea, mediante una aplicación intuitiva para un dispositivo portátil (*smartphone*) accesible para el usuario y con información aprovechable para personal del área automotriz. Posibilitando un mantenimiento preventivo que evite que se produzcan un fallo, notificando al conductor oportunamente y permitiendo que la unidad disponga de las ventajas existentes del *IoV* [2].

2 Estado del Arte

Tras una revisión sistemática de la literatura referente a los sistemas de adquisición de datos vehiculares con comunicación *4G/LTE*, se encontraron varios trabajos que contribuyen al desarrollo de la solución propuesta. En el artículo de Y. Zhou *et al* [14] se presenta un trabajo donde se utilizó un microcontrolador *ARM-Cortex TM-A8* y un módulo *4G/LTE Quectel EC20* producido por *Shanghai Mobile*, que soporta redes *GSM/GPRS*. En vez de usar una red *CAN bus* se plantea el uso del protocolo *Fley-Ray*, el cual tiene ventajas como gran ancho de banda, alta fiabilidad y rendimiento en tiempo real, sin embargo como el proyecto presentado va enfocado en automóviles con más de 6 años de antigüedad y estos no disponen de *FlexRay*, se descarta el uso de este protocolo.

Para enviar los datos a la Nube, utilizaron el protocolo *HTTP* debido a su alta velocidad y bajo índice de pérdida de datos, almacenados en una base de datos realizada en *MySQL*. Después fueron transferidos los datos vehiculares en formato *JSON* desde el servidor hacia una aplicación para el *smartphone* para ser visualizados. Para esto se utilizó el kit de herramientas *Gson* de *Google* para realizar la conversión y el análisis sintáctico de *JSON*.

En el trabajo revisado I. S. BULUT *et al* [15] se plantea una propuesta para detectar el tráfico en ciertas zonas de Istanbul, Turquía. En contraste con el trabajo anterior se utilizó el escáner comercial *OBD-II Renault Megane*, este envía por *Bluetooth* los datos vehiculares a un *smartphone*. Entonces mediante una aplicación diseñada en *Android* establece comunicación con la nube, usando el protocolo *HTTP* con formato *JSON*, muy similar al trabajo anterior.

3 Red CAN Bus

El *ECU* es un dispositivo que controla funciones específicas dentro del vehículo, como el motor, la dirección asistida, las ventanas eléctricas, asientos, puertas, bolsas de aire, sistema de frenado, entre otras más. Esta unidad tiene en general un sistema empujado con un firmware dedicado a cumplir una tarea en específico, la cual involucra comunicarse con otras *ECUs*.

Entonces el *CAN Bus* es una red que permite interconectar los diversos actuadores y sensores del vehículo, que están integrados en la diferentes *ECUs*. Importante no generalizar, ya que existe otros protocolos como *LIN bus* que se enfocan en partes del vehículo como limpiaparabrisas, activación de alguna alarma, entre otras, pero destaca la red *CAN* sobre las demás por comunicar sistemas como en el *ABS* (sistema enfocado a la seguridad y frenado del automóvil) donde no se permiten datos erróneos y se emplean

tiempos críticos. Añadido a esto tiene una compatibilidad electromagnética (por sus siglas en inglés *EMC*, teniendo la capacidad de funcionar sin complicaciones en un entorno electromagnético), dándole más robustez y tolerancia a fallas, considerandos una red prioritaria en el vehículo. Sus características por resaltar son:

1. Se conforma de un par trenzado de dos cables, donde *CAN-HIGH* es de color negro con naranja y *CAN-LOW* es marrón con naranja (figura 1). Sin embargo, cada compañía puede modificar a conveniencia la tonalidad. El que sea trenzado ayuda a reducir el efecto de *Crosstalk* [18].



Fig. 1. Cable típico *CAN Bus*.

2. La figura 2, se muestran los niveles de tensión respectivos al *CAN-HIGH* y *CAN-LOW*. Si llegase afectar una señal parásito, que consecuentemente hiciera variar el nivel de voltaje, ambas líneas serian afectadas de igual manera.

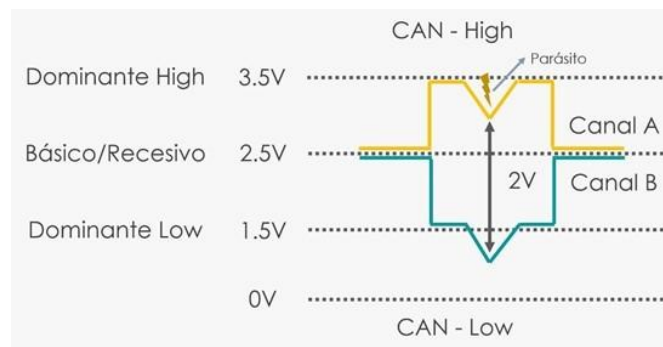


Fig. 2. Compensación de influencias parasitas.

- Una red *CAN Bus* está compuesta por dos resistencias de $120\ \Omega$ conectadas en paralelo. Los dispositivos conectados a sus extremos, se les denomina *nodos* (figura 3).

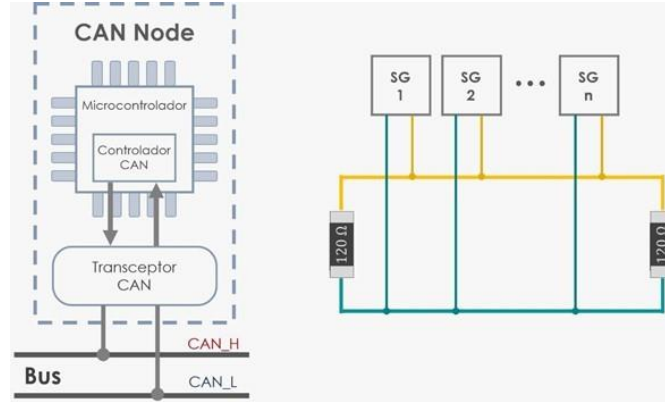


Fig. 3. Composición interna de un nodo (parte izquierda) y una red *CAN bus* (parte derecha).

Un *nodo* (figura 3) se conforma de un microcontrolador que integre internamente un controlador *CAN* ó acceda a uno externo, como puede ser el transceptor MCP2551 (figura 4), encargado de acoplar los niveles de tensión mostrados en la figura 2.



Fig. 4. Transceptor *CAN* MCP2551 superficial montado en una tarjeta para conexión al protoboard.

4 Microcontrolador ARM-Cortex

Se ha optado por usar un microcontrolador *ARM* [19, 20], ya que actualmente son utilizados en una gran variedad de aplicaciones automotrices, ordenadores, tablets, smartphones, etc. Este tiene una arquitectura RISC (del inglés *Reduced Instruction Set Computer*) de 32 bits, aceptando solo instrucciones compactas y ortogonales [19] que se puedan ejecutar en un único ciclo. Además de tener un bajo consumo de energía [20] y poseer varios periféricos, que posibilitan la integración de los elementos antes mencionados.

5 Sistema de Adquisición de Datos Vehiculares

El sistema de adquisición se encuentra en la primera versión en una tarjeta de desarrollo diseñada para que el usuario pueda crear una red *CAN* con facilidad, al disponer de borneras para la fácil conexión con un arnés eléctrico, además de probar todas las funciones correspondientes al protocolo. Se analizaron diferentes modelos de microcontroladores *ARM-Cortex*, se seleccionó el STM32F042C6T, por su reducido

tamaño, bajo costo, reducido consumo energía y se enlistan las características que posee (figura 5):

- 14 *ADC* con una resolución de 12 bits.
- 6 *Timers* configurables para *PWM*, comparación,
- Puerto dedicado para la programación del microcontrolador.
- 14 Pines de propósito general, para el uso libre del usuario.
- Comunicación Serial.
- Protocolo de comunicación *I2C* y *SPI*.
- Comunicación *CAN Bus* con borneras para fácil conexión y transceptor *CAN MCP2551*.
- Un potenciómetro que puede ser conectado a cualquier *ADC* del microcontrolador para hacer pruebas en la misma tarjeta.
- Un sensor de temperatura *LM35D* analógico conectado a un *ADC* para su prueba en la misma tarjeta.
- Tiene un cristal de 8 *MHz*, que puede alcanzar mediante el *PLL*, hasta los 48 *MHz*.



Fig. 5. Tarjeta de adquisición de datos v1.0 vista superior.

Este sistema utiliza el programador *ST-LINK V2*, el cual está incluido en la mayoría de tarjetas de desarrollo de *STMicroelectronics*, este puede programar diferentes familias de microcontroladores *ARM*. Al no integrar el programador en la tarjeta de adquisición, permite reducir costos y a su vez disminuir la contaminación generada.

6 Prueba Realizada en un Automóvil

Para corroborar el correcto funcionamiento del sistema, se hizo una prueba, en la que, se estableció comunicación con un auto *Jetta Hybrid*, por medio del *OBD-II* [17]. El software como el que se programó la tarjeta de desarrollo, fue *KEIL V5*, se escogió este debido a que es gratuito y permite usar el lenguaje C, como se muestra a continuación, el código configura el *CAN bus*, para recibir mensajes del automóvil con un identificador *0x7DE* [15], operando a una frecuencia de 500kHz.

Fragmento del código usado para configurar el periférico *CAN* en código C

```
MX_GPIO_Init();
```

```

MX_USART2_UART_Init();
MX_CAN1_Init();
/* USER CODE BEGIN 2 */
HAL_CAN_MspInit(&hcan1);
__HAL_RCC_CAN1_CLK_ENABLE();
HAL_CAN_Init(&hcan1);
sFilterConfig.FilterMode = CAN_FILTERMODE_IDMASK;

sFilterConfig.FilterScale = CAN_FILTERSCALE_32BIT;
sFilterConfig.FilterIdHigh = 0x0000; //0x0000
sFilterConfig.FilterIdLow = 0x0000; //0x0000
sFilterConfig.FilterMaskIdHigh = 0x0000; //0x0000
sFilterConfig.FilterMaskIdLow = 0x0000; //0x0000
sFilterConfig.FilterFIFOAssignment = CAN_FILTER_FIFO0;
sFilterConfig.FilterActivation = ENABLE;
sFilterConfig.FilterBank=0;
sFilterConfig.SlaveStartFilterBank = 14;
HAL_CAN_ConfigFilter(&hcan1, &sFilterConfig);
HAL_CAN_Start(&hcan1);
/* USER CODE END 2 */
/* Infinite loop */
/* USER CODE BEGIN WHILE */

TxHeader.StdId = 0x7DE; TxHeader.RTR = CAN_RTR_DATA;
TxHeader.IDE = CAN_ID_STD; TxHeader.DLC =8;
uint8_t txData[8] = { 'C', 'A', 'N', ' ', 'B', 'u', 's',
'2' } ;
HAL_StatusTypeDef HAL_RetVal;

if (HAL_CAN_ActivateNotification(&hcan1,
CAN_IT_RX_FIFO0_MSG_PENDING ) != HAL_OK)
{
Error_Handler();
}
/* USER CODE END 2 */
/* Infinite loop */
/* USER CODE BEGIN WHILE */
while (1)
{
/* USER CODE END WHILE */
/* HAL_RetVal = HAL_CAN_AddTxMessage(&hcan1, &TxHeader,
txData, (uint32_t*)CAN_TX_MAILBOX0);
HAL_GPIO_TogglePin(LD3_GPIO_Port, LD3_Pin);
HAL_Delay(400);*/

/* USER CODE BEGIN 3 */
}
/* USER CODE END 3 */

```

A través del osciloscopio, podemos observar la trama del *CAN Bus* transmitida desde el sistema de adquisición de datos vehiculares hacia el automóvil (figura 10). El identificador [10] es un parte de la trama del *CAN Bus* que sirve para etiquetar un nodo, por ejemplo, la mayoría de vehículos existentes tienen el identificador 0x7E8, el cual es un estándar para comunicarnos mediante *OBD-II* [15].



Fig. 11. Fragmento de trama enviada, donde el primero dato es 0x01, siendo este, el que procesa y representa el PIC en los leds.

Se obtuvieron los siguientes datos: *RPM*, velocidad, presión del aceite y temperatura del motor. Para esto el identificador del sistema de adquisición de datos debe ser 0x7DF y el del auto debe responder con el 0x7E8. En la figura 12, se observa en bloques, los dispositivos involucrados en esta prueba.

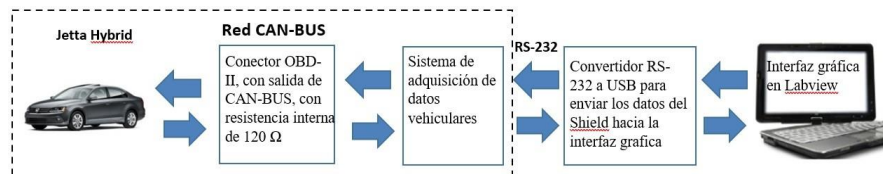


Fig. 12. Diagrama a bloques de los componentes usados para la segunda prueba.

El *OBD-II* especifica que valores se espera recibir y la forma de cómo tratar esta información, para después ser procesada y visualizada, por medio de algún software. Por practicidad se usó *LabVIEW*, donde se le realizó la interfaz gráfica de la figura 14.



Fig. 13. Pruebas realizadas al auto híbrido.

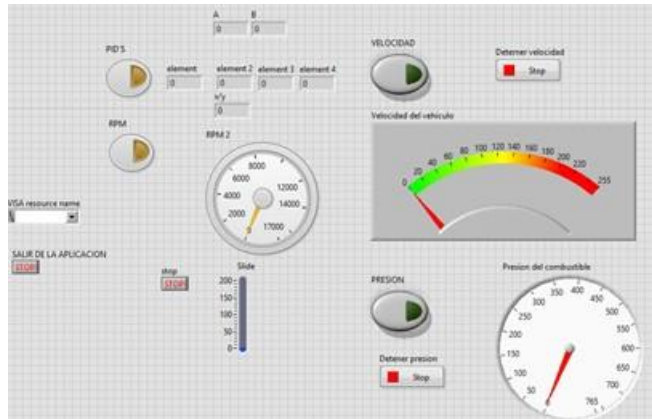


Fig. 14. Interfaz gráfica hecha en *labVIEW*.

7 Resultados

Considérese que algunos de los datos solicitados, se necesitan tratar antes de mostrarlos, por ejemplo para obtener las *RPM*, primero se manda el *PID*, el cual es un código que solicita el dato de las *RPM* y el vehículo responde con dos datos, A y B, entonces con la siguiente fórmula se obtiene el dato real, que debería ser casi igual al representado por el vehículo.

$$rpm = \frac{(256A + B)}{4}$$

Tabla 2. Comparativa de los datos solicitados con los que se visualizan en el cuadro de instrumentación.

Datos enviados desde el sistema de adquisición de datos vehiculares		Datos enviados desde el Jetta Hybrid		Dato registrado en la interfaz gráfica en Labview	Dato visualizado en el cuadro de instrumentación del Jetta Hybrid
Identificador	Datos Solicitado (PID)	Identificador	Datos enviados		

0 x 7 D F	) 0 x 0 C	0 x 7 E 8	A = 0 x 4 9 B = 0 x 5 4	4693 rpm	4967 rpm
0 x 7 D F	0 x 0 D	0 x 7 E 8	A = 0 x 2 0	21km/ h	19km/h
0 x 7 D F	0 x 0 F	0 x 7 E 8	A = 0 x 3 F	23°C	23°C

Hay cierta incertidumbre en los datos enviados y recibidos, puesto que, Labview recibe un dato digital y el visualizado en el *Jetta* es analógico (véase en la tabla 2), la variación obtenida depende del mecanismo de representación del auto, la cual le da un cierto offset, mientras que, la forma de mostrar los datos en Labview depende de la exactitud de como se realicen las operaciones matemáticas, las cuales son descriptas por la norma ISO-14229[17].

8 Trabajos a Futuro

Una vez validado los pasos anteriores, se tiene planeado que el sistema de adquisición de datos automotrices tenga modificaciones como se puede ver en la figura 15, mediante un conector *OBD-II* se acceda a la red *CAN* y a la alimentación del vehículo. Después se almacenen temporalmente estos datos, para posteriormente transmitirlos hacia el módulo *4G/LTE* que será el encargado de establecer comunicación con el servicio de nube en específico. En caso de fallar la conexión con la nube o la transmisión de los datos al módulo *4G/LTE* se propone almacenarlos en una memoria externa *microSD*, interactuando con esta, mediante la interfaz *SPI* [3].

Entonces usando la USART [4] para la transmisión-recepción de datos con el módulo 4G/LTE, almacenado y asignando un formato de geolocalización definido por el GPS que integra el módulo como se puede apreciar en la figura 15. El *SIM card slot*, es una ranura para introducir una tarjeta SIM que opere con las compañías móviles del país y así enviar los datos a la nube. Estas modificaciones se tienen planeadas para una interacción intuitiva con aplicaciones enfocadas en el *IoV*.

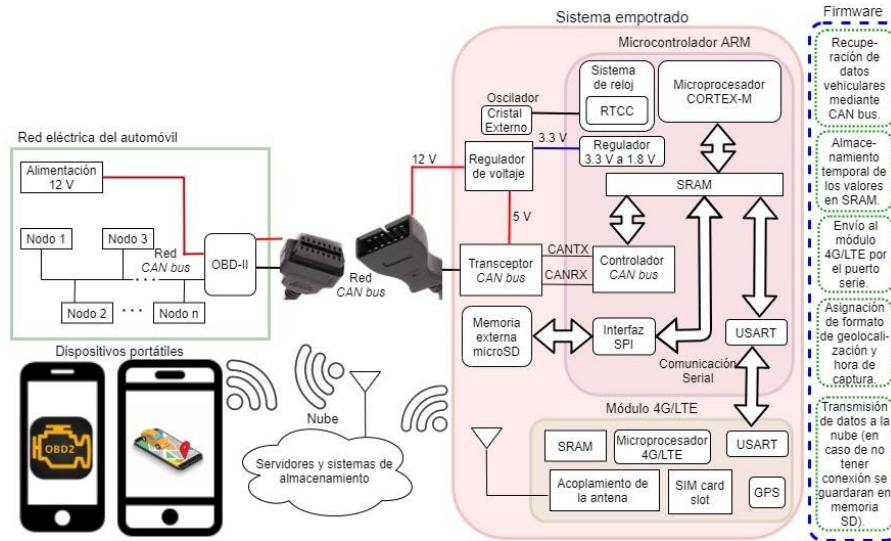


Fig. 15. Sistema completo con la adquisición de datos vehiculares para ser enviados a la nube.

9 Conclusiones

En base a las pruebas realizadas se considera que el sistema empujado es funcional, ya que permite la interacción con cualquier dispositivo que disponga de una red *CAN bus* e inclusive con cualquier tarjeta de desarrollo, sensor y actuador, que disponga de los protocolos ya mencionados (*I2C*, *SPI*, *RS-232* y *CAN*). Con el respaldo de la prueba realizada en el *Jetta Hybrid*, da garantía al sistema de realizar comunicación con los diversos *ECUs* que posee internamente un automóvil. Esta investigación sirve como base para validar la obtención de datos vehiculares y muestra una guía de como visualizarlos antes de enviarlos a un servidor para aplicaciones enfocadas al *IoV*. Como trabajo a posterior se plantea un sistema que fácilmente se conecte al *OBD-II* e integre el módulo 4G para una conexión con la nube y a su vez disponga de un conector *USB* o serial para visualizar por medio de la interfaz creada, datos erróneos, problemas de comunicación con la nube, etc.

Agradecimientos. Con la ayuda de la facultad de ciencias de la electrónica que apoyo con herramienta y instrumentos de medición para que se lograra la primera etapa de este proyecto, que como ya menciono es relacionado con la obtención de señales vehiculares.

Referencias

1. Li, X., Da Xu, L.: A review of internet of things—resource allocation. *IEEE Internet of Things Journal* (2020) 1–1
2. Yang, F., Wang, S., Li, J., Liu, Z., Sun, Q.: An overview of internet of vehicles. *China Communications* **11** (2014) 1–15
3. Signoretti, G., Silva, M., Dias, A., Silva, I., Silva, D., Ferrari, P.: Performance evaluation of an edge obd-ii device for industry 4.0. In: 2019 II Workshop on Metrology for Industry 4.0 and IoT (MetroInd4.0 IoT). (2019) 432–437
4. Winkelhake, U.: The digital transformation of the automotive industry : catalysts, roadmap, practice. Springer, Cham, Switzerland (2018)
5. Elshaer, A.M., Elrakaiby, M.M., Harb, M.E.: Autonomous car implementation based on can bus protocol for iot applications. In: 2018 13th International Conference on Computer Engineering and Systems (ICCES). (2018) 275–278
6. Moustafa, N., Hu, J. In: Security and Privacy in 4G/LTE Network. Springer International Publishing, Cham (2020) 1265–1271
7. Lin, Z., Huang, Y., Xie, Y.J., Wu, Q., Huang, X., Li, J., Liu, X.: Remote monitoring system of track cleaning vehicles based on 4g-network. In: 2020 IEEE International Conference on Networking, Sensing and Control (ICNSC). (2020) 1–5
8. Pyykonen, P., Lumiaho, A., Kuttila, M., Scholliers, J., Kakes, G.: V2x-supported automated driving in modern 4g networks. In: 2020 IEEE 16th International Conference on Intelligent Computer Communication and Processing (ICCP). (2020) 271–275
9. Vannithamby, R., Talwar, S. In: 5G. (2017) 34–51
10. Nugroho, S.A., Ariyanto, E., Rakhmatsyah, A.: Utilization of onboard diagnostic ii (obd-ii) on four wheel vehicles for car data recorder prototype. In: 2018 6th International Conference on Information and Communication Technology (ICoICT). (2018) 7–11
11. Gay, W. In: CAN Bus. Apress, Berkeley, CA (2018) 317–331
12. Nkenyereye, L., Jang, J.: Integration of big data for querying can bus data from connected car. In: 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN). (2017) 946–950
13. Oropeza, K.L.: ARIDRA. 32.5 millones de unidades registradas es el parque vehicular de México durante 2019. (2020)
14. Zhou, Y., Zhang, Y., Liu, M., Li, H., Gao, F.: Design of vehicle remote monitoring system based on 4g and flexray. In: 2018 IEEE 18th International Conference on Communication Technology (ICCT). (2018) 478–482
15. BULUT, I.S., ILHAN, H.: Cloud based vehicle and traffic information sharing application architecture for industry 4.0 (iot). In: 2019 International Conference on Information and Telecommunication Technologies and Radio Electronics (UkrMi- Co). (2019) 1–7
16. BinMasoud, A., Cheng, Q.: Design of an iot-based vehicle state monitoring system using raspberry pi. In: 2019 International Conference on Electrical Engineering Research Practice (ICEERP). (2019) 1–6
17. Srinivasan, A.: Iot cloud based real time automobile monitoring system. In: 2018 3rd IEEE International Conference on Intelligent Transportation Engineering (ICI- TE). (2018) 231–235
18. Chandrakasan, A., Brodersen, R.W. In: SRAM. (1998) 352–366
19. Gay, W. In: Introduction. Apress, Berkeley, CA (2018) 1–16
20. Prabhat, P., Labbe, B., Knight, G., Savanth, A., Svedas, J., Walker, M.J., Jeloka, S., Fan, P.M., Garcia-Redondo, F., Achuthan, T., Myers, J.: 27.2 m0n0: A performance-regulated 0.8-to-38mhz dvfs arm cortex-m33 simd mcu with 10nw sleep power. In: 2020 IEEE International Solid- State Circuits Conference - (ISSCC). (2020) 422–424

Comparativa de Herramientas para Realizar un Análisis a través del Cómputo Forense en CentOS 8.0

Ana C. Zenteno, Ma. del Carmen Santiago, Yeiny Romero, Judith Pérez, Gustavo T. Rubín, Miguel Ángel Pérez Xilo, María de los Ángeles Pájaro Cozatl

Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación, Ciudad Universitaria, 14 sur y Avenida San Claudio, Fraccionamiento Jardines de San Manuel, CP. 72570, Puebla, Pue. México
{ana.zenteno, marycarmen.santiago, yeiny.romero, judith.perez, gustavo.rubin}@correo.buap.mx,
{miguel.perezxi, maria.pajaro}@alumno.buap.mx

Resumen. La informática forense trabaja en el campo de los delitos informáticos causados por un uso inadecuado e ilegal de las tecnologías de información y comunicación. Es importante mantenerse actualizados y prevenidos sobre los últimos avances en ciberseguridad, manteniendo así, total resguardo de nuestra información personal o información privada de empresas. La informática forense ayuda a prevenir ataques a los sistemas de información, así como recrear los hechos en caso de sufrir pérdida de información, dando el origen y la forma en que sucedió, y los daños causados. A su vez esto nos permite actualizar o recrear las políticas de seguridad implementadas en nuestros sistemas. El objetivo de este artículo es compartir información acerca de las herramientas que se utilizan en software libre y como se realiza un informe forense al trabajar en casos de esta índole.

Palabras clave: Informática Forense, Hacking, Ciberseguridad, Recuperación de Datos, Análisis de Datos.

1 Introducción

En el transcurso de los años, el crecimiento de Internet y la necesidad de comunicarse a lo largo y ancho del mundo convirtieron a los sistemas computacionales en un elemento indispensable para el desarrollo de toda empresa u organización. Sin importar el tipo o tamaño de la empresa, existe un elemento de vital importancia para toda empresa, llamado “activo de información”, el cual representa digitalmente cualquier elemento de información que tenga un valor para dicha organización, así sea tan “irrelevante” o “sensible”, la información es poder, y por lo tanto se debe garantizar a toda costa la seguridad de estos datos [1].

La informática forense es el uso de métodos y técnicas científicas probadas, con el propósito de identificar, preservar, validar, analizar, interpretar, documentar y preservar evidencia digital, con el fin de facilitar la reconstrucción de hechos en una investigación legal o ayudar a anticipar o prevenir acciones ilegales en contra de la información digital [2].

1.1 Usos y aplicaciones

Probar que los sistemas de seguridad implementados en los sistemas son los adecuados y brindan bastante seguridad. [3]. En el caso de que el sistema de seguridad haya sido burlado, con la informática forense se consigue realizar un rastro y detectar no sólo cómo se ha producido el ataque, sino que también se puede conocer el nivel de daño [4]

Para llevar a cabo esta disciplina se debe seguir un conjunto de métodos o procedimientos que de manera objetiva y precisa produzcan resultados repetibles y verificables en una investigación para ejecutarlos de manera sistemática, el objetivo de esto es garantizar las políticas de seguridad y la protección de información, así como de las tecnologías que facilitan la gestión de esa información, y poder detectar cualquier clase de evidencia de vulnerabilidad que puedan tener y poder prevenir la pérdida de estos datos o sufrir algún ataque a los dispositivos y brindar una pronta solución. [4] [5].

1.2 Análisis y descubrimiento de evidencia

El análisis comienza al detectar un tipo de ataque informático o la sospecha de manipulación no autorizada de información para lo cual se debe previamente haber realizado la recuperación de los datos del dispositivo o sistema. Dispositivos como disquetes, cd rom, dvd rom, flash drive, USB, discos duros, etc., pueden ser analizados con las diferentes herramientas existentes en el ámbito de la informática forense con el fin de obtener los datos que se encuentren almacenados en estos y que contribuyan a la recreación de los hechos, así como encontrar pistas, cada una de estas herramientas se dedica a cumplir funciones diferentes y en algunos casos son herramientas completas que puede llevar a cabo varios de los procesos indispensables en un trabajo de esta índole [1].

1.3 Recuperación de datos

Este proceso se realiza para acceder y extraer la información contenida en medios de almacenamiento digital, que por motivos de daño o avería no pueden ser accesibles de forma normal, así como aquellos dispositivos que fueron dañados de manera maliciosa al recibir ataques por personas externas [8].

Discos duros, memorias USB, cámaras digitales, Cd o DVD, entre otros, son al menos unos de los muchos dispositivos a los que se pueden analizar para obtener información que nos ayude a recrear los hechos de cómo se produjo el delito [8].

1.4 Informe pericial y su estructura

Siempre se debe realizar el análisis de la información obtenida a una copia exacta creada desde la extracción de los datos del dispositivo o sistema, con el fin de no alterar dichos datos ya que, de ser así, estos no serían aceptados al llevarlos a un juicio. Tan importante es realizar este paso, que de manera correcta una vez realizada la copia de la información los dispositivos deben quedar inalterables, en desuso y en total resguardo para evitar en el peor de los casos la pérdida total de la información que estos contienen, lo cual comprometería por completo la investigación [2].

2 Metodología

2.1 Herramientas utilizadas en la informática forense

Se realizaron pruebas de funcionamiento de 2 herramientas utilizadas en la informática forense, en este caso se usará Volatility, instalada en una máquina virtual utilizando VirtualBox, con un sistema operativo CentOS 8.

Tabla 1. Herramienta Volatility. Características

Nombre	Versión	Funciones	Plataformas
Volatility	2.6	Leer el buffer del teclado de la memoria. Extraer el contenido del portapapeles de Windows. Imprimir lista de conexiones abiertas. Mostrar el árbol de dispositivos. Asociar los objetos del controlador a los módulos del kernel. Extraer archivos de memoria mapeados y almacenados en cache. Imprimir detalles en los ganchos de eventos de Windows. Mostrar los nombres de los servicios en el registro y devolver el SID calculado. Volcar la memoria direccionable para un proceso. Pool scanner para objetos de proceso. Imprimir lista de sockets abiertos.	Windows Linux Mac OS

Para poder analizar una imagen forense de una memoria RAM de manera correcta podemos ejecutar el comando “*volatility --info | more*” el cual nos proporcionara todos los perfiles de sistema que soporta y así en el siguiente paso podremos seleccionar el indicado con respecto al sistema operativo al que pertenece el volcado de memoria que descargamos o hayamos creado. Ver Figura 1.

```

root@localhost:/home/luis/Descargas
Archivo Editar Ver Buscar Terminal Ayuda
[root@localhost Descargas]# volatility --info | more
Volatility Foundation Volatility Framework 2.6

Profiles
-----
VistaSP0x64      - A Profile for Windows Vista SP0 x64
VistaSP0x86      - A Profile for Windows Vista SP0 x86
VistaSP1x64      - A Profile for Windows Vista SP1 x64
VistaSP1x86      - A Profile for Windows Vista SP1 x86
VistaSP2x64      - A Profile for Windows Vista SP2 x64
VistaSP2x86      - A Profile for Windows Vista SP2 x86
Win10x64         - A Profile for Windows 10 x64
Win10x64_10586   - A Profile for Windows 10 x64 (10.0.10586.306 / 2016-04-23)
Win10x64_14393   - A Profile for Windows 10 x64 (10.0.14393.0 / 2016-07-16)
Win10x86         - A Profile for Windows 10 x86
Win10x86_10586   - A Profile for Windows 10 x86 (10.0.10586.420 / 2016-05-28)
Win10x86_14393   - A Profile for Windows 10 x86 (10.0.14393.0 / 2016-07-16)
Win2003SP0x86    - A Profile for Windows 2003 SP0 x86
Win2003SP1x64    - A Profile for Windows 2003 SP1 x64
Win2003SP1x86    - A Profile for Windows 2003 SP1 x86
Win2003SP2x64    - A Profile for Windows 2003 SP2 x64
  
```

Fig. 1. De manera enlistada se muestran los perfiles con los que Volatility cuenta para hacer uso de ellos a la hora de analizar los volcados de memoria de diferentes sistemas operativos y estructuras.

Podemos utilizar el plugin **pslist** con una variante que proporciona un resultado más visual **pstree** el cual también se encarga de mostrar los procesos, pero en este caso de manera jerárquica colocando con cierta sangría a los procesos que descienden o dependen de otros como se muestra en la Figura 2.

```

root@localhost:/home/luis/Descargas
[root@localhost Descargas]# volatility -f cridex.vmem --dtb=0x2fe000 --kdbg=0x80545ae0 --kpcr=0xffdff000 --profile=WinXPSP2x86 pstree
Volatility Foundation Volatility Framework 2.6
-----
Name                               Pid  PPid  Thds  Hnds  Time
-----
0x823c89c8: System                   4      0    53    240  1970-01-01 00:00:00 UTC+0000
. 0x822f1020: smss.exe                368     4     3     19  2012-07-22 02:42:31 UTC+0000
.. 0x82298709: winlogon.exe           608    368    23    519  2012-07-22 02:42:32 UTC+0000
... 0x81e2ab38: services.exe          652    608    16    243  2012-07-22 02:42:32 UTC+0000
.... 0x821dfda0: svchost.exe          1056    652     5     60  2012-07-22 02:42:33 UTC+0000
..... 0x81eb17b8: spoolsv.exe          1512    652    14    113  2012-07-22 02:42:36 UTC+0000
..... 0x81e29ab8: svchost.exe           908    652     9    226  2012-07-22 02:42:33 UTC+0000
..... 0x823001d0: svchost.exe          1004    652    64   1118  2012-07-22 02:42:33 UTC+0000
..... 0x8205bda0: wuauclt.exe          1588   1004     5    132  2012-07-22 02:44:01 UTC+0000
..... 0x821fcd0: wuauclt.exe           1136   1004     8    173  2012-07-22 02:43:46 UTC+0000
..... 0x82311360: svchost.exe           824    652    20    194  2012-07-22 02:42:33 UTC+0000
..... 0x820e8da0: alg.exe               788    652     7    104  2012-07-22 02:43:01 UTC+0000
..... 0x82295650: svchost.exe          1220    652    15    197  2012-07-22 02:42:35 UTC+0000
... 0x81e2a3b0: lsass.exe              664    608     4     330  2012-07-22 02:42:32 UTC+0000
.. 0x822a0590: csrss.exe               584    368     9     326  2012-07-22 02:42:32 UTC+0000
. 0x821dea70: explorer.exe            1484   1464    17    415  2012-07-22 02:42:36 UTC+0000
. 0x81e7bda0: reader_sl.exe           1640   1484     5     39  2012-07-22 02:42:36 UTC+0000
[root@localhost Descargas]#

```

Fig. 2. Resultado ofrecido por el plugin **pstree** el cual los enlista de manera jerárquica para una mejor identificación de prioridad en cuanto a la atención que debe recibir.

2.2 Análisis de los sockets abiertos al momento del volcado.

Es posible visualizar los sockets relacionados a sus respectivos procesos, los cuales estaban abiertos al momento de realizar el volcado o al momento en el que se sufrió un ataque o infección a la máquina. Para lo cual utilizaremos el mismo comando anterior, pero con el respectivo plugin, en este caso **sockscan** y también podemos utilizar el comando **connscan** que también permite visualizar datos de los sockets, pero nos brindan diferente información entre estos 2 plugins (La Figura 3 muestran el uso de estos 2 plugins por separado y sus resultados).

```

root@localhost:/home/luis/Descargas
[root@localhost Descargas]# volatility -f cridex.vmem --dtb=0x2fe000 --kdbg=0x80545ae0 --kpcr=0xffdff000 --profile=WinXPSP2x86 sockscan
Volatility Foundation Volatility Framework 2.6
-----
Offset(P)  PID  Port  Proto  Protocol  Address  Create Time
-----
0x01fd7618 1220 1900  17  UDP  172.16.112.128  2012-07-22 02:43:01 UTC+0000
0x01fdb780 664  500  17  UDP  0.0.0.0  2012-07-22 02:42:53 UTC+0000
0x0203f460  4  138  17  UDP  172.16.112.128  2012-07-22 02:42:38 UTC+0000
0x02076620 1004 123  17  UDP  127.0.0.1  2012-07-22 02:43:01 UTC+0000
0x020c23b0 908  135  6  TCP  0.0.0.0  2012-07-22 02:42:33 UTC+0000
0x02325610 788 1028  6  TCP  127.0.0.1  2012-07-22 02:43:01 UTC+0000
0x02372808 664  0  255  Reserved  0.0.0.0  2012-07-22 02:42:53 UTC+0000
0x02372c50 664 4500  17  UDP  0.0.0.0  2012-07-22 02:42:53 UTC+0000
0x0239cc08  4  445  6  TCP  0.0.0.0  2012-07-22 02:42:31 UTC+0000
0x023f0630 1004 123  17  UDP  172.16.112.128  2012-07-22 02:43:01 UTC+0000
0x023f0d00  4  445  17  UDP  0.0.0.0  2012-07-22 02:42:31 UTC+0000
0x02440d08 1484 1038  6  TCP  0.0.0.0  2012-07-22 02:44:45 UTC+0000
0x02476878  4  139  6  TCP  172.16.112.128  2012-07-22 02:42:38 UTC+0000
0x02477460  4  137  17  UDP  172.16.112.128  2012-07-22 02:42:38 UTC+0000
0x024cd2b0 1220 1900  17  UDP  127.0.0.1  2012-07-22 02:43:01 UTC+0000
[root@localhost Descargas]#

```

Fig. 3. Aplicación del plugin **sockscan** para obtener información de los sockets abiertos al momento de la infección o infiltración en el sistema.

2.3 Uso de Scripts.

En caso de que logremos armar una secuencia de 5 o más procesos relacionados y deseemos realizar esta búsqueda a todos, podemos crear un script que se encargue de

realizar esta búsqueda a todos y cada uno de los archivos binarios de los procesos que indiquemos (La Figura. 4 muestra un sencillo script con la misma estructura recién ejecutada para la búsqueda de palabras, así como la Figura. 5 muestra los resultados del script).

```

root@localhost:/home/luis/Descargas/dump
GNU nano 2.9.8 busqueda.sh
for file in $(ls *.dmp); do echo $file; strings $file | grep bank | sort | uniq;done

```

Fig. 4. El script de búsqueda

```

root@localhost:/home/luis/Descargas/dump
[root@localhost dump]# nano busqueda.sh
[root@localhost dump]# ./busqueda.sh
1484.dmp
*accessbankplc.com*
*access.usbank.com*
*achieveaccess.citizensbank.com*
*achieveaccess.citizensbank.com/exchange/*
*alphabank.com.cy*
*.associatedbank.com*
*associatedbank.com*
*azaniabank.co.tz*
*bankaustria.at*
*bankbyweb*
*banking.calbanktrust.com*
*banking.calbanktrust.com/*
*banking.firsttennessee.biz*

```

Figura 5. Resultados del script de búsqueda

Tabla 2. Foremost es la segunda herramienta con la que se va a trabajar para logran un trabajo más completo. Se muestran las funciones y las plataformas donde lo podemos emplear.

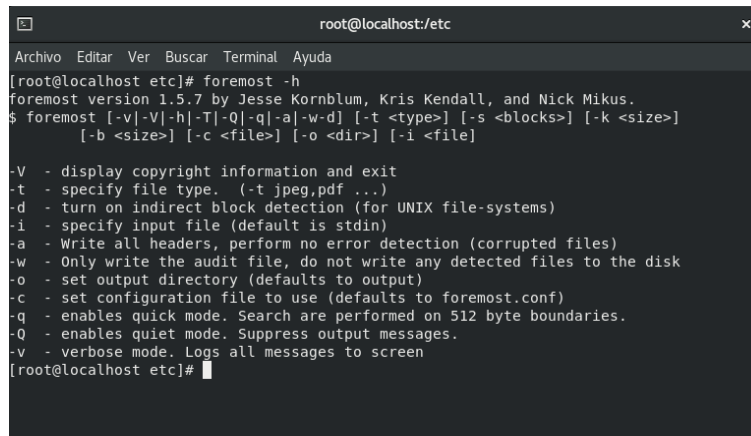
Función(es)	Plataforma(s)
Recuperar archivos utilizando sus encabezados, pies de pagina y estructuras de datos, mediante un proceso llamado <i>escultura de datos</i> o <i>tallado de datos</i> . Leer y copiar secciones directamente desde la unidad en la memoria del equipo. Recuperar datos directamente de discos duros. Recuperar datos de iPhones.	Windows, Linux, Mac OS

2.4 Foremost.

Otra herramienta forense de la que haremos uso en conjunto con Volatility debido a que deseamos implementar un proceso llamado “*tallado de datos*” el cual consiste en analizar los archivos que se utilizaron dentro de los procesos que categorizamos como sospechosos, esta herramienta trabaja recuperando archivos basados en sus encabezados, pies de páginas y estructuras de datos internas, La Figura 11 ejemplifica el uso de comandos importantes el cual nos proporciona una importante cantidad de imágenes de memoria que podemos utilizar para realizar pruebas, debido a que esta herramienta está

presente en casi todos los distros de **Linux**. Esta herramienta puede recuperar datos de tipo: l. [10].

Las pruebas nos permiten hallar algún archivo de cualquier formato que se haya descargado y este haya sido la causa de la infección de la máquina, ya que el proceso que estamos analizando es un navegador llamado **explorer** y un ejecutable de Adobe Acrobat llamado **reader.sl**. ver Figura. 6



```
root@localhost:/etc
Archivo Editar Ver Buscar Terminal Ayuda
[root@localhost etc]# foremost -h
foremost version 1.5.7 by Jesse Kornblum, Kris Kendall, and Nick Mikus.
$ foremost [-v|-V|-h|-T|-Q|-q|-a|-w|-d] [-t <type>] [-s <blocks>] [-k <size>]
[-b <size>] [-c <file>] [-o <dir>] [-i <file>]

-V - display copyright information and exit
-t - specify file type. (-t jpeg,pdf ...)
-d - turn on indirect block detection (for UNIX file-systems)
-i - specify input file (default is stdin)
-a - Write all headers, perform no error detection (corrupted files)
-w - Only write the audit file, do not write any detected files to the disk
-o - set output directory (defaults to output)
-c - set configuration file to use (defaults to foremost.conf)
-q - enables quick mode. Search are performed on 512 byte boundaries.
-Q - enables quiet mode. Suppress output messages.
-v - verbose mode. Logs all messages to screen
[root@localhost etc]#
```

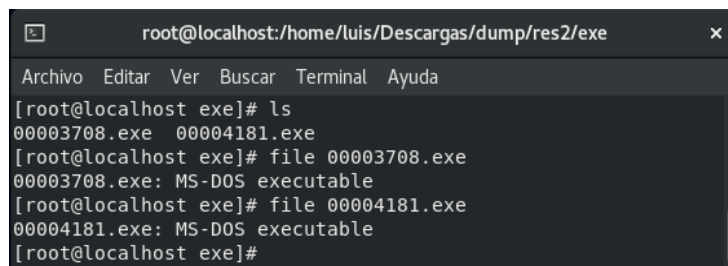
Figura 6. Herramienta Foremost y sus comandos que se pueden emplear.

3 Resultados

Una vez elaborada la parte de investigación, análisis y metodología procedemos a mostrar los resultados obtenidos con cada una de las herramientas dado que se busca generar concientización sobre el problema relacionado con la seguridad de la información analizada, para poder reparar o reforzar el área por donde ocurrió el ataque y de esta manera prevenir que ocurra algo similar en un futuro cercano.

3.1 Análisis de archivos sospechosos y verificación de sus extensiones.

Para realizar un análisis de archivos que pueda contener un malware, primero debemos verificar que el tipo de archivo sea de la extensión que está indicando, para lo cual existe un comando que nos permite realizar esta acción: *file* seguido del nombre del archivo y este nos devolverá los datos del encabezado que pertenece a ese archivo (La Figura 7 muestra un ejemplo de su ejecución).



```
root@localhost:/home/luis/Descargas/dump/res2/exe
Archivo Editar Ver Buscar Terminal Ayuda
[root@localhost exe]# ls
00003708.exe 00004181.exe
[root@localhost exe]# file 00003708.exe
00003708.exe: MS-DOS executable
[root@localhost exe]# file 00004181.exe
00004181.exe: MS-DOS executable
[root@localhost exe]#
```

Fig. 7. Ejecución del comando file para probar herramientas

Para continuar con el análisis podemos hacer uso de un servicio gratuito llamado *jotti's malware scan* con su sitio web: <https://virusscan.jotti.org/> el cual permite subir hasta 5 archivos al mismo tiempo con un límite de 250 MB cada uno, y realiza un análisis con diferentes motores de antivirus para informarnos si existe malware. Vea Figura 8 muestra el uso de este servicio y los resultados obtenidos

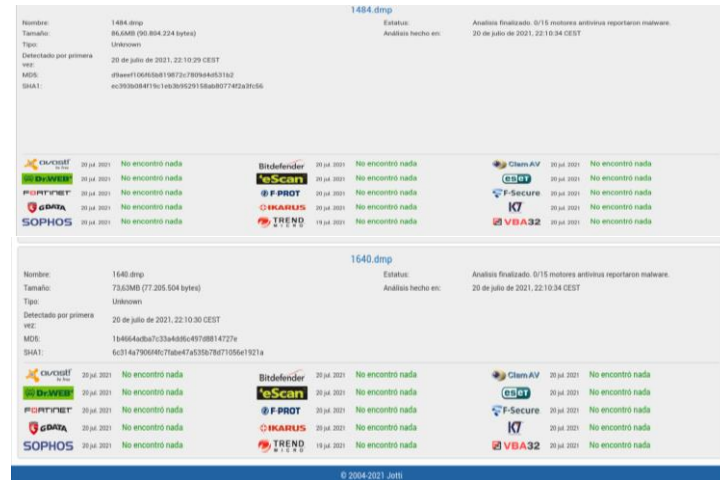


Figura 8. Uso de la herramienta *jotti's malware scan*

4 Conclusiones

Las dos herramientas utilizadas en esta investigación se complementan una con la otra, debido a que con Volatility logramos obtener todo lo que contiene un volcado de memoria RAM y es manejado con sus diferentes plugins, una vez obtenidos los archivos que necesitamos hacemos uso la herramienta Foremost para realizar la búsqueda de archivos que posiblemente contengan malware y con esto determinamos la causa de que un dispositivo o sistema sufrieran una infección.

Se obtuvieron resultados exitosos al analizar los archivos que son útiles al cómputo forense. Por otro lado obtuvimos conocimiento de herramientas que brindan resultados con mayor detalle en cuanto al análisis de los archivos, estas herramientas nos informan acerca de toda la estructura de un archivo y nos indica estrictamente que tipo de malware contiene, son herramientas que se llevan un tiempo de instalación y configuración para lograr hacer un uso adecuado de ellas, sin embargo, se continuará con la experimentación del uso de estas herramientas dado que en el computo forense son buenas herramientas de software libre que aún nos quedan descubrir más de sus bondades en el área la hora de trabajar con estos casos.

Referencias

1. Flore Flores I, Vargas Peña L. Revista Investigación y Tecnología. Informática forense (2016). Facultad de Ingeniería en Ciencias de la Computación y Telecomunicaciones Universidad Autónoma Gabriel René Moreno Santa Cruz – Bolivia. Recuperado el 23 de junio del 2021, sitio web: http://www.revistasbolivianas.org.bo/scielo.php?script=sci_arttext&pid=S2306-05222016000100015&lng=pt&nrm=iso&tlng=es
2. Daniel Esteves. (2019). ¿Qué es la informática forense? Recuperado el 17 de junio del 2021 de Platzi, sitio web: <https://platzi.com/blog/que-es-un-informatico-forense/>

3. SELECT BUSINESS SCHOOL (2020), La importancia de la informática forense. Recuperado el 23 de junio del 2021, de select business school, sitio web <https://escuelaselect.com/informatica-forense/>
4. SUFIJOR, La importancia de la informática forense (2019). Servicios integrales jurídico-forense. Recuperado el 23 de junio del 2021 de <https://www.sijufor.org/informacioacuten-relevante-en-materia-forense/la-importancia-de-la-informatica-forense>
5. Cañone Adrian, Palma Jose. (2018). REDES. Informática forense (6-7). Sitio web: <https://issuu.com/adriancañone63/docs/revista-de-informatica-forense>
6. Meneses Obando O. (2019). Informática forense desde el recurso humano y tecnológico, en las instituciones judiciales que cuentan con el servicio especializado de peritaje informático en Colombia. Universidad Externado de Colombia. Recuperado de <https://bdigital.uexternado.edu.co/bitstream/handle/001/1696/GGAAA-spa-2019-Informatica-forense-desde-el-recurso-humano-y-tecnologico-en-las-instituciones-judicial-es;jsessionid=E37A3460DA8D23D9DAB0734254601C75?sequence=1>
7. Martínez Raúl. (2020). 21 mejores herramientas de informática forense gratuitas. Recuperado el 24 de junio del 2021, de Noticias de seguridad informática, sitio web <https://noticiasseguridad.com/importantes/21-mejores-herramientas-de-informatica-forense-gratuitas/> .
8. Perito Oficial Group. Recuperación de datos de medios informáticos – informática forense. (2021), Recuperado el 23 de junio del 2021, sitio web: <https://peritojudicial.com/recuperacion-de-datos/#Qu%E9%20es%20la%20recuperaci%C3%B3n%20de%20datos>
9. Mesa Elneser, A. M. Vásquez Santamaría, J. E. y Lalinde Pulido, J. G. (2013). Aproximación a la informática forense y el derecho informático: ámbito colombiano. Universidad Católica Luis Amigó. Recuperado de <https://elibro.net/es/ereader/itchilpancingo/127464?page=121>.
10. Solvetic Seguridad. (2019). Cómo usar Foremost Linux y recuperar archivos borrados. Recuperado el 11 de julio del 2021, sitio web: <https://www.solvetic.com/tutoriales/article/7900-como-usar-foremost-linux-y-recuperar-archivos-borrados/>

Sobre Sistemas de Lógica Difusa para la Predicción de Riesgo de Padecer de Diabetes

Luis E. Morales-Márquez, Maya Carrillo-Ruiz, Pedro García-Juárez
Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la
Computación, Ciudad Universitaria. Blvd. Valsequillo y Av. San Claudio, Col.
Jardines de San Manuel, CP 72570, Puebla, Pue. México
{luise.morales, maya.carrilloruiz, pedro.garcia}@viep.com.mx

Resumen. La diabetes es un problema de salud en México y a nivel mundial. Ciertos factores de riesgo son considerados al diagnosticarla, y pueden ser de utilidad para predecir el riesgo de padecimiento. Diversas investigaciones presentan sistemas de predicción de diabetes utilizando lógica difusa, que posteriormente son optimizados por técnicas heurísticas. En este trabajo se enumeran los resultados y métodos utilizados por diferentes autores en la implementación de soluciones para este propósito. Además, con base en los artículos analizados, se presenta la propuesta de un esquema para un sistema difuso con el mismo fin que será implementado a futuro, y que buscará generar resultados comparables con las investigaciones presentadas y pueda ser utilizado por el público en general.

Palabras Clave: Lógica Difusa, Sistema Difuso, Optimización de Funciones de Membresía, Métodos Heurísticos, Diabetes, Riesgo de Padecimiento.

1 Introducción

Cuando se desarrolla cualquier sistema basado en lógica difusa, es de vital importancia tener una adecuada selección de funciones de membresía y los parámetros que las configuran, así como de las reglas difusas que rigen al sistema pues de lo contrario, el software arrojará resultados distantes a la realidad.

En el campo médico con el propósito de facilitar las tareas de decisión o clasificación al momento de tratar a un paciente, para prevenir problemas cardiacos y de diabetes, se han desarrollado diversos sistemas difusos [9,11]. Este trabajo muestra algunas aportaciones realizadas para la predicción de riesgo de padecimiento de diabetes, dichas investigaciones son analizadas y contrastadas para formular una propuesta de desarrollo de un sistema difuso con el fin de detectar pacientes que potencialmente puedan desarrollar esta enfermedad.

Según la Organización Mundial de la Salud (OMS), en su informe mundial sobre diabetes [1] se menciona que más del 33% de los niños y aproximadamente el 70% de los adultos en nuestro país padecen de sobrepeso u obesidad, factor importante en el desarrollo de la diabetes, además el Instituto Nacional de Salud Pública informa en la Encuesta Nacional de Salud y Nutrición [2] que 8.6 millones de personas mayores de 20 años han sido diagnosticadas con diabetes, correspondientes al 10.3% de este sector de la población.

Lo expuesto anteriormente, justifica el apoyarse en sistemas que permitan conocer el

riesgo de dicho padecimiento para llevar a cabo tareas de prevención. Cabe mencionar que muchos de estos sistemas utilizan lógica difusa en conjunto con múltiples algoritmos de optimización basados en el comportamiento de agentes naturales como enjambres o colonias de hormigas para su desarrollo.

A continuación, en la sección 2 se introducen algunos conceptos de lógica difusa, en la sección 3 se describen brevemente las heurísticas utilizadas en los trabajos analizados, la sección 4 presenta estos trabajos relacionados que se han analizado, posteriormente en la sección 5 se propone el esquema de un sistema con el mismo objetivo para implementación a futuro, finalmente en la sección 6 se muestran las conclusiones.

2 Lógica Difusa

La lógica tradicional nos orienta al uso de dos valores al momento de definir la pertenencia de objetos a determinado conjunto, sin embargo, este enfoque no siempre refleja la realidad que se describe. Para reducir las fallas derivadas de la aproximación clásica se introduce el concepto de conjunto difuso propuesto por Zadeh en 1965 [3].

Un conjunto difuso se entiende como un conjunto de pares ordenados, donde cada dupla está formada por el elemento y su valor de membresía en el conjunto. El valor de membresía de un elemento x en el conjunto A está dado por su *función de membresía* $f_A(x)$, el grado de pertenencia de x sólo puede estar en el intervalo $[0,1]$.

Las funciones de membresía más usadas son las triangulares, pero también pueden encontrarse de tipo gamma, lineales, sigmoidales, etc. El uso de ellas, para solucionar un problema específico, depende directamente del juicio de un experto en el tema. Como ejemplo, considere el mapeo de las funciones de membresía para la temperatura en grados centígrados, donde las variables lingüísticas o categorías a usar serán *frío*, *fresco*, *cálido* y *caliente*, y el universo discurso está definido entre 0° y 80° , la Figura 1, muestra el resultado de definir estas variables lingüísticas, recuerde que la selección de la forma de las funciones puede variar en función del propósito [4].

El proceso de asignación de valor de pertenencia de x a un conjunto A recibe el nombre de *fuzzyficación*, una vez capturado este valor, se inicia el proceso de *inferencia*, donde se aplican una serie de reglas de implicación lógica, provenientes del razonamiento deductivo empleando un sistema de razonamiento. Esta etapa generalmente utiliza más de una variable para obtener como respuesta la clasificación a la que debe pertenecer un objeto. La traducción de la salida a un valor cuantificable en lógica clásica, recibe el nombre de *defuzzyficación*.

3 Métodos Heurísticos

Los problemas de optimización tienen una diversa cantidad de soluciones potenciales, el objetivo es encontrar los valores adecuados para las variables de decisión involucradas que permitan obtener el máximo o mínimo valor de una *función objetivo*. Diversos

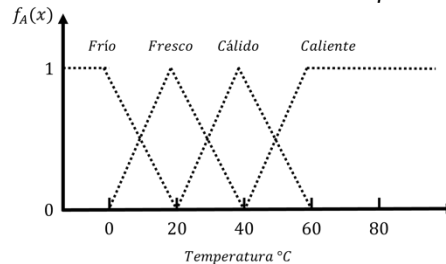


Fig. 1. Funciones de membresía para Temperatura.

problemas poseen una complejidad algorítmica elevada, pero aun así pueden solucionarse de manera eficiente. Para ello, se generan procedimientos que puedan encontrar las soluciones más adecuadas sin que necesariamente sean óptimas.

Un método exacto siempre genera una solución óptima, pero un método heurístico simplemente busca aportar una solución aproximada al problema sin comprometerse a ser la óptima, entre las principales razones para utilizar técnicas heurísticas se encuentran: a) No existe ningún método exacto para encontrar la solución al problema, b) Existe un método exacto, pero computacionalmente es muy costoso, c) El problema no exige de la solución exacta, y d) Se incluyen condiciones difíciles de modelar.

Los métodos heurísticos son utilizados en una amplia gama de problemas y tienen como ventajas tener una solución potencialmente buena desde el punto de partida, poder usarse para solucionar solamente una parte del problema, además de ser sencillos de comprender [5]. A continuación, se describen algunos de ellos.

3.1 Búsqueda Cuckoo

Este algoritmo está basado en el comportamiento parasitario de las crías de pájaro Cuckoo, esta especie inserta sus huevos en nidos comunales o de otras especies de aves con el objetivo de que el huésped los críe, si este huésped detecta la intrusión, tirará el huevo del Cuckoo o simplemente abandonará el nido.

Para maximizar la posibilidad de eclosionar, los huevos de Cuckoo, suelen imitar el color de otras especies. Si eclosionan antes, instintivamente las crías parásito, arrojarán a otros huevos del nido y tratará de imitar su llamado.

La búsqueda se fundamenta en 3 reglas: 1) El Cuckoo elige un nido al azar y deja un solo huevo, 2) los huevos de mejor calidad o imitación darán pie a la siguiente generación, 3) El número de nidos siempre es fijo, y se elige una probabilidad entre 0 y 1 de ser descubierto por el anfitrión, si esto se cumple, se elige entre abandonar el nido y crear uno nuevo, o elegir un huevo a tirar. El proceso se repite de manera iterativa evaluando la calidad de las soluciones en el nido [6].

3.2 Optimización por enjambre de partículas

El algoritmo nace como una simulación del comportamiento social, en específico la coreografía de las bandadas de pájaros en busca de alimento. Originalmente se desarrolló para optimizar funciones bidimensionales, pero se ha extendido a otro tipo de problemas.

Cada partícula se mueve en el espacio de búsqueda de manera iterativa obedeciendo una ecuación de velocidad y el punto actual en donde se encuentra la partícula, cada partícula actualiza localmente la mejor solución que ha encontrado individualmente, además de compararla con la mejor solución de todo el grupo, si la supera, la comunica al resto de la comunidad, esta tarea afecta a la ecuación de movimiento por lo que el enjambre deberá reagruparse al rededor la mejor solución después de prolongadas iteraciones.

Los enjambres cumplen con cinco principios básicos [7]: a) Principio de proximidad, la población debe realizar cálculos simples sobre el espacio en que se encuentra en poco tiempo. b) Principio de calidad, la población debe poder responder a variaciones en la calidad del ambiente. c) Principio de diversas respuestas, la población no debe hacer su trabajo en canales excesivamente estrechos. d) Principio de estabilidad, la población no debe cambiar su forma de trabajo o comportamiento aun cuando haya variaciones en el ambiente y e) Principio de adaptabilidad, se rompe la regla 4 solamente si vale la pena en término de costo computacional.

3.3 Optimización por colonia de hormigas

Como su nombre lo indica, se basa en la inteligencia colectiva de los grupos de hormigas en búsqueda de fuentes de comida. La forma adecuada de pensar en el funcionamiento es modelar un problema en forma de un grafo a partir del cual se busca el camino más corto hacia un destino, al encontrarse con un obstáculo, la primera hormiga se encuentra en una disyuntiva para elegir una nueva ruta a tomar, cada ruta tiene en principio la misma probabilidad de ser elegida, así que aleatoriamente la hormiga dará el siguiente paso. Mientras hace su recorrido, la hormiga secreta feromona, elemento utilizado para indicar al resto del grupo el camino que ha tomado.

La feromona se disipa con el paso del tiempo, por lo que una ruta más larga demorará más en ser completada y cuando la hormiga termine, este camino contendrá poca feromona. Puesto que un camino más corto logra mantener la feromona por más tiempo, las hormigas siguientes potencialmente elegirán esta ruta al ser más atractiva, siguiendo este proceso, eventualmente la colonia completa tendrá un acuerdo mutuo del camino que deben seguir para hallar la fuente de alimento [8].

4 Estado del Arte

Después de múltiples éxitos obtenidos sobre control industrial, se decide orientar los sistemas difusos hacia el área de la salud para apoyar al diagnóstico de enfermedades, principalmente problemas cardíacos y de diabetes. Un primer ejemplo se da en el 2015 por parte de Jain y Raheja llamado *“Improving the Prediction Rate of Diabetes using Fuzzy Expert System”* [9]. Donde a partir de la Pima Indians Diabetes Database, un experto selecciona los seis datos más relevantes de los pacientes.

El mismo experto diseña funciones triangulares de membresía y la predicción de riesgo de padecimiento como salida del sistema, además de seis reglas que dan funcionamiento al sistema difuso construido con el método de inferencia de Mamdani. El sistema no es optimizado por alguna heurística, en su lugar se opta por normalizar los datos de entrada en el intervalo [0,1] y finalmente, se prueba la efectividad del sistema calculando el porcentaje de precisión respecto a la presentada en el corpus.

El sistema muestra un 87.2% de precisión, superior a otros sistemas hechos por computadora que alcanzan 85.03% de aciertos, incluso superando sistemas médicos de especialistas que oscilan entre 77.6% y 81.7%, lo que lo califica como un sistema considerablemente mejor que otros revisados y suficientemente confiable.

Por otra parte, en el mismo año, se presentan avances para el diagnóstico de enfermedades del corazón en *“Fuzzy Membership Function Generation using DMS-PSO for the Diagnosis of Heart Disease”* [10], Kumar, Chanda, Rafiqul, Kundu y Aminul utilizan diversas técnicas para el mejoramiento de su sistema. Partiendo de la base de datos de la UCI, reducen los 14 datos correspondientes a los pacientes a través de coeficientes de correlación entre parámetros, aquellos que muestren una relación mayor se seleccionan, además los datos faltantes se calculan usando medias aritméticas.

Se calculan todas las posibles reglas de inferencia que tengan sentido usando los datos para un total de 82 reglas válidas, mismas que se ponderan y se seleccionan aquellas con puntuación mayor, las funciones de membresía se diseñan triangulares por el equipo de desarrollo basándose en literatura especializada. Se aplica la técnica de Optimización por Enjambre de Partículas Multienjambre para optimizar el vértice superior de cada triángulo.

Se revisa la precisión respecto a la establecida en el corpus, obteniendo cerca del 85% de aciertos, resultado mejor que una prueba realizada con todos los parámetros y sin optimización la cual obtuvo 77%, otra comparativa se establece con todos los parámetros aplicando optimización llegando hasta 82%, finalmente con parámetros reducidos sin

optimización, se obtiene 75%. La especificidad y sensibilidad arrojada por el sistema propuesto obtiene 88% y 78% respectivamente, resultados satisfactorios.

En 2017 se publica “*Cuckoo Search Optimized Reduction and Fuzzy Logic Classifier for Heart Disease and Diabetes Prediction*”[11], realizado por Reddy y Khare. Este estudio pretende establecer una predicciones para problemas cardiacos y diabetes. Usando teoría de conjuntos aproximados y binarización de los datos del UCI, de 13 parámetros se eligen para ser normalizados en el intervalo [0,1] y optimizados usando Búsqueda Cuckoo.

Las funciones de membresía son diseño de un experto, sin embargo, no son optimizadas de ninguna manera. Para evaluar el desempeño, se revisa la precisión respecto al corpus, así como la sensibilidad y especificidad, la puntuación es de 91%, 94% y 90% respectivamente. Al establecer una comparación contra un sistema previo basado en búsqueda por luciérnagas o murciélagos, se nota que las tres métricas fueron mejores pues se supera el 68%, 79% y 84% de dicho estudio, y el 72.6%, 100% y 0% de un sistema ya existente en la industria, con otras bases de datos se percibe un rendimiento menor debido a la reducción de datos en más del 50% perdiendo información representativa.

La Optimización por Búsqueda Cuckoo se hace presente nuevamente en el 2017, Sahu, Verma y Reddy realizan una investigación llamada “*Diabetes Classification using Fuzzy Logic and Adaptive Cuckoo Search Optimization Techniques*”[12] donde se obtienen datos de un laboratorio local de patología, sin embargo no se especifican dichos datos, posteriormente se aplica Reducción por Preservación Local de Proyecciones y las reglas de inferencia se calculan utilizando Búsqueda por Colonia de Hormigas, este conjunto de reglas se optimiza a través de Búsqueda Cuckoo, los autores no especifican el número de reglas, además, se menciona es que las funciones de membresía adoptaron forma triangular provenientes de un experto que no son optimizadas.

Se evalúa usando precisión de los resultados respecto al corpus consultado, así como la sensibilidad y especificidad en 10 ejecuciones y se compara contra otra implementación del mismo sistema optimizando las reglas de inferencia mediante un híbrido de Búsqueda por luciérnagas y murciélagos. En los 3 criterios el sistema Cuckoo muestra un mejor rendimiento que la variante Luciérnaga-murciélago, sin embargo, los resultados en cada ejecución son inestables variando el porcentaje de precisión del sistema entre 64% y 74%.

En 2020, Bressan, Flami de Azevedo y Molina de Souza presentan “*A Fuzzy Approach for Diabetes Mellitus Type 2 Classification*”[13], que muestra un clasificador difuso para catalogar a un grupo de personas como pacientes que requieren atenciones básicas o que requieren atenciones especializadas. Se utilizan datos de pacientes proporcionados por el Unified Health System de Brasil.

Un experto elige los datos que considera esenciales y usando un árbol de decisión con algoritmo C4.5 se generan las reglas de inferencia. Este estudio genera dos sistemas, uno para cada grupo de pacientes a clasificar, las funciones de membresía se eligen triangular a través de un Sistema de Inferencia Neurodifuso Adaptativo.

Se simula 10000 pacientes para ser clasificados por el sistema, estos resultados se compararon contra un sistema experto basado en conocimiento de especialistas, la clasificación es consistente con la del sistema experto por lo que los resultados se consideran aceptables.

Finalmente, en 2020 Risqy, Cantika y Fitra llevan a cabo “*Fuzzy Tsukamoto Membership Function Optimization Using PSO to Predict Diabetes Mellitus Risk Level*”[14] para desarrollar un sistema basado en el método de inferencia de Tsukamoto con el objetivo calcular el riesgo de padecimiento de diabetes, el corpus utilizado solamente consta de dos pacientes reales, dichos datos constan de tres factores de riesgo, un experto diseña funciones de membresía triangulares cuyos vértices superiores se optimizan por Enjambre de Partículas, además, el experto proporciona 6 reglas de inferencia para el control del sistema.

Para comprobar el funcionamiento se obtiene el error cuadrático medio del porcentaje de riesgo predicho por el sistema y el porcentaje realizado por un estudio especializado, el sistema arroja un 0.012% de error, contra un 25.9% de error del sistema no optimizado.

Las propuestas anteriores se comparan en las Tablas 1 y 2.

5 Sistema Propuesto

En continuación al camino abierto por las anteriores propuestas, se pretende desarrollar un sistema difuso capaz de predecir el riesgo de padecimiento de diabetes de un paciente con base en datos clínicos, como edad, IMC o probabilidad de padecimiento según historial familiar, donde los parámetros de las funciones de membresía serán optimizados por más de una técnica heurística buscando maximizar el porcentaje de aciertos en las predicciones respecto a las mostradas en el corpus, además de establecer una comparación del rendimiento entre las heurísticas que se implementarán a través de métricas comparables con las de los trabajos desarrollados previamente.

Tabla 1. Comparación de los trabajos en el estado del arte considerando mecanismo de reducción de datos y métodos de obtención de reglas y funciones de membresía

	Tipo de reducción de datos	Método de obtención de reglas de inferencia	Método de obtención de funciones de membresía
Fuzzy Tsukamoto Membership Function Optimization Using PSO to Predict Diabetes Mellitus Risk Level	Ninguna	Diseñadas por un experto	Triangulares diseñadas por experto
A Fuzzy Approach for Diabetes Mellitus Type 2 Classification	Datos seleccionados por un experto	Obtenidas a través de árbol de decisión con algoritmo C4.5	Triangulares obtenidas por Sistema de Inferencia Neuro Difuso Adaptativo
Diabetes Classification using Fuzzy Logic and Adaptive Cuckoo Search Optimization Techniques	Reducción por Preservación Local de Proyecciones	Obtención por Colonia de Hormigas	Triangulares diseñadas por un experto.
Improving the Prediction Rate of Diabetes using Fuzzy Expert System	Datos seleccionados por un experto	Diseñadas por un experto	Triangulares diseñadas por un experto.
Cuckoo Search Optimized Reduction and Fuzzy Logic Classifier for Heart Disease and Diabetes Prediction	Reducción por Conjuntos aproximados	Diseñadas por un experto	Triangulares diseñadas por un experto
Fuzzy Membership Function Generation using DMS-PSO for the Diagnosis of Heart Disease	Cálculo de coeficiente de correlación usando R- Squared	Calculadas con todos los casos coherentes posibles	Triangulares calculadas con base en literatura consultada

Partiendo de la investigación realizada, se elegirán los factores de riesgo que se consideren más importantes y las reglas del sistema de razonamiento, se propondrá un sistema difuso sin ningún tipo de optimización para, posteriormente optimizarlo hasta

obtener la configuración que presente mejores resultados. La optimización consistirá en mejorar los parámetros de las funciones de membresía a través del uso de más de una técnica heurística tanto en las funciones de entrada como de salida, el sistema de razonamiento no será modificado, además, la forma de evaluación del sistema consiste en revisar la tasa de aciertos obtenidos respecto a la predicción que se encuentra en el corpus de prueba, así como la cantidad de falsos positivos y falsos negativos.

Una vez teniendo el sistema óptimo, se tomarán datos de otro corpus, cuyos datos serán preparados seleccionando los tributos requeridos y normalizando los valores de los mismos, estos datos se usarán como entrada en el sistema optimizado entregando el resultado de la predicción. La Figura 2 ilustra el esquema del sistema propuesto.

El problema de optimización abordado consiste en contrar los parámetros de las funciones de membresía para el sistema difuso que maximicen la cantidad de aciertos al momento de realizar predicciones con datos en el conjunto de prueba, liegras variaciones pueden afectar el resultado, de manera que la exploración se lleva acabo en un espacio de soluciones continuo.

Tabla 2. Comparación de los trabajos en el estado del arte considerando criterio de efectividad, factores de riesgo y optimización utilizada

	Criterios para valuación de efectividad	Cantidad de Factores de riesgo utilizados y corpus.	Tipo de optimización
Fuzzy Tsukamoto Membership Function Optimization Using PSO to Predict Diabetes Mellitus Risk Level	Error cuadrático medio entre el corpus y la predicción	3 atributos. Datos de dos pacientes reales.	Funciones de membresía: Optimización por Enjambre de Partículas
A Fuzzy Approach for Diabetes Mellitus Type 2 Classification	Precisión respecto al corpus	7 atributos Corpus del unified health system de brasil.	Ninguna
Diabetes Classification using Fuzzy Logic and Adaptive Cuckoo Search Optimization Techniques	Precisión en 10 corridas, sensibilidad y especificidad respecto a un sistema con otra heurística	Características no especificadas provenientes de un laboratorio patológico local.	Reglas de inferencia: Optimización por Búsqueda Cuckoo
Improving the Prediction Rate of Diabetes using Fuzzy Expert System	Precisión	6 atributos Corpus del Pima Indians Diabetes Database[15]	Datos de entrada: Normalización binaria 0,1
Cuckoo Search Optimized Reduction and Fuzzy Logic Classifier for Heart Disease and Diabetes Prediction	Precisión, ensibilidad y especificidad respecto a otros trabajos relacionados.	9 atributos Corpus del repositorio de UCI[16] y pacientes del Sree Diabetic Care Cente (India)	Datos de entrada: Normalización en el intervalo [0,1] Búsqueda Cuckoo
Fuzzy Membership Function Generation using DMS-PSO for the Diagnosis of Heart Disease	Precisión	5 atributos Corpus del repositorio de UCI[16]	Funciones de membresía: DM-PSO Reglas de inferencia: Ponderación y selección.

Si se tuvieran n variables lingüísticas, donde cada una de ellas tuviera m_i parámetros por optimizar, con $i = 1, \dots, n$, y además, se subdividiera el espacio de soluciones en únicamente p posibles elementos con p considerablemente grande, un intento de fuerza bruta requeriría de explorar una cantidad exponencial de p^{nm_i} soluciones, estas condiciones reflejan que la solución no puede ser encontrada en un tiempo polinomial, sin embargo, si que puede ser verificada rápidamente al evaluar un conjunto de datos, lo que lo categoriza como un problema NP-Completo[17].

Buscar la configuración exacta que reduzca los errores de predicción al mínimo sería muy costoso computacionalmente, además de que no se puede garantizar la existencia de una configuración libre de errores y la tarea modelar el problema como una expresión matemática implica un elevado grado de dificultad.

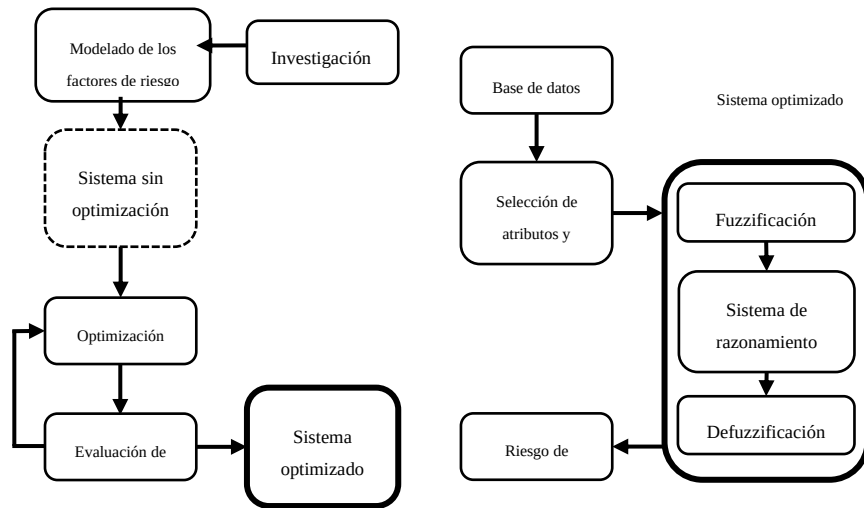


Fig. 2. Diagrama esquemático de la construcción y optimización del sistema propuesto, y diagrama esquemático del funcionamiento del sistema propuesto

Lo anterior justifica el uso de una técnica heurística para encontrar valores de los parámetros de las funciones de membresía pues con una aproximación adecuada logrará el objetivo sin necesidad de recurrir a un algoritmo de fuerza bruta y en un tiempo considerablemente reducido.

6 Conclusiones y Trabajo Futuro

Como hemos visto, los sistemas enfocados a la predicción de diabetes y otras enfermedades utilizando lógica difusa apoyada por técnicas de optimización heurística están vigentes en la actualidad y se ven respaldados por investigación en control difuso para múltiples propósitos en el pasado. Además, su efectividad en general es aceptable, si hablamos de la predicción de riesgo de enfermedades. Sin embargo, existen áreas de oportunidad en las que trabajos futuros podrían ayudar para mejorar la construcción de sistemas médicos y no solo sistemas de predicción.

Diferente selección de factores de riesgo y diferentes técnicas de selección generan resultados diferentes, pues cada investigación elige el conjunto de datos que le resulte adecuado, por lo que establecer una lista de los factores de riesgo definitivos no es posible aún. Como trabajo futuro se implementará la propuesta del sistema esquematizado en la Figura 2 y sus resultados se contrastarán con los trabajos resumidos en este artículo.

Referencias

1. Organización Mundial de la Salud.: Informe Mundial Sobre la Diabetes. P.39, (2016).
2. Instituto Nacional de Salud Pública.: Encuesta Nacional de Salud y Nutrición. p.18, (2018).
3. Zadeh, L. A.: Fuzzy Sets. In: Information and Control, pp. 8(3), 338-353, (1965).
4. Ross, T. J.: Fuzzy Logic with Engineering Applications, John Wiley & Sons, 3rd. Ed., (2010).
5. Martí, R.: Procedimientos Metaheurísticos en Optimización Combinatoria., Dpto. de Estadística e Investigación Operativa, Facultad de Matemáticas, Universidad de Valencia: <https://www.uv.es/rmarti/paper/docs/heur1.pdf>
6. Yang, X. S., Deb, S.: Cuckoo search via Lévy flights. In: World Congress on Nature & Biologically Inspired Computing (NaBIC), pp. 210-214, (2009).
7. Kennedy, J., Eberhart, R.: Particle Swarm Optimization. In: Proceedings of IEEE International Conference on Neural Networks, pp. 1942-1948, (1995).
8. Dorigo, M., Maniezzo, V. & Colomi, A.: Ant system: optimization by a colony of cooperating agents. In: IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), pp. (26) 29-41, (1996).
9. Jain, V. & Raheja, S.: Improving the Prediction Rate of Diabetes using Fuzzy Expert System. In: International Journal of Information Technology and Computer Science (IJITCS), p. 7(10), 84, (2015).
10. A. K. Paul, P. C. Shill, M. R. I. Rabin, A. Kundu & M. A. H. Akhand.: Fuzzy membership function generation using DMS-PSO for the diagnosis of heart disease. In: 18th International Conference on Computer and Information Technology (ICCIT), pp. 456-461, (2015).
11. Gadekallu, T. R., &Khare, N.: Cuckoo Search Optimized Reduction and Fuzzy Logic Classifier for Heart Disease and Diabetes Prediction. In: International Journal of Fuzzy System Applications (IJFSA), pp. 6(2), 25-42, (2017).
12. Sahu, N, Verma, T., Reddy, G.T.: Diabetes classification using fuzzy logic and adaptive cuckoo search optimization techniques. In: International Journal on Future Revolution in Computer Science & Communication Engineering (IJFRCSCE), pp. 3(9): 252-255, (2017).
13. Bressan, G. & Flámia Azevedo, B. & Souza, R.: A Fuzzy Approach for Diabetes Mellitus Type 2 Classification. In: Brazilian Archives of Biology and Technology. 63., (2017).
14. Risqy, F., Pradini, S., Cantika N. P. & Fitra A. B.: Fuzzy Tsukamoto Membership Function Optimization Using PSO to Predict Diabetes Mellitus Risk Level. In: Proceedings of SIET '20: 5th International Conference on Sustainable Information Engineering and Technology. pp. 101-106, (2020).
15. Pima Indians Diabetes Database, <https://www.kaggle.com/uciml/pima-indians-diabetes-database>, accesado el 28 de agosto del 2021.
16. UCI Machine Learning Repository, Diabetes Dataset, <https://archive.ics.uci.edu/ml/datasets/diabetes>, accesado el 28 de agosto del 2021.
17. Jiménez, H. & Lavalle, J.: Análisis y Diseño de Algoritmos: Traducción de partes del libro Fundamentals of Algorithmics de Brassard and Bratley, <http://aleteya.cs.buap.mx/~jlavalle/ada/adl.pdf>, accesado el 9 de septiembre del 2021.

Análisis de Herramientas para el Testeo de Vulnerabilidades de Aplicaciones Web

Ricardo E. García, Ana C. Zenteno, Ma. del Carmen Santiago, Yeiny Romero, Judith Pérez, Gustavo T. Rubín

Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación, Ciudad Universitaria, 14 sur y Avenida San Claudio, Fraccionamiento Jardines de San Manuel, CP. 72570, Puebla, Pue. México
ricardo.garciacru@alumno.buap.mx, {ana.zenteno, marycarmen.santiago, yeiny.romero, judith.perez, gustavo.rubin}@correo.buap.mx

Resumen. La realización de pruebas de penetración permite identificar las vulnerabilidades en aplicaciones web las cuales afectan directamente la confiabilidad que los usuarios. pues su información personal se puede ver gravemente comprometida. Con el fin de mitigar estas brechas en la seguridad se crean metodologías de desarrollo seguro. Así mismo, existen herramientas de testeo para detectar fallas, analizarlas y corregirlas antes de que los sitios salgan al mercado. En este artículo se realiza un análisis de herramientas para el testeo de vulnerabilidades de aplicaciones web, tales como Wikto y Nikto que se enfocan en analizar al servidor donde están alojadas las aplicaciones y otras como ZAP de OWASP que centran su trabajo tanto del lado de la aplicación como del servidor, se concluye que OWASP es una herramienta óptima que además cuenta con un organismo de prestigio que lo respalda.

Palabras Clave: Seguridad Informática, Vulnerabilidades, Sitios Web, XSS, OWASP, Herramientas de Testeo.

1 Introducción

La seguridad de la información describe actividades relativas a la protección de la información y los activos de las infraestructuras de la información contra riesgos de pérdida, uso inadecuado, revelación o daño. Los riesgos de estos activos pueden ser calculados mediante el análisis de amenazas a sus activos y vulnerabilidades. Las normas, modelos y estándares para la implementación de programas y controles apropiados para mitigar los riesgos, tales como ISO 27000, los modelos de ITIL, COBIT y estándares del NIST.

Muchas de los sitios que se visitan diariamente están expuestos al robo de información valiosa, intrusión en su privacidad. Cualquier interacción inocente o navegación desinformada por medio de las redes sociales, emails entre otros sitios puede resultar en la posibilidad de exponerse a un ataque cibernético.

Según Shahriar las aplicaciones web contienen vulnerabilidades, las cuales pueden ser brechas de seguridad tal como el robo de información confidencial, ruptura de la integridad de los datos o afectar la disponibilidad de la aplicación web [1]. Lo que lleva a desarrollar software de acuerdo a los estándares, muchos administradores de aplicaciones web están en constante presión para desarrollar aplicaciones y servicios web de forma rápida y que además cumplan con los requerimientos mínimos necesarios para proteger al usuario final.

Entre las vulnerabilidades que se encuentran presente en los sitios web con más frecuencia y que además presentan ataques con mayor frecuencia se listan en el Top Ten Project de OWASP, presentado cada 3 años de donde se destaca XSS, que es una vulnerabilidad bastante empleada y que se encuentra en la mayoría de las aplicaciones web. Esta le permite al atacante insertar un código malicioso dentro de las páginas que visitan los internautas. También resaltan los ataques de inyección SQL que son técnicas que emplean la inyección de código para explotar las vulnerabilidades de un determinado sitio web. Es utilizado comúnmente para atacar aplicaciones que son controladas únicamente por datos, donde se colocan sentencias de SQL en un campo de entrada para su ejecución. Por otro lado, se tiene Clickjacking, en esta vulnerabilidad se utilizan capas transparentes o en marca de agua para poder engañar al usuario que de clic en un determinado botón o enlace en otro sitio cuando pretendía hacer clic en la pagina de nivel superior. De esta forma es posible secuestrar, por así decirlo, los clic que están destinado a las página original y los enruta a otros sitios web que no son del mismo sitio web al que se quería acceder originalmente. Por último, el uso de componentes con vulnerabilidades conocidas, componentes como librerías, framework y otros módulos de software que se ejecutan con los mismos privilegios a los de la aplicación. Si alguna vulnerabilidad es encontrada en alguno de estos componentes se puede explotar hasta el punto de perder fácilmente la información valiosa de los usuarios o que el servicio se vea interrumpido [2].

Por esta razón, la mayoría de los arquitectos y desarrolladores de sitios web hacen uso de los certificados SSL que son un estándar de seguridad global que permite las transferencia de datos cifrados entre un navegador y un servidor web[3] con el fin de disminuir el riesgo de robo y manipulación de información confidencial (como números de tarjetas de crédito, nombres de usuario, contraseñas, correos electrónicos, etc.) por parte de hackers y ladrones de identidades.

“Hoy en día, más del 70% de los ataques en contra de los sitios web o aplicaciones web de alguna compañía proviene de la “Capa de aplicación” y no de la capa de red o de la capa de sistema” (Gartner Group). Para esto se plantea el uso de metodologías de desarrollo que tiene como objetivo el desarrollo de mejores aplicaciones, un mejor proceso de desarrollo que identifique salidas de cada fase de forma que se pueda planificar y controlar un proyecto y que guie hacia un proceso estándar en la organización.

Por ejemplo, Microsoft Trustworthy Computing SDL, una metodología que se encuentra vigente y en constante evolución, donde se proponen como objetivos principales reducir el número de defectos y/o errores de diseño y de código relacionados con la seguridad y reducir la severidad del impacto de cualquier efecto residual [4].

También se tiene a Comprehensive, Lightweight Application Security Process (CLASP), una metodología diseñada para aplicarse en cada fase de ciclo de vida metodologías de seguridad. En la cual se comprenden un conjunto de actividades enfocadas en la seguridad que pueden integrarse en cualquier proceso de desarrollo de software [4].

Y por último se tiene Open Web Application Security Project (OWASP), esta metodología tiene como objetivo ofrecer una metodología de libre acceso y utilización y que pueda ser utilizada como material de referencia para los arquitectos de software, desarrolladores, fabricantes y profesionales de la seguridad. Tiene como objetivos principales la definición de estándares sea, el estándar de verificación de seguridad en aplicaciones y el estándar de seguridad en la web y el desarrollo de herramientas útiles para identificar los fallos y corregirlos [5].

Además de las metodologías para el desarrollo seguro de software existen herramientas de análisis de vulnerabilidades que facilitan al pentester realizar ataques controlados a los sitios para exponer y explotar las vulnerabilidades de los sitios web, entre ella se encuentra Vega[6] que es una herramienta de análisis de vulnerabilidades open source desarrollada por Subgraph y una plataforma para testear la seguridad de

aplicaciones web. Funciona como proxy, interceptando las peticiones que se realizan durante la navegación y analizándolas para obtener información. Nikto[7] es una herramienta open source para Linux que escanea servidores web. Realiza pruebas exhaustivas sobre archivos potencialmente peligrosos, CGIs, comprueba versiones no actualizadas y problemas específicos de cada versión. También se tiene a Wikto [8] una versión de Nikto para Windows que añade algunas funcionalidades extras, por ejemplo, lógica difusa de comprobación de errores, minería de datos, minería de directorios asistida por Google y monitorización en tiempo real de petición/respuesta. Wapiti [9] es una herramienta para auditar la seguridad de las aplicaciones web. Realiza análisis de caja negra en busca de scripts y formularios para actuar como un fuzzer e inyectar payloads con el objetivo de encontrar vulnerabilidades.

Y por último en estas herramientas tenemos a OWASP Zed Attack Proxy Project [10] una herramienta open source para la realización de tests de penetración desarrollada por OWASP, que permite detectar vulnerabilidades en aplicaciones web. Dentro de sus funcionalidades destaca que funge como un proxy que intercepta todas las peticiones que se realizan entre el servidor y el navegador y las respuestas recibidas. También permite establecer puntos de ruptura para cambiar peticiones y respuesta sobre la marcha. De igual manera permite el análisis activo de vulnerabilidades, ZAP realiza ataques conocidos contra las aplicaciones web seleccionadas con el objetivo de encontrar vulnerabilidades potenciales. Incluso es capaz de realizar un análisis pasivo de vulnerabilidades, este tipo de análisis no modifica las respuestas y no ralentiza la exploración ya que se realiza en segundo plano. Del mismo modo puede detectar automáticamente nuevos recursos (spider) a partir de una lista de enlaces, esta herramienta identifica todos los hipervínculos de la página y accede a ellos repitiendo el proceso de forma recursiva.

2 Análisis de la Herramienta

2.1 Análisis de seguridad

Desde la perspectiva que existen diferentes herramientas con diferentes fines se debe conocer cuáles son las funcionalidades que deben cumplir las herramientas mencionadas para poder ser consideradas buenas. Esto puede cambiar dependiendo del fin con el que sean usadas.

Por lo tanto, una herramienta debe de ser capaz de crear un árbol del sitio que está siendo analizado, el árbol de direcciones debería contener tantos los ficheros y directorios bajo los cuales está construido también poder registrar los archivos a los cuales no se pueda acceder, toda la información necesaria para evaluar la seguridad del sitio. Esta funcionalidad es conocida como Spidering.

También debe de ser capaz de poder reemplazar peticiones, es decir, detener la ejecución o el envío de paquetes al servidor desde el navegador para cambiar la información enviada sin que se reenvíen al destino.

Del mismo modo, para detectar vulnerabilidades como SQL injection, Local/Remote File Inclusion, XML External entity, XSS, entre otros, la técnica de fuzzing se vuelve indispensable. Esta técnica consiste en proporcionar datos inválidos, inesperados o aleatorios en las entradas del servidor web y analizar su respuesta para estudiar cómo afectan las misma a la lógica implementada en el servicio.

Algunas de las funcionalidades mencionadas anteriormente se pueden resumir en dos tipos de escaneo, el escaneo pasivo y el escaneo activo. El escaneo pasivo analiza de forma pasiva las peticiones y respuestas que pasan a través del proxy utilizando un enfoque no intrusivo. Se consigue la información sin interactuar directamente con el objetivo mediante el uso de técnicas tales como la ingeniería social sniffing de red,

búsquedas por internet o vigilancia de instalaciones para recabar información sobre usuarios, accesos, infraestructura, etc. Y el reconocimiento activo emplea un enfoque intrusivo para intentar identificar cierto tipo de vulnerabilidades mediante el uso de ataques conocidos. Este tipo de escaneo puede generar errores como la denegación del servicio web, ejecución de código o cualquier otro tipo de problema no esperado. Tiene como propósito identificar objetivos directamente con el objeto al interactuar con los mismos y de esta manera identificar los que se encuentran activos, para luego practicar un análisis más profundo, dicho análisis profundo tiene como objetivo detectar las potenciales vulnerabilidades encontradas en las etapas anteriores que están a nivel de infraestructura como a nivel de aplicación para poder mitigarlas.

2.2 Análisis de herramientas

Se recopiló información sobre herramientas y sus características y así identificar el análisis de vulnerabilidades que identifica en el mapeo como se muestra en la Tabla 1 donde se muestran las más importantes de las aplicaciones web marcadas en el top ten de OWASP, ya sea para auditar una aplicaciones web o para que un arquitecto o desarrollador pueda comprobar su funcionalidad y robustez. Herramientas como Nikto[7] y Wikto[8] centran sus especificaciones en analizar el servidor en el cual las páginas web se encuentran alojadas ya que la parte más crítica de un sitio web es su servidor web pues el servidor web es el responsable de aceptar las solicitudes de los visitantes, comprenderlas y dar respuesta a dichas solicitudes.

Tabla 1. Análisis comparativo entre las diferentes herramientas que ayudan a identificar vulnerabilidades en los sitios web.

Vulnerabilidad/funcionalidad	Vega	Nikto	Wikto	Wapiti	OWASP ZAP
Spidering	✓	✓	✓		✓
Fuzzing					✓
Reemplazo de peticiones				✓	✓
Local/Remote File inclusion	✓				✓
XSS	✓			✓	✓
XML external entity	✓				✓
Escaneo Pasivo					✓
Escaneo Activo		✓	✓		✓
HTTP		✓	✓	✓	✓
HTTPS		✓	✓	✓	✓

Ahora bien, el reemplazo de peticiones consta de detener la ejecución o el envío de paquetes al servidor desde el navegador para cambiar la información enviada sin que se reenvíen al destino, de donde únicamente OWASP ZAP[10] y Wapiti[9] lo implementan. Por otro lado, LFI, Inclusion de archivos locales del inglés Local File Inclusion[14] es una técnica que consiste en incluir ficheros locales, es decir, archivos que se encuentran en el mismo servidor de la web con este tipo de fallo a diferencia de Remote File Inclusion[14] o inclusión de archivos remotos (RFI) que incluye archivos alojados en otros servidores. Esto se produce como consecuencia de un fallo en la programación de

la página, filtrando inadecuadamente lo que se incluye al usar funciones en PHP para incluir archivos.

Por otro lado el XSS[15], Cross Site Scripting, es una de las vulnerabilidades más explotadas en los sitios web, según OWASP Top Ten Project[2]. Los ataques XSS son un tipo de inyección en la cual un atacante logra ejecutar código en los navegadores de los usuarios que acceden a un sitio web legítimo. Por otro lado el XXE[16], XML External Entity, es una inyección de código en una aplicación que analiza datos XML. La inyección ocurre mediante un concepto llamado entidad que almacena cualquier tipo de dato. Esta entidad, funciona como el termino conocido en programación de variable. La entidad, puede también acceder al contenido local o remoto a través de declarar un identificador del sistema donde éste puede acceder a una URL cuando se procesa la entidad.

Mientras que el escaneo pasivo y escaneo activo son funcionalidades desarrolladas fuertemente por OWASP ZAP[10] que brindan información sobre los usuarios, accesos, infraestructura del sitio de forma automática.

Y por último los protocolos de comunicación que hace referencia al tipo de tráfico que se pueda intersectar desde las herramientas, siendo solo Vega[6] el único que no cuenta con esta funcionalidad. HTTP[17] es el acrónimo de Hypertext Transfer Protocol, es un protocolo de transferencia de hipertexto, es un protocolo de comunicación que permite la transferencia de información en internet y HTTPS[18], Hypertext Transfer Protocol Secure, utiliza una combinación de protocolos de comunicación (HTTP + SSL(Secure Sockets Layer (capa de sockets seguros))/TLS(Transport Layer Security)) que hace que cualquier tipo de información que se transmita en la red sea cifrada y nadie pueda acceder a ella, únicamente el navegador y el servidor web.

De igual manera en la Tabla 1 se muestran algunas de las características más buscadas a la hora de decidir qué herramienta de análisis de seguridad utilizar, ya sea para auditar un sitio web o para que un arquitecto o desarrollador pueda comprobar fallos en el sitio en cuestión. Herramientas como Nikto[7] y Wikto[8] centran sus especificaciones en analizar el servidor en el cual las páginas web se encuentran alojadas ya que la parte más crítica de un sitio web es su servidor web pues el servidor web es el responsable de aceptar las solicitudes de los visitantes, comprenderlas y dar respuesta a dichas solicitudes.

Por otra parte, hay herramientas que no solo centran sus funcionalidades en analizar el servidor, si no hacerlo desde el navegador por el cual está siendo visitado el sitio, enviando solicitudes al servidor y analizando las respuestas siendo el caso de OWASP ZAP[10] que cumple con las características fundamentales y mayormente utilizadas en la detección de brechas en la seguridad. Además de ser una herramienta respaldada por un organismo bastante importante en el mundo de la seguridad informática.

3 Resultados

Tal y como hemos podido comprobar en la Tabla 1. Una de las mejores herramientas que detectan vulnerabilidades en los sitios web es OWASP por ser una herramienta muy completa y que cuenta con las características necesarias un análisis completo. Además de ser una herramienta que está en constante actualización desarrollada por un organismo importante dentro del campo de la seguridad informática.

Referencias

1. Petukhov A.; Kozlov D.: Detecting security vulnerabilities in web application using dynamic Analysis with penetration testing. <https://owasp.org/www-pdf-archive/OWASP-AppSecEU08-Petukhov.pdf>. Accedido el 05 de Septiembre de 2021.
2. Open Web Application Security Project. Top ten Project. <https://owasp.org/www-project-top-ten/>. Accedido el 05 de Septiembre de 2021.
3. Todo lo que debe saber sobre certificados SSL. Obtenido de Todo lo que debe saber sobre certificados SSL: https://www.verisign.com/es_LA/website-presence/online/ssl-certificates/index.xhtml. Accedido el 05 de Septiembre de 2021
4. Castellaro M.; Romaniz S. C.; Pessolani P. A.: Hacía la ingeniería de software seguro. XV Congreso Argentino de Ciencias de la Computación, pp. 1257-1266 (2009)
5. OWASP: <https://owasp.org/>. Accedido el 05 de Septiembre de 2021.
6. We live security: Vega: <https://www.welivesecurity.com/la-es/2015/03/03/como-auditar-la-seguridad-sitio-web-vega/>. Accedido el 05 de Septiembre de 2021
7. Nikto: <https://cirt.net/nikto2>. Accedido el 05 de Septiembre de 2021
8. Wikto: <http://research.sensepost.com/tools/web/wikto>. Accedido el 05 de Septiembre de 2021
9. Wapiti: <http://wapiti.sourceforge.net/>. Accedido el 05 de Septiembre de 2021
10. OWASP ZAP: <https://owasp.org/www-project-zap/>. Accedido el 05 de Septiembre de 2021
11. Manual de instalación OWASP ZAP: https://drive.google.com/file/d/1omz_f6-e2RHwSrf-J785WVcd0yMBX-6D/view?usp=sharing.

Análisis de Comentarios de Cambio Climático en Twitter bajo Técnicas de Procesamiento de Lenguaje Natural

M. Beatriz Bernábe¹, Gerardo Martínez¹, Carmen Ceron¹, Alberto J.L. Carrillo²
Benemérita Universidad Autónoma de Puebla, ¹ Facultad de Ciencias de la Computación, ²Facultad de Filosofía y Letras, Puebla, Pue. México
beatriz.bernabe@gmail.com, gmartinez54@hotmail.com,
academicaceron@gmail.com, acarrillo_mx@yahoo.com

Resumen. Cuando se extraen grandes cantidades de información relacionada con comentarios en plataformas web como las redes sociales, el uso de herramientas en Procesamiento de Lenguaje Natural (PLN) produce resultados alentadores si se analizan problemas donde el lenguaje humano es el tema central. En particular, la percepción que los usuarios de Twitter tienen acerca del cambio climático, se ha analizado definiendo previamente diccionarios básicos para obtener información contenida en Tweets. Esta información identificada como opiniones, se ha extraído aleatoriamente utilizando un proceso de descargas con una API de Python. Con el fin de que la información sea clasificada en sentimientos positivos y negativos, el algoritmo Naive Bayes ha sido muy útil. En este punto, se identifica que la apreciación sobre el cambio climático de los usuarios de Twitter se encuentra sesgada hacia la opinión negativa, lo cual hace suponer que la influencia pesimista del tema climático incide en el juicio de los twittereros. Por tanto, a partir de los resultados obtenidos, se observó la necesidad de extrapolar el valor de las funciones de Twitter como herramienta de comunicación científica que ayude a informar a los usuarios y al público general para que perciban la verdadera magnitud del problema sobre el cambio climático.

Palabras Clave: Análisis de Sentimientos, Cambio Climático, PLN.

1 Introducción

Con la información disponible en la red, es posible desarrollar distintos análisis tanto de marketing como de tendencias o estimaciones. Entre las técnicas más populares para abordar problemas de esta magnitud, se encuentran aquellas enmarcadas dentro PLN, entendida como una rama de la Inteligencia Artificial que ayuda a las computadoras a entender, interpretar y manipular el lenguaje humano (natural). Con el afán de establecer la brecha entre la comunicación humana y el entendimiento de las computadoras, PLN toma elementos prestados de muchas disciplinas, incluyendo la ciencia de la computación y la lingüística computacional.

Una de las aplicaciones del PLN es el análisis de comentarios de Twitter respecto a un tema específico. En este artículo, se ha mirado el problema del cambio climático, el cual ha causado divergentes y contradictorias opiniones, en particular de aquellas personas que no están convencidas del cambio climático. En este contexto, los autores han establecido la conjetura siguiente: El análisis del pensamiento sobre el cambio climático en las redes sociales, puede ser influenciado asertivamente cuando se proponen programas o campañas de conciencia y a su vez diversificarse a otros medios.

En la literatura Sueca, pueden encontrarse reportes que se basan en el uso del marco de la teoría de las representaciones sociales y partir de ahí, se han propuesto análisis cuantitativos considerando diversos medios como lo es un periódico sensacionalista y

otro en noticias de televisión de servicio público, entonces bajo los resultados del análisis de sentimientos, las representaciones emocionales del cambio climático pueden mejorar la participación pública en el tema, pero también desviar la atención del cambio climático como el fenómeno abstracto [1].

El presente artículo comparado con la propuesta de Höijer [1], identificó que ambos comparten el análisis de las emociones sobre el cambio climático, entendiendo como emociones el miedo, esperanza, culpa, compasión o nostalgia, las cuales se bifurcan en sentimientos negativos y positivos, pero en este caso, nuestra contribución es en idioma español.

En particular, para el presente proyecto, primero se señalaron palabras claves del problema para proceder a descargas de Twitter usando hashtags. Una vez contando con información suficiente, se aplicó el algoritmo de clasificación Naive Bayes, el cual al apoyarse en el Teorema de Bayes con una suposición de independencia entre los predictores, genera la bifurcación en sentimientos positivos y negativos, además, Naive Bayes es fácil de construir y particularmente útil para conjuntos de datos muy grandes.

El trabajo se encuentra estructurado como sigue; la presente introducción como sección 1. La sección 2 se encarga de exponer aspectos alrededor del cambio climático. En la sección 3 se describe el diseño de un algoritmo de descargas y del análisis de sentimientos en Python para PLN. La sección 4 presenta los resultados y finalmente se describen las conclusiones.

2 Cambio Climático y Twitter

La percepción pública de la ciencia es un aspecto social abordado desde mediados del siglo XX a través de encuestas a la población, como son los Eurobarómetros en Europa, los cuestionarios de la National Science Foundation (NSF), etc. Los sondeos buscan establecer medidas sobre el interés, conocimiento y actitudes o críticas de las personas respecto a la ciencia [2][3][4]. En específico, la comunicación y/o divulgación de la ciencia persigue la ósmosis entre la ciencia y el público sobre distintos temas como el problema del cambio climático.

Actualmente, la comunicación científica intenta involucrar al público en el diálogo donde la tarea informativa resalta pragmáticamente en las redes sociales debido a la gran cantidad de usuarios activos, sin embargo, se advierte la necesidad de entendimiento sobre la incidencia e impacto de los nuevos entornos virtuales para estudiar la forma en que los debates públicos se mueven respecto a temas científicos [5][6]. Destaca en este punto que el acceso a los contenidos se produce sin mediadores y en consecuencia, el análisis en redes sociales puede constituir un complemento útil a las encuestas clásicas de percepción social de la ciencia. Más aún, ha sido de interés evaluar la implicación de los usuarios en discusiones abiertas debido a la facilidad de acceso a los contenidos y el tipo de información disgregada.

2.1 Twitter

En las últimas décadas, la comunicación se ha extendido en la web, principalmente en las redes sociales a través de distintas formas, por ejemplo, por opiniones y comentarios. En este contexto, los usuarios tienen acceso a diversidad de opiniones con la posibilidad de incidir en las ideas la gente, y a esta capacidad de influenciar sobre los pensamientos, sentimientos y acciones de los usuarios se le llama influencia social [7] [8].

En la dinámica de esta propuesta, el usuario se encuentra en un rol de emitir-recibir una opinión y/o sentimiento, e incluso, aquellos usuarios/consumidores de distinta

magnitud, toman decisiones sobre el consumo de productos y marcas a partir de la comunicación e información propagada por las redes [9].

Uno de los principales espacios de las redes sociales que consiente el intercambio de opiniones es Twitter, cuyo principal fin es el intercambio de mensajes de manera bidireccional por parte de los usuarios. Twitter es un servicio que permite mensajes de hasta 280 caracteres. Por otra parte, el número de usuarios activos en la red social ha experimentado un crecimiento considerable aunque en los últimos años parece más estable, de esta manera, se estima que Twitter se encuentra entre las redes sociales más importantes llegando a alcanzar un 24% de uso entre la población de los Estados Unidos en el año 2019 [10].

A pesar de la disponibilidad de una gran cantidad de papers alrededor de Twitter y PLN, poco existe sobre análisis de cambio climático con PLN.

Es aquí donde nuestra contribución es simple pero no por ello menos importante, los autores analizamos las opiniones de usuarios de Twitter sobre cambio climático proponiendo implícitamente un procedimiento sencillo para aplicarlo a otros escenarios [11][12].

En Twitter por ejemplo, se encuentran enlazados temas de carácter científico, lo cual se traduce en asunto de estudio para la investigación en comunicación pública de la ciencia [13]. También se ha observado que en Twitter ha sido útil para la difusión de temas, centrarse en recolectar seguidores y en segundo plano, las búsquedas de palabras clave para visibilizar contenidos científicos.

2.2 El Contexto del cambio climático en twitter

El público se ha pronunciado activamente en Twitter en asuntos académicos relacionados a la energía nuclear, el debate climático y la nanotecnología [14, 15]. Esta observación responde al análisis de información de una discusión polarizada donde se extraen tweets mediante la búsqueda de palabras clave con el fin de analizar las emociones de los usuarios. Por ejemplo, en el debate del cambio climático se ha observado de distintas maneras que los usuarios optan cada vez más las redes sociales como medio para expresarse [16, 17]; además existen estudios donde el análisis de sentimientos es una buena manera de investigar el fondo subyacente de la información extraída [1].

Una investigación en Twitter señaló tres tipos de comunidades entre los twitters: 1) los que perciben la emergencia climática, 2) los negacionistas y 3) los imparciales. En tal manuscrito se explica que los hashtags más utilizados sobre la crisis del clima se asocian con contenidos científicos, discusiones geográficas y con preocupaciones sociales o tecnológicas [18]. En consecuencia de la lectura, los autores de este artículo, abordamos un similar problema respecto a las opiniones del público sobre cambio climático en Twitter.

Es importante comenzar con una descripción plausible de cambio climático. En términos amplios, se define como la modificación del clima que ha tenido lugar respecto de su historial a escala regional y global. En general, se trata de cambios de orden natural, pero actualmente se les encuentra asociados con el impacto humano sobre el planeta. Se trata de un fenómeno complejo con numerosas variables como los ciclos del agua, los vientos solares y la posición de la Luna. Todas estas variables generan modificaciones sobre las condiciones atmosféricas que motivan la gran complejidad que caracteriza al clima de la Tierra. Esto explica de alguna manera las grandes dificultades para intentar definir el cambio climático. Por otra parte, el calentamiento global también interviene en las precipitaciones, nubosidad y muchos otros parámetros [19].

Existen distintas conjeturas sobre las causas del cambio climático, por ejemplo la cuenta de variaciones propias del Sol (vientos, "manchas solares", fenómenos meteorológicos propios, orbitales, el impacto de meteoritos, deriva continental,

composición atmosférica, corrientes oceánicas, campo magnético terrestre y efectos causados por el hombre, etc.). Se ha planteado que el planeta Tierra podría responder reforzando o moderando los efectos para recuperar un equilibrio natural. En cualquiera de estos escenarios, la mayor parte de los cambios observados tienen incidencia en la calidad de vida de las personas [18].

Si bien, el cambio climático no es sinónimo de calentamiento global dado que responde a distintas causas con variabilidad de implicaciones, no es de sorprender que cambio climático se halle relacionado al aumento de la temperatura promedio en la atmósfera y en los océanos, por tanto, no debe confundirse calentamiento global con cambio climático. El primero indica el aumento de la temperatura media de la Tierra, que es propiciado por el efecto invernadero, es decir, el que es originado por la emisión de gases contaminantes a la atmósfera. De este modo, tanto el calentamiento global como el efecto invernadero no son sino algunos de los factores que influyen en el cambio climático [19].

3 Desarrollo

Cuando se dispone de grandes cantidades de información que ha sido extraída para analizarla con distintos propósitos, es inevitable aplicar las correctas herramientas para conseguir el resultado esperado. Si la información obtenida son comentarios de la red social Twitter a través de la API de esta red social, se hace necesario homogeneizar tal información en un formato que facilite su tratamiento y acceso [20][21].

En la Fig. 1 se aprecia el procedimiento general del análisis que inicia con la extracción de información en Twitter hasta el análisis de sentimientos.

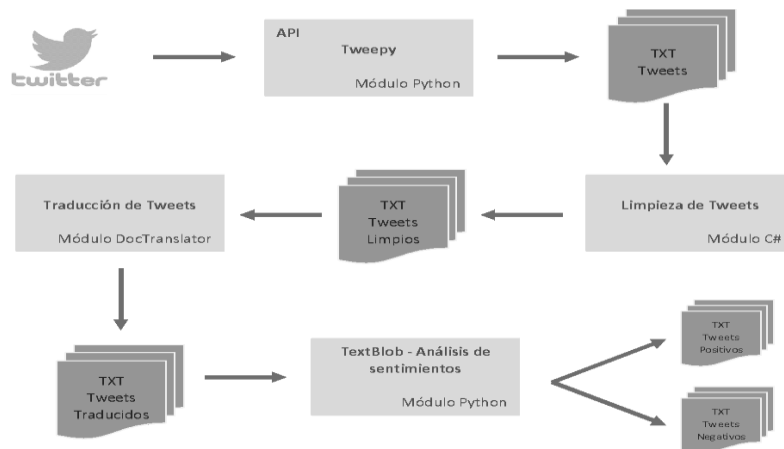


Fig. 1. Procedimiento de Análisis de comentarios en cambio climático (elaboración propia)

Una aproximación metodológica de la Fig. 1 donde se pueden enumerar tareas para el análisis de sentimientos se reducen a 4 etapas: • Tarea 1 - Extraer y categorizar entidades, • Tarea 2 - Extraer y categorizar aspectos, • Tarea 3 - Extraer y categorizar a los autores de la opinión, • Tarea 4 - Extraer el momento temporal: se trata de detectar el momento en el que la opinión fue emitida, • Tarea 5 - Clasificar la polaridad a nivel de aspecto [22].

3.1 Extracción de información vía descargas usando hashtags

Con el fin de asegurarse que los comentarios de las descargas corresponden al cambio climático, se dispuso que los hashtags se asocien directamente con palabras significativas del problema con el propósito de que los tweets (opiniones) de los usuarios sean identificados correctamente.

Solo se consideraron aquellos comentarios que tenían coincidencias con las palabras clave. De aproximadamente 10, 000 comentarios totales de casi 1 mes de descargas, se observó que la mayoría eran re-tweets y otros excluyentes del contexto, por tanto, también se eliminaron. El fragmento principal del código se muestra en el algoritmo 1.

Algoritmo 1. Hashtag.

```
#Términos de búsqueda
search_words=["#Contaminación", "#CambioClimático", "#ConcienciaAmbienta
l"]
#Nombre de archivos
archivos=["Contaminacion.csv", "Cambio.csv", "Conciencia.csv"]
```

Los comentarios generados se almacenan en tres archivos .csv diferentes para proceder a autenticar las claves de la API y descargar los comentarios con el algoritmo 2 de descargas.

Algoritmo 2. Descargas.

```
for nombre,busqueda in zip(archivos,search_words):
    busqueda = busqueda + " -filter:retweets"
    with open(nombre, 'a', encoding = 'utf-8') as f:
        csvWriter = csv.writer(f)
        for tweet in tweepy.Cursor(api.search, q=busqueda,
tweet_mode = "extended").items(500):
            tem = ' '.join(re.sub("(@[A-Za-z0-9]+)|([^0-9A-Za-z
\t]) |(\w+:\/\/\S+)", " ", tweet.full_text).split())
            csvWriter.writerow([tweet.created_at, tem])
```

Después de conseguir todos los comentarios almacenados y con un mismo formato, se ajustaron para obtener las palabras que utilizan los twitteros, las cuales deben aparecer mínimo 30 veces con al menos 4 caracteres. Los comentarios almacenados con un formato uniforme se procesaron para filtrar aquellos que contengan al menos una de las palabras representativas del problema. El procedimiento de la contabilización de palabras de los comentarios se plantea en el procedimiento de obtención de palabras, algoritmo 3.

Algoritmo 3. Obtener palabras.

```
with open("palabras.csv", "w", encoding = "utf-8") as f:
    writer = csv.writer(f)
    for key, value in lista.items():
        if value > 30 and len(key) > 3:
            writer.writerow([key,value])
```

En el algoritmo 4 se indica la manera en que se almacenan las palabras que superan las 30 réplicas con el fin de crear el diccionario.

Algoritmo 4. Creación de un diccionario.

```

for nombre in archivos:
    with open(nombre, encoding = "utf-8") as f:
        csvReader = csv.reader(f)
        for row in csvReader:
            if len(row) > 0:
                tem = row[1]
                tem = tem.lower().split()
                for clave in tem:
                    if clave[0] != "#":
                        if clave not in lista.keys():
                            lista[clave] = 1
                        else:
                            lista[clave] += 1

```

3.2 Creación de diccionarios positivo y negativo: análisis de sentimientos

Dado un texto escrito en lenguaje natural, el análisis de sentimientos o de opinión identifica la emoción de un conjunto de palabras en sentimiento positivo o negativo. Algunas aplicaciones de clasificación de opiniones con orientación semántica positiva (pulgares hacia arriba) u orientación semántica negativa (pulgares hacia abajo) han favorecido a la comprensión de este tema [23].

El análisis sentimientos sobre las opiniones de películas con el uso de Naive Bayes, clasificación de entropía máxima (MaxEnt) y máquinas de vectores de soporte probando diferentes enfoques de n-gramas, ha alcanzado buenos resultados permitiendo tratar la subjetividad, el sentimiento y las emociones respecto al gusto sobre películas [24].

Cuando se inició en la comunidad PLN una producción importante de artículos sobre análisis de sentimientos, solo se consideraba la polaridad positiva o negativa debido a que la neutralidad no era sencilla de explicar y este concepto sigue vigente [25]. En el presente documento, se determinó que apelando a los inicios y tradición de análisis de sentimientos, solo categorizamos los positivos y negativos, los cuales capturan las emociones ilustradas en la Fig. 2.



Fig. 2. Polaridad de los sentimientos

3.3 Naive Bayes

En este apartado, se expone el procedimiento para la clasificación de comentarios negativos o positivos. Las palabras con etiqueta negativa se deben almacenar en el diccionario llamado pesimista y las palabras positivas en el diccionario optimista.

El clasificador Naive Bayes asume que el efecto de una característica particular en una clase es independiente de otras y utiliza el Teorema de Bayes (1).

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)} \quad (1)$$

Donde (1) indica:

$P(h)$: es la probabilidad de que la hipótesis h sea cierta (independientemente de los datos). Esto se conoce como la probabilidad previa de h .

$P(D)$: probabilidad de los datos (independientemente de la hipótesis). Esto se conoce como probabilidad previa.

$P(h|D)$: es la probabilidad de la hipótesis h dada los datos D . Esto se conoce como la probabilidad posterior.

$P(D|h)$: es la probabilidad de los datos d dado que la hipótesis h era cierta. Esto se conoce como probabilidad posterior.

En caso de que se tenga una sola característica, el clasificador Naive Bayes calcula la probabilidad de un evento en los siguientes pasos:

- Calcular la probabilidad previa para las etiquetas de clase dadas.
- Determinar la probabilidad de probabilidad con cada atributo para cada clase.
- Asignar estos valores en el teorema de Bayes y calcular la probabilidad posterior.
- Identificar la clase que tiene una probabilidad más alta.

El algoritmo 5 describe brevemente a Naive Bayes cuya función central es la clasificación. La entrada para el procedimiento son los diccionarios creados por el algoritmo 4 o bien, los tweets con algunos ajustes sencillos. Una vez leídas las palabras de los diccionarios o frases, se calcula la probabilidad sobre un comentario para que pueda clasificarse como negativo o positivo. Para ello se utiliza la fórmula Naive Bayes de Python que genera dos valores asignados a los comentarios: menor a cero si un comentario es negativo o mayor a cero si el comentario es positivo y cero si el resultado no pudo clasificarse [21].

Algoritmo 5. Naive Bayes

```
matrizN = [x[:-1] for x in negative]
matrizP = [x[:-1] for x in positive]
tamPA = 0
tamPB = 0
for x in tem:
    if x in matrizP:
        tamPA += 1
    elif x in matrizN:
        tamPB += 1
pA = len(tem) / len(matrizP)
pB = len(tem) / len(matrizN)
pRA = (pA * pRA) / (pA + pB)
pRB = (pB * pRB) / (pA + pB)
res = pRA - pRB
```

4 Resultados

Para asegurarnos que todos los algoritmos descritos anteriormente fueron correctos y que se clasificaban la mayoría de los comentarios de manera adecuada, se diseñó una script que se ejecuta con comentarios preestablecidos. Este procedimiento se repite distintas veces.

4.1 Descargas

Con el propósito de seleccionar tweets que revelaran el sentir de los usuarios de Twitter respecto al cambio climático, previamente se seleccionó una muestra de 30 personas para que expresaran una sola palabra relacionada con cambio climático, entonces se seleccionaron aquellas palabras más representativas del problema y se procedió a etiquetar hashtags con las palabras escogidas.

Es importante conocer si las palabras utilizadas como hashtags fueron convenientes para poder obtener la mayor cantidad de comentarios-tweets relacionados con el corpus del problema. La Fig. 3 señala la frecuencia de los hashtags que se observaron dentro de los tweets descargados.

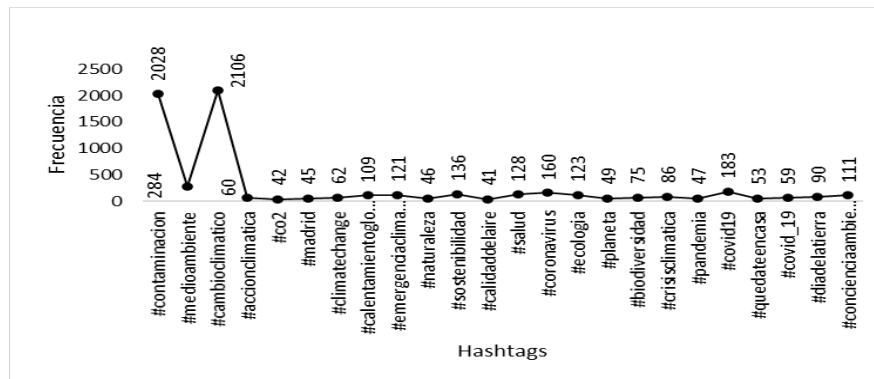


Fig. 3. Descargas de los tweets con los hashtags

La gráfica de la Fig.3 nos muestra que dos hashtags de los que se propusieron al principio del estudio son los más mencionados.

4.2 Diccionario total

Como resultado de los procedimientos de la sección anterior, se generaron los siguientes diccionarios, tabla 1.

Tabla 1. Diccionarios: cambio climático y contaminación.

Cambio climático		Contaminación	
Palabras	Frecuencia	Palabras	Frecuencia
cambio	143	minería	211
climático	124	cianuro	208
nuestro	210	envenena	208
planeta			
ciudades	95	calidad del aire	180
futuro	89	viviendo tiempos difíciles	386
científicos	79	plástico	202
mundo	79	microplásticos	187
capital	78	peces	187
indonesia	72	muerdes	34
hundiendo			
costeras	69	evolucion	80
crisis del	132	niveles de dióxido	249
agua			
region	127	nitrogeno	92
artico			

derretimien o	55	salud	43
Groenlandia	52	contaminacion	132
sgk-planet	46	ciudades	73
global	42	ambiental	31
emisiones	39	nuestro_planeta	85
climatica	39	reduccion atmosferica	87
medio	75	confinamiento	35
ambiente			
medidas de	35		
países			
pandemia	34		
2020	33		
carbono	32		
tierra	31		

Los diccionarios de las tabla1 son el resultado de haber descargado aproximadamente 10,000 comentarios en un mes.

4.3 Sentimientos positivos y negativos

El procedimiento de clasificación que generó los sentimientos negativos y positivos se puede apreciar en la tabla 2:

Tabla 2. Sentimientos positivos y negativos.

Positivos	Negativos
esperanza	contaminantes
mejora	contaminación
conciencia	riesgo
adecuado	muerres
bueno	problema
buena	problemas
optimista	impacto
positivo	consecuencias
proteger	pandemia
genial	crisis
magnifico	derretimiento
feliz	derrite
optimistas	envenenar
solidaridad	matando
bien	mala
luchando	malo
arreglar	impotencia
solucionar	pésimo
mejor	horror
oportunidades	negativo
impulso	enojado
impulsar	triste
sustentable	miedo
verde	vergüenza
queridos	culpables
amados	sufrimiento
bienestar	dolor

conforme	daños
increíble	alerta
construir	malestar
reconstruir	inconforme
amor	cansado
	cansada
	duele
	destruir
	destruyendo
	morir

En la gráfica de la Fig. 4 se puede observar el resultado de las clasificaciones de todos los comentarios descargados usando Naive Bayes, “Si” refiere a los comentarios clasificados como positivos, “No” son negativos y “SC” son comentarios sin clasificación o neutrales [26][27][28].

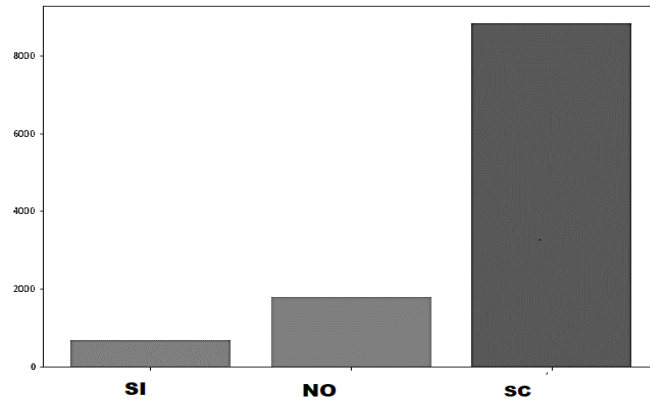


Fig. 4. Distribución de clasificaciones por Naive Bayes

Los comentarios neutrales corresponden a enunciados no relacionados con las palabras clave. En general incluyen pronombres confusos, calificativos disociados, abuso de contracciones y conectores etc., por tanto son inconsistentes y se encuentran limitados para ser clasificados. Por otra parte, el tamaño de los diccionarios iniciales creados por la encuesta a usuarios y que dio lugar a definir los hashtags, no son grandes y ello impacta en la identificación de comentarios de tweet para descargarlos. Si el número de hashtags es mayor, entonces la posibilidad de descargar mayor número de tweets con contenido semántico crece y en este trabajo, se pudo observar que a menos tamaño de las palabras clave, menor es el número de tweets con significado y mayor es el número de comentarios desatinados e ilógicos para el problema que se trata. En este punto, muchos vocablos y comentarios neutrales se generaron y la lista no es posible incorporarla.



Fig. 5. Distribución de clasificaciones por Naive Bayes

5 Conclusiones

Las distintas aplicaciones del PLN ayudan a resolver problemas donde el lenguaje natural/humano es el centro de discusión. Sin embargo, el estudio no es relevante si los datos no son suficientes y las redes sociales son una plataforma muy atractiva para recoger millares de datos.

En problemas donde la sensibilidad y conciencia toman un lugar importante, es útil la aplicación de técnicas PLN para conseguir resultados que incidan en la propuesta de programas sociales y/o gubernamentales cuando la gente no percibe una gravedad sustancial.

Una conclusión relacionada con los resultados de Naive Bayes, consiste en que a pesar de que son mayores los comentarios negativos, no existe gran diferencia entre los sentimientos positivos y negativos, entonces es conveniente mejorar la participación pública en el tema informando objetivamente a los usuarios web debido a que hemos identificado que las emociones se consideran fenómenos discursivos. En particular, el análisis de sentimientos negativos señala una inclinación al negacionismo y consecuentemente, otras técnicas deben ser aplicadas para su evaluación

A partir de los resultados obtenidos, es importante valorar las funciones que desempeña Twitter como herramienta de comunicación científica y subrayar que, en la aplicación de PLN en Twitter, se puede conseguir una distinción entre al menos dos aproximaciones de opiniones, aquellas enfocadas en la información que perturba y/o confunde y aquella objetiva, las cuales se pueden atribuir al análisis de sentimientos negativos y positivos respectivamente.

Considerando el estudio en Twitter donde se identificó tres tipos de personas [18], (los que perciben la emergencia climática, los negacionistas y los imparciales), es posible plantear que de acuerdo a la Fig. 4, los resultados obtenidos en este trabajo se relacionan de la siguiente forma:

El porcentaje de sentimientos negativos (SI) corresponden a emergencia climática, los negacionistas se vinculan directamente a los sentimientos negativos (NO) y los imparciales se asocian a los neutrales (SC), que son comentarios irrelevantes.

Es importante que otras herramientas de aprendizaje automático se apliquen a los datos generados no solo por opiniones, sino también a datos duros de emisiones contaminantes.

Finalmente, aclaramos que no se utilizó algún mecanismo para evaluar la calidad del modelo, no estaba planeado dentro del objetivo del trabajo, por tanto, esa tarea es amplia y merece otro reporte que se traduce en futura investigación y ahí consideraremos una ampliación de los diccionarios para capturar mejor calidad en las descargas de tweets y evitar el bloque de comentarios neutrales.

La sensibilidad del modelo no se analizó debido que en este manuscrito se analizaron sentimientos en negativos y positivos considerando la regla de Bayes, no así la matriz de confusión ni validación cruzada, para ello, un estudio sobre noticias fake de cambio climático comprendería validación cruzada.

Los comentarios neutrales no se analizaron, suponemos que muchas de estas opiniones neutrales tienen una connotación moral, emocional o afectiva que no se puede clasificar. Por tanto, hemos estudiado un documento donde los autores hablan de diccionarios morales y partir de ahí, explicaremos en un artículo posterior la manera en que la neutralidad puede dividirse en palabras morales [29].

Referencias

1. Höijer, B.: Emotional anchoring and objectification in the media reporting on climate change. *Public Understanding of Science*, Vol. 19, No. 6, pp. 717–731 (2010).
2. University of Michigan. Survey Research Center.; National Association of Science Writers.: *The Public Impact of Science in the Mass Media. A Report on a Nation-wide Survey for the National Association of Science Writers.* (1958).
3. Bauer, M.W.; Allum, N.; Miller, S.: What can we learn from 25 years of PUS survey research? Liberating and expanding the agenda. *Public Understanding of Science*, Vol. 16, No. 1, pp. 79–95 (2007).
4. Pardo, R.: La cultura científico-tecnológica de las sociedades de la modernidad tardía. *Treballs de la Societat Catalana de Biologia*, pp. 35–63 (2001).
5. Kapoor, K.K.; Tamilmani, K.; Rana, N.P.; Patil, P.; Dwivedi, Y.K.; Nerur, S.: Advances in social media research: Past, present and future. *Information Systems Frontiers*, Vol. 20, No. 3, pp. 531–558 (2018).
6. Shan, L.; Regan, Á.; De Brún, A.; Barnett, J.; Sanden, M.; Wall, P.; McConnon, Á.: Food crisis coverage by social and traditional media: A case study of the 2008 Irish dioxin crisis. *Public Understanding of Science*, Vol. 23, No. 8, pp. 911–928 (2014).
7. Lozares, C.: La teoría de redes sociales. *Revista de sociología*, pp. 103–126 (1996).
8. Cordero, R.; Lahuerta, O.: Redes sociales: un antes y un después en el comportamiento humano, *Fundación Telefónica* (2018).
9. Kozinets, R.V.: The Field behind the Screen: Using Netnography for Marketing Research in Online Communities. *Journal of Marketing Research*, Vol. 39, No. 1, pp. 61–72 (2002).
10. Wikipedia, https://es.wikipedia.org/wiki/Pew_Research_Center (2019).
11. García, J.F.; Martínez, J.L.; García, E.; Reyna, O.: Revisión de técnicas de pre-procesamiento de textos para la clasificación automática de tweets en español. *Revista de Cómputo Aplicado*, Vol. 1, No. 2, pp. 1–11, (2017).
12. Diaz, J.D.; Suarez, M.J.: Análisis social aplicando técnicas de lenguaje natural a información extraída de Twitter. *Scientia et Technica Año XXIV, Universidad Tecnológica de Pereira*, Vol. 24, No. 03, pp. 496–503 (2019).
13. Büchi, M.: Microblogging as an extension of science reporting. *Public Understanding of Science*, Vol. 26, No. 8, pp. 953–968 (2017).
14. Li, R.; Crowe, J.; Leifer, D.; Zou, L.; Schoof, J.: Beyond big data: Social media challenges and opportunities for understanding social perception of energy. *Energy Research & Social Science*, Vol. 56 (2019).
15. Veltri, G.A.: Microblogging and nanotweets: Nanotechnology on Twitter. *Public Understanding of Science*, Vol. 22, No. 7, pp. 832–849 (2013).
16. Schäfer, M.S.: Online communication on climate change and climate politics: a literature review. *Wiley Interdisciplinary Reviews Climate Change*, Vol. 3, No. 6, pp. 527–543 (2012).
17. Walter, S.; Lörcher, I.; Brüggemann, M.: Scientific networks on Twitter: Analyzing scientists’ interactions in the climate change debate. *Public Understanding of Science*, Vol. 28, No. 6, pp. 696–712 (2019).
18. Pearce, W.; Holmberg, K.; Hellsten, I.; Nerlich, B.: Climate Change on Twitter: Topics, Communities and Conversations about the 2013 IPCC Working Group 1 Report. *PLoS ONE*, Vol. 9, No. 4 (2014).
19. Definición del cambio climático, *Manos Unidas*, <https://www.manosunidas.org/observatorio/cambio-climatico/definición-cambio-climatico>.

20. Yener, Y.: Step by Step: Twitter Sentiment Analysis in Python <https://towardsdatascience.com/step-by-step-twitter-sentiment-analysis-in-python-d6f650ade58d> (2020).
21. Morrissey, M.; Wasser, L.; Farmer, C.: Lesson 2. Automate Getting Twitter Data in Python Using Tweepy and API Access, *Earth Lab*, <https://www.earthdatascience.org/courses/use-data-open-source-python/intro-to-apis/twitter-data-in-python/> (2020).
22. Sobrino, J.C.: Análisis de sentimientos en twitter, Máster Universitario en Ingeniería Informática, *Universitat Oberta de Catalunya* (2018).
23. Turney, P.: Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 417-424. (2002).
24. Pang, B.; Lee, L.; Vaithyanathan, S.: Thumbs up? Sentiment classification using machine learning techniques. *Association for Computational Linguistics*, pp. 79-86 (2002)
25. Russell, J.A.: A circumplex model of affect. *Journal of Personality and Social Psychology*, Vol. 39, No. 6, pp. 1161–1178 (1980)
26. Bakliwal, A.; Arora, P.; Madhappan, S.; Kapre, N.; Singh, M.; Varma, V.: Mining sentiments from tweets. *Association for Computational Linguistics*, pp. 11-18 (2012)
27. Hurtado, L., Pla, F., Buscaldi, D.: ELiRF-UPV en TASS 2015: Análisis de sentimientos en Twitter. *TASS Workshop at SEPLN 2015*, pp. 75-79 (2015)
28. Adjei, G.: Understanding Naive Bayes & its applications in text classification (Part I). *HEARTBEAT*, <https://heartbeat.fritz.ai/understanding-naive-bayes-its-applications-in-text-classification-part-1-ec9caea4baae> (2020). Accedido el 14 de Agosto de 2020.
29. Brady, W.J.; Wills, J.A.; Jost, J.T.; Tucker, J.A.; Van Bavel, J.J.: Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 114, No. 28, pp. 7313–7318 (2017).

Fundamentos del Álgebra Lineal en la Computación Cuántica

Héctor D. Ramírez, Roberto Contreras, Nelva B. Espinoza, Yusdivia Huerta
Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la
Computación
Edif. CC02, Ciudad Universitaria Av. San Claudio y 14 sur, Colonia San Manuel,
C.P. 72570, Puebla, Pue. México
{hector.ramirezhe, roberto.contreras, nelva.espinoza}@correo.buap.mx,
yusdivia.huerta@alumno.buap.mx

Resumen. En este trabajo se pretende visualizar el cómo la matemática formula una pequeña parte de la mecánica cuántica que es base importante en la computación cuántica, tratando de evitar adentrarnos en la física. Veremos los conceptos básicos de los sistemas de computación, así como la representación matemática del *qubit*, la unidad fundamental de la computación cuántica, dando la importancia en el concepto de superposición de forma matemática.

Palabras Clave: Bit, *Qubit*, Computación Cuántica.

1 Introducción

La idea de la computación cuántica fue introducida por primera vez en 1982, por Richard Feynman, quien consideró la simulación de sistemas en mecánica cuántica por otros sistemas cuánticos [4]. Sin embargo, la computación cuántica solo era hipótesis y fue que en 1985 se logró un avance fundamental, en la que David Deutsch, publicó un artículo en el que describía una computadora cuántica universal [2], probablemente haciendo pasar a la historia la máquina universal descrita por Alan Turing, en Cambridge, en los años 30 [5][7].

En la computación cuántica, a diferencia de la computación actual donde cada bit puede estar en un estado discreto y alternativo a la vez, la unidad fundamental de almacenamiento es el bit cuántico, denominado *qubit*, donde cada uno de ellos puede tener múltiples estados simultáneamente en un instante determinado, así reduciendo el tiempo de ejecución de algunos algoritmos de manera considerable. Uno de los algoritmos con esta característica es el algoritmo de Shor [3][8], el cual es un algoritmo para descomponer en factores un número N en tiempo polinomial.

En este trabajo, resaltaremos la importancia del álgebra lineal para dar la definición y representación matemática del *qubit*, incluyendo un conciso análisis matemático del concepto de superposición; además de que es ampliamente utilizada para describir estados *qubit*, operaciones cuánticas y para predecir lo que hace una computadora cuántica en respuesta a una secuencia de instrucciones. Todo esto centrado en mostrar a los estudiantes que cursan la materia de Álgebra Lineal que se pueden visualizar los conceptos que se estudian ahí como: vectores, espacios vectoriales, combinaciones lineales, operaciones entre matrices, entre otras, aplicados a los conceptos básicos de la computación cuántica.

2 Qubit

El elemento fundamental de la computación clásica es el bit el cual es la unidad básica de información. El estado de un bit está completamente determinado y puede tomar dos valores 0 y 1. Es importante resaltar que el estado en el que se encuentra un bit es el estado que se observa del bit. Esto es, el estado de un bit es físicamente observable, y se puede almacenar en un dispositivo de almacenamiento digital. Esto da lugar a un sistema de numeración binaria de almacenamiento, ya que se permite representar datos usando solo los dígitos 0 y 1.

La unidad fundamental de la computación cuántica es el *qubit* [1], el cual representa el estado de una partícula cuántica. Los posibles estados de un *qubit* están representados por dos estados ortogonales denotados por $|0\rangle$ y $|1\rangle$ los cuales son los análogos al 0 y 1 en la computación clásica. Estos dos estados ortogonales, pueden representarse de la siguiente forma:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (1)$$

Sin embargo, la parte cuántica del *qubit* es que $|0\rangle$ y $|1\rangle$ no son los únicos estados posibles. Un *qubit* puede estar en todo un rango continuo de estados comprendidos entre $|0\rangle$ y $|1\rangle$, es decir, en una superposición de ambos estados. Usando el álgebra lineal, el estado de un *qubit* se describe como un vector y es representado por una matriz de una sola columna $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$, en donde α y β son números complejos y se les denomina amplitudes de probabilidad¹. También se le conoce como vector de estado cuántico y debe cumplir con el requisito de que $|\alpha|^2 + |\beta|^2 = 1$. Los estados del sistema son denotados por $|\psi\rangle$ y se les denomina kets.

Los estados $|0\rangle$ y $|1\rangle$ representados de la forma (1), cumplen que:

- Son linealmente independientes,
- Generan al espacio vectorial $\mathbb{C}^2$,

Por lo que forman una base ortonormal del espacio vectorial $\mathbb{C}^2$. Esto implica que los *qubits* cuentan con una estructura algebraica de Espacio Vectorial, misma que se encuentra dotada de una operación interna de adición y una operación de producto por un escalar definida sobre el campo $\mathbb{C}$.

Los *qubits* por tener una estructura de espacio vectorial complejo, cumplen las siguientes propiedades:

	$ x\rangle + (y\rangle + z\rangle) = (x\rangle + y\rangle) + z\rangle$	Propiedad Asociativa
i)	$ x\rangle + 0\rangle = 0\rangle + x\rangle = x\rangle$	Existencia del neutro aditivo
ii)	$ x\rangle + -x\rangle = 0\rangle$	Existencia del inverso aditivo
iii)	$ x\rangle + y\rangle = y\rangle + x\rangle$	Propiedad conmutativa

- | | | |
|------|--|-------------------------------------|
| iv) | $a(b x\rangle) = (ab) x\rangle$ | Propiedad asociativa escalar |
| v) | $a(x\rangle + y\rangle) = a x\rangle + b y\rangle$ | Propiedad distributiva escalar |
| vi) | $(a + b) x\rangle = a x\rangle + b y\rangle$ | Propiedad distributiva |
| vii) | $1 x\rangle = x\rangle$ | Existencia de neutro multiplicativo |

Como ya se mencionó, un *qubit* puede encontrarse en superposición de los estados $|0\rangle$ y $|1\rangle$. Esto se representa mediante una combinación lineal de los vectores base de la forma:

$$\alpha|0\rangle + \beta|1\rangle. \quad (2)$$

El valor $|\alpha|^2$ representa la probabilidad de que al medir el *qubit*, este se encuentre en el estado $|0\rangle$. De la misma forma, $|\beta|^2$ determina la probabilidad de que al medir al *qubit*, este se encuentre en el estado $|1\rangle$. Las siguientes matrices representan vectores de estado cuántico válido:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \frac{1}{\sqrt{2}} \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

2.1 Representación de estado de n qubits.

El sistema cuántico de n *qubits* se representa como un vector de 2 elevado a la n *qubits*. Los cuatro estados $|00\rangle, |01\rangle, |10\rangle, |11\rangle$, en los que se puede representar un sistema de dos *qubits* se representa por el producto tensorial $\otimes$ de los vectores de los dos estados base de un *qubit*, de tal manera que:

$\begin{aligned} & 0\rangle \otimes 0\rangle \\ &= 00\rangle = \\ &\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ 0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \end{aligned}$	$\begin{aligned} & 1\rangle \otimes 1\rangle \\ &= 11\rangle = \\ &\begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \end{aligned}$
$\begin{aligned} & 0\rangle \otimes 1\rangle \\ &= 01\rangle = \end{aligned}$	$\begin{aligned} & 1\rangle \otimes 0\rangle \\ &= 10\rangle = \end{aligned}$

$ \begin{aligned} & \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ 0 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \end{aligned} $	$ \begin{aligned} & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ 1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \end{aligned} $
---	---

Representación de un Sistema de dos *qubits* mediante el product tensorial.

En un sistema de dos *qubits* se tiene que un estado de superposición es de la forma:

$$\begin{bmatrix} \alpha_{00} \\ \alpha_{01} \\ \alpha_{10} \\ \alpha_{11} \end{bmatrix},$$

de tal manera que se cumple:

$$|\alpha_{00}|^2 + |\alpha_{01}|^2 + |\alpha_{10}|^2 + |\alpha_{11}|^2 = 1. \quad (3)$$

De esta forma, la combinación lineal de los cuatro estados de dos *qubits* es:

$$|\psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle.$$

Donde $|\alpha_{00}|^2$ representa la probabilidad de que el *qubit* se encuentre en el estado $|00\rangle$, $|\alpha_{01}|^2$ será la probabilidad de que se encuentre en $|01\rangle$, y así sucesivamente.

De manera general, si se cuenta con n *qubits* en cada vector se obtiene un sistema con 2^n estados cuánticos

$$\underbrace{|000 \dots 0\rangle}_{n \text{ veces}} |000 \dots 10\rangle |000 \dots 100\rangle \dots |111 \dots 1\rangle$$

En donde cada uno de estos vectores forman una base para el espacio vectorial $\mathbb{C}^{2^n}$.

3 Operaciones Matriciales Sobre un Qubit.

Lo que se ha mencionado anteriormente es tratar de explicar qué es un *qubit* y cuáles son las propiedades que cumplen. En este apartado nos centraremos en visualizar algunas operaciones cuánticas, las cuales hacen uso de matrices, que modifican el estado de un *qubit* [6]. Sin embargo, debemos de mantener las propiedades de los *qubits*, ya que una operación sobre un *qubit* debe de ser de nuevo un *qubit*. De esta forma, si se quiere aplicar un operador A a un estado cuántico cualquiera $|\psi\rangle$, el resultado de dicha operación es:

$$A|\psi\rangle = |\psi'\rangle.$$

Dado que debemos de cumplir la normalización se debe de satisfacer que

$$\langle \psi' | \psi' \rangle = 1.$$

Si se sabe que

$$\langle \psi' | = \langle \psi | A^\dagger,$$

entonces

$$\langle \psi' | \psi' \rangle = \langle \psi' | A^\dagger A | \psi' \rangle = 1. \quad (4)$$

La relación (4) se cumple si y sólo si $A^\dagger A = I$, esto es, A debe de ser un operador unitario [9]. Por lo que se concluye que para que un operador A actúe sobre un *qubit* es necesario y suficiente que sea un operador unitario. Si se cuenta con la representación matricial es necesario y suficiente que la matriz sea unitaria, es decir, que su inversa coincida con la transposición conjugada de la matriz.

A continuación, se presentan algunas operaciones cuánticas básicas que se pueden aplicar a un *qubit*, las cuales pueden ser representadas con la multiplicación de matrices.

3.1 Operación X o puerta NOT

Esta operación es la única que se le puede aplicar a un *qubit* que cambia el estado $|0\rangle$ al estado $|1\rangle$ y viceversa. El operador que realiza este cambio de estado de un *qubit* está representado por la matriz denominada de Pauli y definida por:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (5)$$

Veamos que en efecto la operación dada por (5) cambia el estado de un *qubit* $|0\rangle$ al estado $|1\rangle$ y viceversa.

$$X|0\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

$$X|1\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

De manera general, si se le aplica el operador X al estado cuántico $\alpha|0\rangle + \beta|1\rangle$, entonces obtenemos que:

$$X \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}.$$

3.2 Operaciones Z y H

La operación Z , también conocida como phase-flip, se define como:

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (6)$$

La forma que actúa Z dada en (6) sobre los estados $|0\rangle$ y $|1\rangle$ es de la siguiente manera:

$$Z|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$$Z|1\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = -\begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Esto es, Z deja al estado $|0\rangle$ sin modificar y al estado $|1\rangle$ le cambia el signo, obteniendo $-|1\rangle$.

La operación H , también conocida como transformación de Hadamard, es definida por:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (7)$$

H dado en (7) actúa sobre los estados $|0\rangle$ y $|1\rangle$ es de la siguiente manera:

$$H|0\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad (8)$$

$$H|1\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (9)$$

Los estados resultantes al aplicar la operación H tienen una notación particular y es:

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \quad (10)$$

Es importante notar que la operación H , partiendo de un estado conocido ($|0\rangle$ o $|1\rangle$), genera una superposición de estados con probabilidad de $\frac{1}{2}$ cada uno.

El siguiente diagrama conmutativo muestra cómo obtener cada uno de los estados y como se relaciona a través de las operaciones analizadas, figura 1.

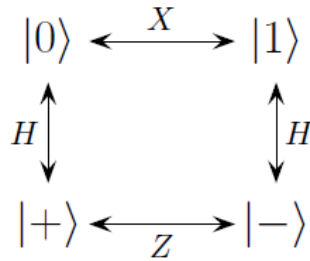


Fig. 1. Diagrama que relaciona las operaciones X , H y Z

Veamos como se comportan algunas de estas operaciones aplicadas al estado $|0\rangle$. La operación por realizar es:

$$HZH|0\rangle. \quad (11)$$

Por lo obtenido en (8), tenemos que:

$$H|0\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}. \quad (12)$$

Aplicando la operación Z a (12), obtenemos:

$$ZH|0\rangle = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \frac{|0\rangle + |1\rangle}{\sqrt{2}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Finalmente, para obtener lo que se pide en (11), tenemos:

$$HZH|0\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle.$$

Con ello, podemos observar que aplicar las operaciones HZH es lo mismo que aplicar una operación X .

4 Conclusiones

Como se pudo observar, se puede considerar al álgebra lineal como el lenguaje de la computación cuántica. Aunque no es necesaria para implementar o escribir programas cuánticos, es ampliamente utilizada para describir estados *qubit*, operaciones cuánticas y para predecir lo que hace una computadora cuántica en respuesta a una secuencia de instrucciones.

Así como es importante conocer los conceptos básicos de la física cuántica para entender la computación cuántica, el conocer algo de álgebra lineal básica puede mostrar a los estudiantes los conceptos necesarios para entender cómo funcionan los algoritmos cuánticos. De la misma forma, no debemos perder de vista la importancia que existe el conocer los vectores y la multiplicación de matrices para realizar operaciones cuánticas.

Referencias

1. G. Benenti, G. Casati, and G. Strini, *Principles of quantum computation and information*, vol. 1: Basic concepts, World Scientific, 2004.
2. D. Deutsch, "Teoría cuántica, el principio de Church-Turing y la computadora cuántica universal", *Proc. R. Soc. London A* **400**, 97-117 (1985).
3. Ekert, R. Jozsa: *Quantum Computation and Shor's Factoring Algorithm*. Reviews of Modern Physics, Vol. 68, No. 3, pp.733-753, (1996).
4. RP Feynman, "Simulación de la física con computadoras", *Int. J. Theor. Phys.* **21**, 467-488 (1982).
5. M. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem" *Proc. London Math. Soc.* 2, 230 (1936)
6. https://www.researchgate.net/publication/337758188_Computacion_cuantica_Aplicaciones_practicas_que_la_computacion_clasica_no_puede_solucionar
7. M. R. Álvarez, «Evolución de la computación cuántica y los retos para la seguridad de la información,» 2016.

8. A., Scott, “Shor, i’ll do it” en Cryptograph blog, 2007, disponible en la siguiente liga:
<http://www.scottaaronson.com/blog/?p=208>.
9. Sheldon, Linear algebra done right, Springer, Cham, 2015.

Diseño de un Robot Sanitizante anti COVID-19 con Navegación Autónoma

Ma. del Carmen Santiago, Ana C. Zenteno, Judith Pérez, Yeiny Romero,
Gustavo T. Rubín, Hermes Moreno Álvarez, Michel Yadira Montelongo
Flores, Alejandro Vargas Pérez

Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la
Computación, Ciudad Universitaria, 14 sur y Avenida San Claudio, Fraccionamiento
Jardines de San Manuel, CP. 72570, Puebla, Pue. México

{marycarmen.santiago, ana.zenteno, judith.perez, yeiny.romero,
gustavo.rubin}@correo.buap.mx, ymontelongo@uach.mx,
alejandro.vargasp@alumno.buap.mx

Resumen. El COVID-19 ha generado varios retos ante la sociedad, causando incertidumbre e inseguridad a las personas para regresar a una vida normal. Uno de los mayores retos ha sido en regresar de forma segura a los espacios públicos, viéndose afectada directamente la educación en nuestro país. En este trabajo se propone un robot sanitizante inteligente para garantizar un regreso seguro a las instalaciones de CU-BUAP utilizando un conjunto de sensores de aproximación para que el robot pueda navegar de forma autónoma y también sensores de temperatura y una cámara digital para determinar el uso correcto del cubrebocas, todo esto, utilizando algoritmos de Machine Learning. Se describen los módulos funcionales que integran el robot, así como los aspectos clave del desarrollo de los algoritmos y la programación de cada módulo y del sistema completo.

Palabras Clave: Machine Learning, Robot Sanitizante.

1 Introducción

Actualmente se vive una situación donde el contagio del COVID-19 es una amenaza latente para la sociedad. Existen diversos protocolos de sanidad que disminuyen el riesgo de contagios, entre los cuales se encuentran, el uso correcto de cubrebocas y la constante sanitización de los espacios, para reducir la transmisión del virus, sin embargo, algunas personas no hacen uso correcto o total del cubrebocas, así como la dificultad de sanitizar las áreas en todo momento, ocasionando un riesgo mayor de contagio a los individuos que lo rodean [1] [7][8][9].

La propagación del virus a nivel comunitario se genera por transmisión de persona a persona y sucede con mayor frecuencia entre contactos que se encuentren a una distancia menor a un metro. Este tipo de transmisión ocurre a través de gotas respiratorias, que se producen al hablar, toser o estornudar. Estas gotas también pueden permanecer en objetos y superficies durante horas o días [2].

El COVID-19 es un virus susceptible a los desinfectantes en comparación con otros virus, de esta manera, la limpieza y desinfección de superficies con productos idóneos es de suma importancia como estrategia de prevención y control para la transmisión de COVID-19 en entornos comunitarios [3][9].

Dentro de las consideraciones generales para la limpieza y desinfección de superficies según la documentación oficial de Salud de Gobierno de México se encuentran las siguientes:

- Se debe limpiar y desinfectar todas las áreas, como oficinas, baños y áreas comunes
- Enfocarse principalmente en las superficies que se tocan con frecuencia como: manijas, barandales, puertas, etc.
- La mayoría de los desinfectantes comunes de uso doméstico son eficaces para las tareas de desinfección, como: el hipoclorito de sodio.
- Se deberán de seguir las instrucciones del fabricante, utilizando las recomendaciones de concentración, forma de aplicación y tipo de contacto.
- No se deberán utilizar concentraciones de desinfectantes superiores a las recomendadas, por ejemplo, las altas concentraciones de cloro pueden provocar la corrosión de las superficies, así como efectos perjudiciales para la salud, como la irritación de la piel o mucosas hasta intoxicación.
- En el caso de las superficies porosas como pisos alfombrados y cortinas, quite la contaminación visible si existiera; si los objetos son lavables, se debe realizar siguiendo las instrucciones del fabricante [4][9].

Para facilitar la tarea de sanitización se han implementado robots como el “Hyper Light” perteneciente a la empresa “steganini group”, el cuál es un robot encargado de prevenir la infección y combatir virus utilizando el método de desinfección UVGI [5]. Este robot se utiliza para neutralizar a los virus en una zona cerrada utilizando la luz UV y evitar la propagación de ellos [6].

Como solución a la situación planteada anteriormente, se propone un robot inteligente con sistema de reconocimiento facial, capaz de sanitizar los espacios públicos dentro de la universidad sin causar problemas a las personas que se encuentran a su alrededor, así como, detectar si las estas utilizan correctamente el cubrebocas, propiciando un ambiente libre de contagios en zonas públicas. El sistema propuesto cuenta con diferentes etapas para su correcto funcionamiento, como lo son el movimiento del robot, el uso eficiente de sensores, la programación para la interacción con los usuarios y la oportuna reacción a los diversos eventos que puedan suceder.

2 Descripción Funcional del Robot Sanitizante

Las tareas que realizará el robot mientras se encuentre trabajando, será recorrer el área establecida, mientras al mismo tiempo está sanitizando con rociador la zona y al momento que detecte algún objeto, determinar si este es un obstáculo para corregir su trayectoria, o, si es alguna persona, detener los rociadores y proseguir con su camino, además estará interactuando con las personas que encuentre durante su recorrido midiendo la temperatura y verificando que estén utilizando correctamente el cubrebocas mediante reconocimiento facial.

Para la elaboración del proyecto es importante tomar en cuenta qué características van a tener los distintos componentes de cada módulo con la finalidad de cumplir con las tareas propuestas. A continuación, se describe cada módulo y sus características.

2.1 Módulo de percepción del entorno dinámico

El objetivo de este módulo es controlar el movimiento del robot a partir de la información procedente de sensores ultrasónicos que sean capaces de detectar objetos a la distancia que se requiera ajustar y que generen el menor ruido posible en la medición, con el objetivo de ejecutar un plan establecido y realizar una determinada tarea reaccionando en tiempo real ante cambios en su entorno, utilizando Inteligencia Artificial y Machine Learning.

2.2 Módulo motriz

El módulo está compuesto por cuatro llantas, ocho sensores ultrasónicos distribuidos de forma geométrica, en la superficie exterior del robot y dos motores eléctricos ensamblados en las dos llantas delanteras como se puede observar en la figura 1.

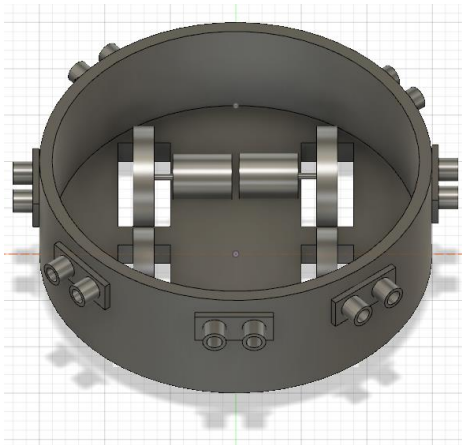


Fig. 2. Propuesta del módulo motriz, donde se observan ocho sensores ultrasónicos colocados alrededor de la base y los dos motores eléctricos que generan el movimiento a las llantas.

El objetivo es que el vehículo lleve a cabo de forma autónoma movimientos previamente planificados y que le permitan reaccionar de acuerdo a la percepción del entorno.

En la figura 2 se muestra el diagrama del algoritmo de control del vehículo propuesto en el que se identifican dos bucles. El del lado izquierdo es el encargado de llevar el control de la dirección del vehículo, generando las señales de control para los actuadores a fin de que el vehículo tome en cada instante la dirección apropiada de acuerdo con su entorno. El bucle de la derecha está a cargo de controlar la velocidad robot, generando la señal de control para los actuadores que realizan la propulsión del vehículo.

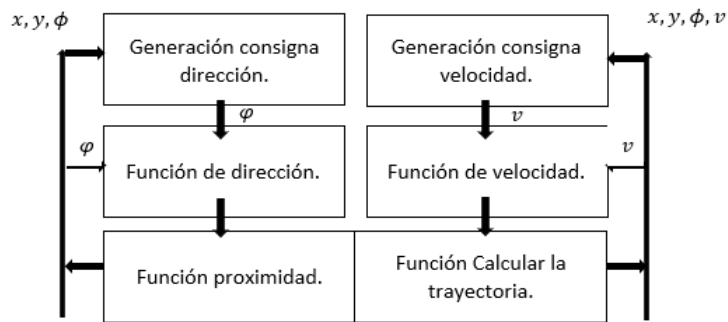


Fig. 2. Diagrama del algoritmo del control de velocidad y desplazamiento. La posición del vehículo se expresa por x, y ; ϕ la orientación y v la velocidad del robot.

En la figura 3 se muestra el algoritmo para la ejecución del módulo de movimiento a partir de la percepción de su entorno generando diversas alternativas dependiendo del escenario en que se encuentre el robot en su entorno.

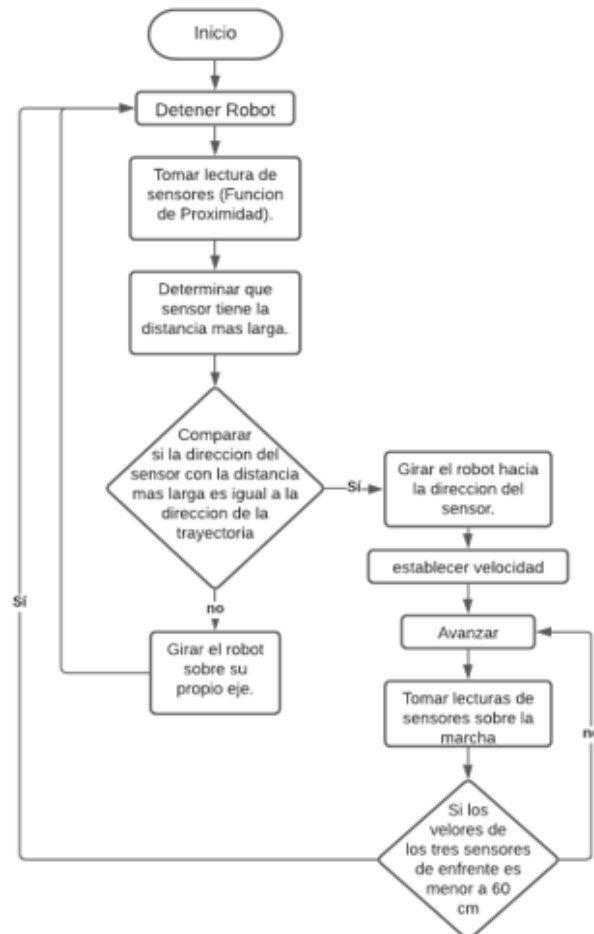


Fig. 3. Algoritmo de control de movimiento a partir de la lectura de los sensores dentro del entorno del robot.

2.3 Módulo de reconocimiento facial

El uso de este módulo tiene como objetivo reconocer que el robot se encuentra frente a una persona y verificar que esté utilizando de manera correcta el cubrebocas.

Para la implementación de este módulo el robot utilizará como apoyo una cámara con lectura de imágenes en alta definición para obtener resultados certeros y con poco margen de error.

El módulo de reconocimiento facial para la verificación de que la persona tiene colocado su cubrebocas correctamente, se utiliza un algoritmo el cuál involucre procesamiento de imágenes, reconocimiento facial, entrenamiento del sistema y ejecución de resultados, para generar un análisis de la imagen captada por la cámara y de esta forma dictaminar un resultado. A continuación, en la figura 3, se muestra el orden en el que se realizó este algoritmo.

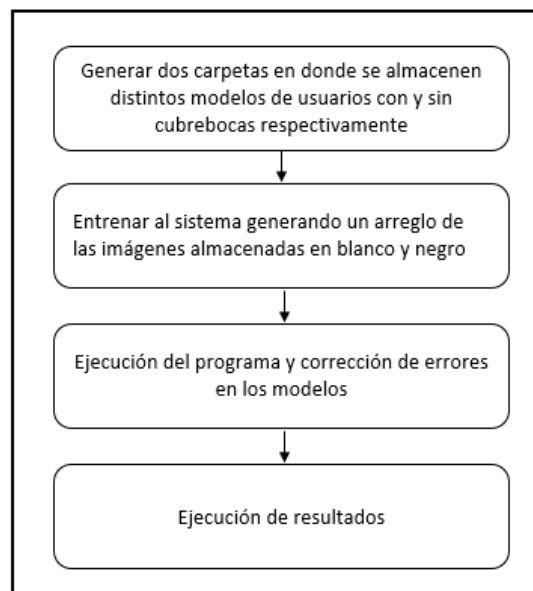


Fig. 3. Pasos para generar el algoritmo para el sistema de reconocimiento del correcto uso de cubrebocas.

Una vez implementados estos pasos en el desarrollo de código en Python. Con la ayuda de librerías de OpenCV, NumPy y MediaPipe, que permiten manejar las imágenes de acuerdo a las condiciones en las que se encontrará el robot, permitiendo, un procesamiento óptimo, reduciendo el margen de error en los resultados, siendo así que nos brinden herramientas como reconocimiento en automático de un rostro, convirtiendo las imágenes en blanco y negro, así como la detección de figuras en el entorno [10][11][12]. Vea figuras 4, 5 y 6.



Fig. 4. Uso de la librería NumPy para la detección de objetos.



Fig. 5. Uso de la librería OpenCV para convertir imágenes en blanco y negro.



Fig. 6. Uso de la librería MediaPipe para reconocer rostros automáticamente.

2.4 Módulo para determinar la temperatura del usuario

Este módulo es utilizado para determinar la temperatura de las personas que vaya encontrando el robot en su trayectoria con ayuda de sensores ultrasónicos, además cuenta con una cámara que servirá de apoyo para que el sensor de temperatura corporal se encuentre en una posición correcta de acuerdo a la ubicación de la persona y pueda determinar si su temperatura es menor a 37°C. En la figura 7 se muestra el funcionamiento de este módulo.

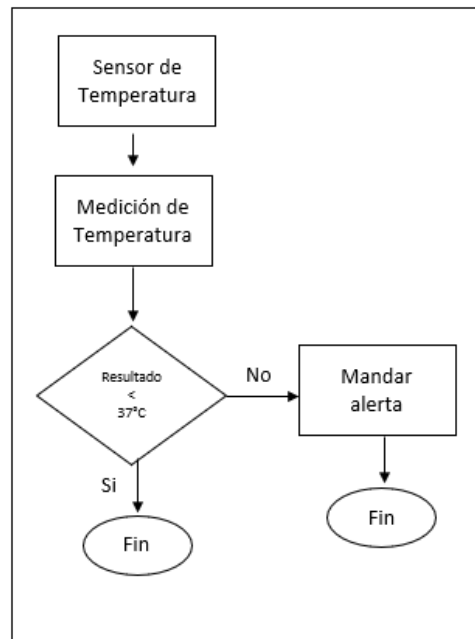


Fig. 7. Funcionamiento del módulo para la medición de la temperatura.

2.5 Módulo sanitizante

Este módulo es el encargado de sanitizar las zonas que se encuentren dentro del recorrido del robot, utilizando un sistema para controlar la apertura y cierre de electroválvulas en las tuberías portadoras del flujo sanitizante cuando sea necesario.

Cada determinado tiempo el robot activará las electroválvulas para rociar el área, teniendo la capacidad de detener a tiempo el flujo sanitizante para no afectar a las personas cercanas a su recorrido, todo esto haciendo uso de la cámara y los sensores ultrasónicos, como se muestra en la figura 8.

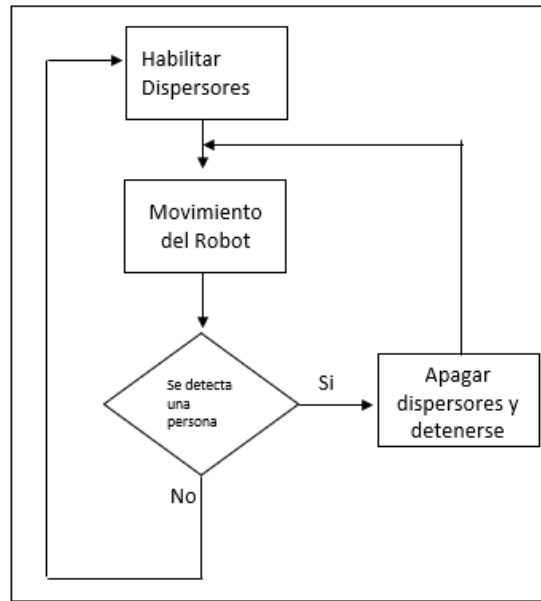


Fig. 8. Forma de operación para el módulo sanitizante.

3 Resultados

3.1 Resultados del módulo de reconocimiento facial

A partir de la implementación de los procesos explicados anteriormente en el módulo de reconocimiento facial, en la figura 9 y 10 se muestra el funcionamiento del módulo, utilizando el conjunto de modelos almacenados en él para determinar si una persona utiliza o no el cubrebocas.

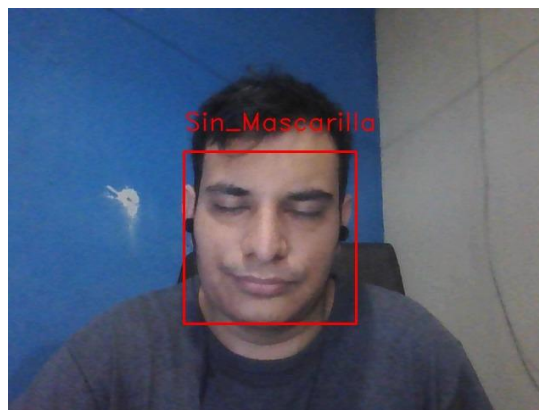


Fig. 9. Resultado del sistema cuando detecta que no se utiliza el cubrebocas.



Fig. 10. Resultado del sistema cuando detecta que se utiliza el cubrebocas.

3.2 Implementación del algoritmo final del proyecto

Teniendo en cuenta los módulos descritos anteriormente que son necesarias para que el robot cumpla con sus tareas, se debe realizar un enlace entre estos para que el resultado sea el esperado. A continuación, en la figura 11, se muestra el ciclo de funcionamiento de todos los módulos en conjunto de acuerdo a las diversas situaciones en las que se encuentre el robot.

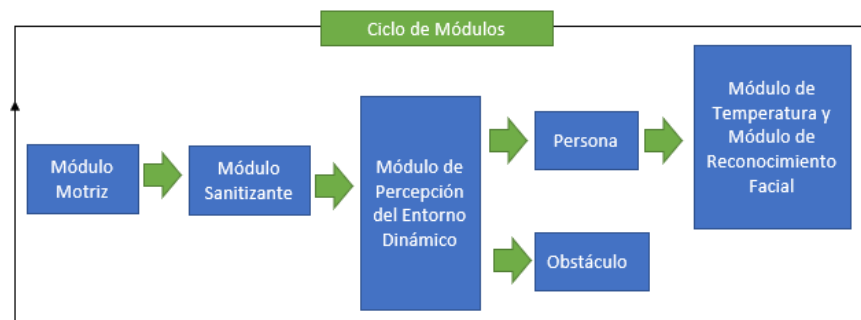


Fig. 11. Ciclo de los procesos en que el robot ejecutará los módulos.

4 Conclusiones

En la actualidad la sociedad se está enfrentando a nuevos retos a causa de la pandemia, uno de ellos, es garantizar la seguridad de las personas para regresar a lugares públicos como escuelas, universidades, plazas comerciales, etc. El proyecto propuesto en este artículo actualmente se encuentra en la etapa de diseño e implementación que tiene como finalidad proponer una solución a este problema dando cara a la situación actual en la que vivimos, garantizando seguridad a los estudiantes para su pronto regreso a las instalaciones de CU-BUAP.

Hay que tomar en cuenta que este proyecto puede desarrollarse de forma autónoma, es decir, sin la necesidad de utilizar librerías o tecnologías que ya se están utilizando en otros proyectos, pero ante los retos que nos ha venido trayendo el COVID-19 en los

últimos años, el uso de estas tecnologías es de suma importancia para brindar soluciones rápidas y eficientes, proyectando mejoras del proyecto mediante el desarrollo de herramientas y algoritmos propios a mediano y largo plazo.

Referencias

1. Organización Mundial de la Salud. Protégase a sí mismo y a los demás contra la COVID-19. Disponible en: <https://www.who.int/es/emergencias/diseases/novel-coronavirus-2019/advice-for-public>. Consultado el 23 de Noviembre del 2021.
2. Organización Mundial de la Salud. Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations(2020). Disponible en: <https://apps.who.int/iris/bitstream/handle/10665/332168/WHO-2019-nCoV-Disinfection-2020.1-spa.pdf?sequence=1&isAllowed=y>. Consultado el 23 de Noviembre del 2021.
3. Gobierno de México (Junio 11, 2021). Limpieza y desinfección de espacios comunitarios durante la pandemia por SARS-CoV-2. Disponible en: https://coronavirus.gob.mx/wp-content/uploads/2020/06/Limpieza_desinfeccion_espacios_comunitarios_durante_pandemia_SARS-CoV-2.pdf. Consultado el 23 de Noviembre del 2021.
4. Centros para el Control y Prevención de Enfermedades. Cómo limpiar y desinfectar su establecimiento (15 de Noviembre del 2021). Disponible en: <https://espanol.cdc.gov/coronavirus/2019-ncov/community/disinfecting-building-facility.html>. Consultado el 23 de Noviembre del 2021.
5. Dispositivos Médicos CANIFARMA. Robots de desinfección, procedimientos automatizados hacen frente a coronavirus (3 de Marzo del 2020). Disponible en: <https://dispositivosmedicos.org.mx/robots-de-desinfeccion-procedimientos-automatizados-hacen-frente-a-coronavirus/>. Accedido el 23 de Noviembre del 2021.
6. Stefanini Group. Robot de desinfección Hyper Light (3 de Marzo del 2020). Disponible en: <https://stefanini.com/es/soluciones/respuesta-stefanini-ante-el-covid-19/robot-de-desinfeccion-hyper-light>. Consultado el 23 de Noviembre del 2021.
7. Mofeyisara Oluwatoyin Omobowale, Eniola Adetola Bamgboye, Akinfemi Akinyode, Olugbenga Samuel Falase, Taiwo Olabode Ladipo, Olufunmilayo Salami, Akindele Olupelumi Adebisi. Contextual interpretation of COVID-19 pandemic among public space users in Ibadan Metropolis, Oyo State, Nigeria: An ethnographic review. Disponible en: <https://doi.org/10.1371/journal.pone.0259631>. Consultado el 25 de Noviembre del 2021.
8. Muh-Yong Yen, Jonathan Schewartz, Chwan-Chuen King, Chung-Ming Lee, Po-Ren Hsueh, Society of Taiwan Long-term Care Infection Prevention and Control. Recommendations for protecting against and mitigating de COVID-19 pandemic in long term care facilities. Disponible en: <https://doi.org/10.1016/j.jmii.2020.04.003>. Consultado el 25 de Noviembre 2021.
9. Oscar Omondi Donde, Evans Atoni, Anastasia Wairimu Muia, Paul T. Yillia. COVID-19 pandemic: Water, sanitation and hygiene (WASH) as a critical control measure remains a major challenge in low-income countries. Disponible en: <https://doi.org/10.1016/j.watres.2020.116793>. Consultado el 25 de Noviembre 2021.
10. José Antonio Taquía Gutiérrez. El procEsamiEnto dE imágEnEs y su potEncial aplicación En EmprEsas con EstratEgia digital (15 de Octubre del 2017). Disponible en: <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwj4I3D1rP0AhWcnGoFHeDyC3wQFnoECAIQAQ&url=https%3A%2F%2Fdialectnet.unirioja.es%2Fdescarga%2Farticulo%2F6230450.pdf&usg=AOvVaw0oMNRjHxUHbPeHX2d8km1q>. Consultado el 23 de Noviembre del 2021.

11. Charles R. Harris¹, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke & Travis E. Oliphant. Array Programming with NumPy. Disponible en: <https://doi.org/10.1038/s41586-020-2649-2>. Consultado el 25 de Noviembre del 2021.
12. D. Kavitha, B.P. Rishi Kiran, B. Niteesh, S. Praveen. Multiple Object Recognition Using OpenCV. Disponible en: <https://doi.org/10.47059/revistageintec.v11i2.1795>. Consultado el 26 de Noviembre 2021.

Aportaciones de las Ciencias Computacionales Durante la Pandemia COVID19

se terminó de editar en Diciembre de 2021 en la

Facultad de Ciencias de la Computación

Av. San Claudio y 14 Sur Jardines de San Manuel

Ciudad Universitaria

C.P. 72570

Aportaciones de las Ciencias Computacionales Durante la Pandemia COVID19
Coordinado por Gustavo Trinidad Rubín Linares

