UNIVERSIDAD POLITÉCNICA SALESIANA

SEDE GUAYAQUIL

CARRERA DE COMPUTACIÓN

Identificación de modelos de aprendizaje automático para la detección de ransomware en empresas basado en bibliografía.

Trabajo de titulación previo a la obtención del
Título de Ingeniero en ciencias de la computación

AUTOR: EVELYN IVONE MITE FLORES

TUTOR: MSc. VALVERDE LANDIVAR GALO ENRIQUE

Guayaquil – Ecuador

2025

CERTIFICADO DE RESPONSABILIDAD Y AUTORÍA DEL TRABAJO DE TITULACIÓN

Yo, **Evelyn Ivone Mite Flores** con documento de identificación N° **0957883226**: manifiesto que:

Soy el autor y responsable del presente trabajo; y, autorizo a que sin fines de lucro la Universidad Politécnica Salesiana pueda usar, difundir, reproducir o publicar de manera total o parcial el presente trabajo de titulación.

Guayaquil, 22 de enero de 2025

Atentamente,

Evelyn Mite

Evelyn Ivone Mite Flores

0957883226

CERTIFICADO DE CESIÓN DE DERECHOS DE AUTOR DEL TRABAJO DE TITULACIÓN A LA UNIVERSIDAD POLITÉCNICA SALESIANA

Yo, **Evelyn Ivone Mite Flores** con documento de identificación No. **0957883226**, expreso mi voluntad y por medio del presente documento cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autor(a) del Artículo Académico: **"Identificación de modelos de aprendizaje automático para la detección de ransomware en empresas basado en bibliografía"**, el cual ha sido desarrollado para optar por el título de: **Ingeniero de Sistemas, en la Universidad Politécnica Salesiana**, en la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En concordancia con lo manifestado, suscribo este documento en el momento que hago la entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Guayaquil, 22 de enero de 2025

Atentamente,

Evelyn Mite

Evelyn Ivone Mite Flores

0957883226

CERTIFICADO DE DIRECCIÓN DEL TRABAJO DE TITULACIÓN

Yo, Galo Enrique Valverde Landivar con documento de identificación N° 0912511532, docente de la Universidad Politécnica Salesiana, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: Detección de ransomware mediante aprendizaje automático, realizado por Evelyn Ivone Mite Flores con documento de identificación N° 0957883226, obteniendo como resultado final el trabajo de titulación bajo la opción Artículo Académico que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Guayaquil, 22 de enero de 2025

Atentamente,

Galo Enrique Valverde Landivar

0912511532

DEDICATORIA

Este trabajo es dedicado a mi madre, cuyo cariño y dedicación siempre han sido mi mayor fuente de inspiración, y a pesar de que ya no está a mi lado, su herencia perdura en cada paso que doy. A mi padre, por su constante respaldo, y a mi hermana, quien me proporcionó su apoyo financiero y me facilitó continuar con mis estudios. Además, dedico esta obra a todos aquellos que me acompañaron y respaldaron durante este trayecto, y a los amigos que transformaron esta travesía académica en una vivencia más enriquecedora y provechosa.

AGRADECIMIENTO

Estoy profundamente agradecido a todos los que me han acompañado en este viaje académico, en particular a los docentes, cuyo consejo y guía han sido esenciales para potenciar mi experiencia en el trabajo y en el ámbito profesional. A mi familia y amigos, cuyo respaldo ininterrumpido me ha facilitado el desarrollo en todos los aspectos de mi vida. También quiero expresar mi agradecimiento a aquellos individuos que, con su estímulo y motivación, se transformaron en mi pilar de apoyo en los momentos de incertidumbre, cuando pensaba que no podría continuar. Agradezco a todos ellos, pude seguir y lograr mis objetivos.

RESUMEN

Se lleva a cabo un estudio detallado de los modelos de aprendizaje automático empleados en la identificación de ransomware en ambientes corporativos, fundamentado en un amplio estudio bibliográfico. La investigación analiza el progreso desde las técnicas convencionales fundamentadas en firmas, hacia soluciones más avanzadas que utilizan inteligencia artificial. Diversos algoritmos son reconocidos y evaluados, sobresaliendo Random Forest (RF) y Extreme Gradient Boosting (XGB), que han evidenciado tasas de detección que superan el 97% en la detección de amenazas.

El estudio persigue tres metas concretas: el análisis de investigaciones anteriores acerca de métodos de ML, la consolidación de descubrimientos acerca de beneficios y restricciones de cada método, y la comparación de estrategias en diversas plataformas tecnológicas. Los hallazgos indican que los métodos híbridos, que fusionan análisis estático y dinámico, resultan especialmente eficaces para identificar variantes desconocidas de ransomware.

El estudio experimental llevado a cabo presenta matrices de confusión y evaluaciones de exactitud que muestran retos importantes, tales como el desbalance en los grupos de datos y la exigencia de mejorar la detección en la categoría minoritaria (ransomware). Se reconocen sectores críticos de mejora, entre ellos la escalabilidad de soluciones en infraestructuras corporativas y la adaptación a ambientes con recursos escasos como IoT.

Las conclusiones destacan la necesidad de un enfoque integrado que combine modelos de ML con marcos regulatorios como el NIST CSF, señalando la importancia de desarrollar soluciones adaptativas que puedan responder a la creciente sofisticación de los ataques que incorporan técnicas de inteligencia artificial ofensiva.

Palabras claves: Aprendizaje automático, Detección de ransomware, Análisis híbrido de malware, Ciberseguridad empresarial, Random Forest.

ABSTRACT

This study carries out a detailed examination of machine learning models used in ransomware identification in corporate environments, based on extensive bibliographic research. The investigation analyzes the progression from conventional signature-based techniques toward more advanced solutions that utilize artificial intelligence. Various algorithms are identified and evaluated, with Random Forest (RF) and Extreme Gradient Boosting (XGB) standing out, having demonstrated detection rates exceeding 97% in threat detection.

The study pursues three specific goals: the analysis of previous research on ML methods, the consolidation of findings regarding benefits and limitations of each method, and the comparison of strategies across various technological platforms. The findings indicate that hybrid methods, which combine static and dynamic analysis, are particularly effective in identifying unknown ransomware variants.

The experimental study conducted presents confusion matrices and accuracy evaluations that reveal significant challenges, such as imbalance in data sets and the need to improve detection in the minority class (ransomware). Critical areas for improvement are identified, including the scalability of solutions in corporate infrastructures and adaptation to resource-constrained environments such as IoT.

The conclusions highlight the need for an integrated approach that combines ML models with regulatory frameworks such as NIST CSF, emphasizing the importance of developing adaptive solutions that can respond to the growing sophistication of attacks incorporating offensive artificial intelligence techniques.

Key words: Machine Learning, Ransomware Detection, Hybrid Malware Analysis, Enterprise Cybersecurity, Random Forest.

ÍNDICE DE CONTENIDO

1. INTRODUCCIÓN

El ransomware se ha establecido como una de las amenazas más sofisticadas, persistentes y disruptivas dentro del panorama de la ciberseguridad actual. Este tipo de malware tiene la

capacidad de encriptar datos críticos y exigir rescates económicos considerables, representando un desafío significativo para las pequeñas y medianas empresas (pymes), las cuales suelen carecer de los recursos técnicos y financieros necesarios para prevenir o mitigar estos ataques de manera efectiva (Cen et al., 2024). El impacto de un ataque de ransomware no se limita a la interrupción inmediata de las operaciones, sino que también abarca daños a largo plazo, como la pérdida de confianza de los clientes, daños reputacionales y complicaciones legales, lo que dificulta enormemente la recuperación de las empresas afectadas.

El progreso continuo en las estrategias utilizadas por los cibercriminales ha demostrado las restricciones de los métodos convencionales de identificación de malware, como los sistemas basados en firmas. A pesar de que estos métodos son eficaces en la detección de amenazas conocidas, han probado ser insuficientes ante las versiones más avanzadas de ransomware que emplean técnicas de evasión sofisticadas, como la ofuscación de código y la simulación de procesos legítimos (Das et al., 2024). En este escenario, es esencial crear soluciones innovadoras que no solo faciliten la detección proactiva de estas amenazas, sino que también sean capaces de ajustarse al ritmo rápido de su desarrollo.

El aprendizaje automático (ML, en inglés) se ha presentado como un instrumento revolucionario en la batalla contra el ransomware. Los modelos de Machine Learning proporcionan la habilidad de examinar grandes cantidades de datos en tiempo real, detectar patrones inusuales y modificar de manera dinámica su comportamiento ante nuevas estrategias empleadas por los atacantes. Por ejemplo, algoritmos como las Redes Neuronales, los Árboles de Decisión y la Máquina de Apoyo Vectorial (SVM) han mostrado un desempeño destacado en la identificación precoz de ransomware, consiguiendo no solo disminuir el tiempo de respuesta, sino también atenuar las consecuencias devastadoras de estos ataques (Alzahrani et al., 2022). Esta habilidad para fusionar rapidez y exactitud en la detección sitúa al ML como un recurso esencial para las pequeñas y medianas empresas, cuyos recursos escasos demandan soluciones sumamente eficaces y escalables.

Pese a los progresos tecnológicos, los atacantes también han empezado a incluir instrumentos de inteligencia artificial y aprendizaje automático en sus tácticas, incrementando de esta manera la complejidad de los ataques. Esta circunstancia ha motivado a los científicos a indagar en métodos híbridos que fusionen análisis estáticos y dinámicos, además de técnicas sofisticadas como el seguimiento de llamadas a la API, el estudio del tráfico en la red y la representación

de datos complejos. Estos procedimientos no solo incrementan la exactitud en la identificación, sino que también posibilitan prever y atenuar futuros patrones de ataque, reforzando la posición de defensa cibernética (Zakaria et al., 2025).

Uno de los retos más significativos en la identificación de ransomware es la detección de malware modificado, creado específicamente para eludir los sistemas de detección tradicionales. La mezcla de métodos de análisis sofisticados con algoritmos de aprendizaje supervisado y no supervisado, tales como la Regresión Logística y las Redes Neuronales Convolucionales, ha probado ser especialmente eficaz para tratar este problema, consiguiendo resultados alentadores en cuanto a exactitud y velocidad de detección **Hossain et al. 2024**

En este contexto, el Marco de Ciberseguridad (CSF) del Instituto Nacional de Estándares y Tecnología (NIST) emerge como una orientación crucial para la administración de riesgos en el ámbito cibernético. Este marco estructura sus controles en cinco áreas clave: Identificar, Salvaguardar, Identificar, Actuar y Recuperar. Su implementación, ajustada a las demandas particulares de las pequeñas y medianas empresas, posibilita robustecer su posición de seguridad y potenciar su habilidad para resistir amenazas avanzadas como el ransomware (Ashley et al., 2022). La aplicación de este tipo de estrategias regulatorias en combinación con enfoques fundamentados en ML ofrece una ruta completa hacia la salvaguarda de las empresas en el entorno actual. Este estudio detallado de metodologías estáticas, dinámicas e híbridas no solo permitirá reconocer las restricciones presentes de los modelos de ML, sino también sugerir áreas de optimización que favorezcan la creación de soluciones más sólidas y adaptativas para salvaguardar los datos y asegurar la continuidad operativa de las compañías (Falowo, Edinam Botsyoe, et al., 2024)

2. REVISIÓN DE LITERATURA

2.1. EVOLUCIÓN Y CARACTERÍSTICAS DEL RANSOMWARE

La protección frente al ransomware se centra en salvaguardar sistemas operativos y archivos de

alteraciones no deseadas, mientras que la prevención intenta prevenir que el ransomware llegue a las víctimas. Estas estrategias comprenden métodos como el análisis estático, el aprendizaje automático y la gestión de accesos a archivos. Los creadores de ransomware enfrentan el desafío de esquivar las estrategias de mitigación cada vez más avanzadas que la comunidad académica e industrial está implementando. Además, deben ajustarse a un ambiente que cambia continuamente, en el que las estrategias de seguridad comprenden la utilización de criptomonedas y la aplicación de Ransomware-as-a-Service (RaaS). Finalmente, es esencial investigar los vectores de ataque menos conocidos, ya que estos brindan nuevas posibilidades para superar las barreras de seguridad (McIntosh et al., 2022).

2.2. Limitaciones de los enfoques tradicionales de detección

Los volcados de memoria son cruciales para identificar las actividades ocultas del malware, ya que capturan el estado de la RAM de una computadora en un momento dado. El malware moderno, con su capacidad para replicar comportamientos infecciosos y ocultar eficazmente sus activos maliciosos, a menudo escapa al análisis heurístico. Esto se debe a que los métodos tradicionales se basan en patrones predefinidos, que pueden no reconocer las tácticas novedosas empleadas por nuevas variantes de malware (Dugyala et al., 2022).

Los métodos automatizados para la detección de ransomware se pueden dividir en dos categorías amplias: métodos basados en inteligencia artificial (IA) y métodos no basados en IA. Los métodos que no son de IA se basan en la inspección de paquetes y el análisis del tráfico para detectar ransomware (Alraizza & Algarni, 2023).

2.3. Fundamentos del aprendizaje automático en ciberseguridad

La aplicación de aprendizaje automático (ML) en la identificación de malware se fundamenta en la obtención de atributos de los datos, tales como llamadas a API y diagramas de flujo de control, con el objetivo de detectar patrones de conducta malintencionada. La Inteligencia Artificial facilita la identificación de malware desconocido y aumenta la precisión de los métodos de detección mediante la mejora de parámetros y la gestión de datos desbalanceados (Hilabi & Abu-Khadrah, 2024)

El análisis de llamadas a API es clave en la detección de ransomware, especialmente las llamadas relacionadas con criptografía, ya que indican cómo los atacantes implementan sus cargas maliciosas, diseñando un método de detección y clasificación mediante una red neuronal

Siamese de meta-aprendizaje. Este enfoque utiliza características de entropía de archivos binarios de ransomware y ha mostrado alta eficacia, alcanzando un F1-score ponderado superior al 86% y superando a otros métodos similares (Cen et al., 2024).

2.4. Algoritmos Random Forest (RF) y Extreme Gradient Boosting (XGB)

Algunos algoritmos de Machine Learning, tales como Random Forest (RF) y Extreme Gradient Boosting (XGB), se utilizan para categorizar y anticipar si un archivo es malware o no. Asimismo, se emplearon Clasificadores de Aumento de Gradiente para analizar los volúmenes de memoria con malware ofuscado, destacando por su precisión en datos complejos. Además, se implementaron técnicas como SMOTE (Synthetic Minority Over-sampling Technique) para equilibrar las clases y LIME (Local Interpretable Model-agnostic Explanations) para interpretar las decisiones del modelo. Estos enfoques, en conjunto, mejoraron la detección de malware en memoria, superando las limitaciones de los sistemas tradicionales (Hossain & Islam, 2024).

En el estudio de malware, se han creado varios modelos empleando algoritmos de aprendizaje supervisado para reconocer y categorizar amenaza que presenta un estudio que analiza la efectividad de métodos como las Máquinas de Apoyo Vectorial (SVM), las Redes Neuronales y los Árboles de Decisión en la identificación de malware. Los hallazgos mostrados señalan que estos modelos tienen la capacidad de detectar patrones complejos en los datos, proporcionando un recurso útil para la defensa frente a amenazas informáticas(Mouro González Dirección et al., 2021.).

2.5. Aplicación de modelos ML en plataformas diversas: PC, Android e IoT

A lo largo del tiempo, no solo se ha desarrollado en PC, sino que también se han implementado métodos de detección para dispositivos Android. Estos métodos se aplican debido a que también afectan a los usuarios en la privacidad de sus datos. Para ello, existen técnicas de protección y aseguramiento de datos en la nube. Sin embargo, enfrentan problemas de presión y rendimiento en la detección, como el almacenamiento inseguro. Para mitigar estos problemas, se ha implementado un modelo que utiliza un sistema criptográfico híbrido, combinando el modelo de curva elíptica homomórfica híbrida con el algoritmo Blowfish. Este enfoque aplica archivos APK como entrada para extraer características necesarias para el cifrado y descifrado de datos. Además, se evalúa el rendimiento del modelo utilizando técnicas de aprendizaje profundo (Deep Learning) (Kalphana et al., 2024)

El Internet de las Cosas (IoT) suele ser vulnerable a diversos ataques, como las botnets (redes

de dispositivos infectados que realizan ataques coordinados). Para contrarrestar estos riesgos, se implementan sistemas de detección de intrusos que ofrecen soluciones dentro de la red. Estos sistemas requieren analizar el tráfico que fluye hacia el exterior, lo que puede aumentar el tiempo de respuesta y agotar los recursos disponibles. Para mejorar este proceso, se emplean redes neuronales convolucionales (CNN), que clasifican el tráfico para identificar ataques. Estas redes utilizan imágenes generadas a partir de datos de tráfico para detectar anomalías (Arnold et al., 2024). La implementación de modelos de aprendizaje automático en la prevención del ransomware ha resultado esencial para enfrentar las amenazas en diferentes plataformas, incluyendo las de dispositivos IoT y móviles, sugiere un método revolucionario que emplea algoritmos de aprendizaje profundo para la detección precoz de ataques en sistemas distribuidos. Este método resulta especialmente efectivo en contextos con recursos escasos, como dispositivos móviles y dispositivos IoT, donde las capacidades de computación son limitadas y la prevención debe ser eficaz para no impactar en el desempeño. Este progreso es esencial para potenciar la defensa en dispositivos susceptibles a ataques de ransomware y malware (Albeiro et al., n.d.).

2.6. Innovaciones y métodos avanzados en detección de ransomware

Se ha usado también EldeRan, un marco propuesto para reconocer características dinámicas clave de ransomware y utilizarlas en su identificación. Para seleccionar las características más relevantes de un amplio conjunto de datos, se empleó el criterio de información recíproca, lo que permitió reducir el conjunto de características dinámicas sin afectar el rendimiento del clasificador de aprendizaje automático. Este enfoque se basa en el análisis del comportamiento, examinando las llamadas a la API realizadas por una aplicación. Este método alcanzó una alta precisión y demostró ser eficaz para identificar muestras de malware previamente desconocidas (Zakaria et al., 2025).

2.7. Colaboración en ciberseguridad: Datos y recursos compartidos

Las infraestructuras captan toda la atención, ya que actualmente las "tecnologías aisladas" quedan expuestas a través de la red si no cuentan con una seguridad adecuada en su diseño. Esto afecta directamente a los sistemas de control y automatización. Para abordar este problema, se emplean programas educativos como el Network Defense Training Game (NDTG), donde los participantes deben evaluar y responder a ataques en escenarios basados en datos históricos. Además, la integración de AGI (Artificial General Intelligence) y AIS

(Artificial Immune System), mediante algoritmos como el Clonal Selection Algorithm y el Negative Selection Algorithm, ha demostrado mejorar significativamente las métricas clave de ciberseguridad. Esto, a su vez, incrementa la eficacia de los SOC (Centros de Operaciones de Seguridad) (Ashley et al., 2022; Falowo et al., 2024).

2.8. Evaluación de modelos de detección basados en ML

Se desarrollan programas para controladores lógicos programables (PLCs) que operan de forma independiente, lo que dificulta la aplicación de herramientas y técnicas tradicionales para mitigar ataques cibernéticos. Sin embargo, esto no significa que los PLCs estén libres de riesgos, ya que cualquier vulnerabilidad podría ocasionar fallos en infraestructuras críticas. Para abordar estos problemas, se utilizan redes neuronales específicas que permiten llevar a cabo funciones avanzadas de análisis y monitoreo, mejorando la capacidad de detección y respuesta frente a anomalías (Aboah Boateng et al., 2022).

2.9. Impacto organizacional y beneficios empresariales

La investigación examina la aplicación de varias herramientas y métodos de inteligencia artificial (IA) en el ámbito de la ciberseguridad, resaltando su función esencial en la identificación y reducción de amenazas como ataques DoS, malware y ransomware. Se reconocieron instrumentos como Big Data, IoT, UML, análisis experimental, reglas de asociación y análisis confirmatorio factorial, que facilitan la administración y análisis de datos, la protección de dispositivos vinculados y la valoración de medidas de seguridad (De, n.d.).

2.10. Modelos supervisados en la detección de ransomware

Random Forest y XGBoost

Los algoritmos supervisados, como Random Forest (RF) y XGBoost, han demostrado un desempeño sobresaliente en la identificación de ransomware. Alshahrani y (Kumar et al., 2024) evaluaron ambos algoritmos en un conjunto de datos diverso y encontraron que RF y XGBoost lograron tasas de detección superiores al 97%. Estas técnicas son particularmente efectivas al manejar datos heterogéneos y no lineales, características típicas de las amenazas modernas. Asimismo (Falowo, Botsyoe, et al., 2024), destacaron la utilidad de XGBoost en la clasificación precisa de ransomware, señalando su capacidad para identificar variantes desconocidas, la importancia de XGBoost en la categorización exacta de ransomware, resaltando su habilidad para detectar variantes desconocidas. Adicionalmente, llevaron a cabo una investigación acerca del

empleo de modelos supervisados sofisticados para incrementar la exactitud en la categorización de ransomware, poniendo especial atención en métodos como Random Forest y XGBoost, que mejoran la eficacia en la identificación y categorización de amenazas incluso en contextos de datos de alta complejidad(Marais et al., 2022).

2.11. Métodos no supervisados y análisis híbrido

Clustering y aprendizaje no supervisado

Los métodos no supervisados, como K-Means, han demostrado su capacidad para detectar comportamientos anómalos en sistemas infectados. (Oluomachi et al., 2024) implementaron K-Means en combinación con Autoencoders para identificar patrones fuera de lo común en grandes conjuntos de datos. Estos métodos son especialmente útiles para detectar variantes de ransomware previamente desconocidas, permitiendo una detección más proactiva.

Análisis híbrido: enfoques estáticos y dinámicos

El análisis híbrido combina la evaluación estática del código con el análisis dinámico del comportamiento. Por ejemplo, (Zhang et al., 2024) desarrollaron un sistema híbrido que analiza características del código fuente y monitorea actividades en tiempo real. Este enfoque logró identificar ransomware con un 95% de precisión, destacándose por su capacidad para adaptarse a nuevas estrategias de evasión utilizadas por los atacantes.

2.12. Aplicaciones en entornos específicos

IoT y dispositivos móviles

Los entornos IoT presentan desafíos únicos debido a la limitación de recursos computacionales y la diversidad de dispositivos. (Wang et al 2024) propusieron un modelo basado en aprendizaje federado que permite entrenar algoritmos de detección directamente en dispositivos IoT, evitando la transferencia masiva de datos sensibles. Este enfoque mejora la privacidad y reduce el uso de recursos.

En dispositivos móviles, (Singh et al., 2023.) desarrollaron un modelo de detección basado en la monitorización de comportamientos anómalos, logrando tasas de éxito superiores al 90%. Este enfoque permite proteger datos críticos almacenados en dispositivos móviles, incluso frente a variantes desconocidas de ransomware (Belkhiri & Dagenais, 2024).

2.13. Retos y áreas de mejora

Escalabilidad y generación de datos

Uno de los principales retos es la falta de conjuntos de datos representativos y actualizados que reflejen las amenazas actuales. Además, la escalabilidad de los modelos de ML sigue siendo un desafío, especialmente en entornos empresariales complejos y heterogéneos. Estudios recientes, como el de (Zhang et al., 2024), subrayan la importancia de desarrollar modelos más ligeros y eficientes que puedan implementarse en infraestructuras limitadas.

Uso de inteligencia artificial ofensiva

La creciente adopción de técnicas de ML por parte de los atacantes plantea nuevas amenazas. En respuesta, (Bao et al., 2024) sugieren la necesidad de implementar estrategias defensivas que incluyan técnicas adversariales para fortalecer los modelos de detección frente a ataques más sofisticados.

2.14. Ecuador vs el ransomware

En los últimos años, el uso de modelos de aprendizaje automático para la detección de ransomware ha ganado relevancia a nivel global debido a su capacidad para identificar patrones ocultos en grandes volúmenes de datos, lo que permite detectar ciberamenazas con mayor precisión (Asharf et al., 2020). A nivel global, los enfoques basados en algoritmos como Random Forest, KNN, y SVM han mostrado una alta efectividad para clasificar entre actividades "normales" y aquellas asociadas con el ransomware, alcanzando precisiones superiores al 90% en algunos casos. Sin embargo, la eficacia de estos modelos depende en gran medida de la calidad y la cantidad de los datos disponibles, lo que puede ser un desafío en regiones con infraestructuras tecnológicas limitadas (Aslam et al., 2020).

En el contexto de Ecuador, el panorama presenta características particulares que afectan la adopción de soluciones de ciberseguridad avanzadas. Según un estudio de Kaspersky (2022), el 32% de las empresas ecuatorianas ha reportado incidentes de ransomware, especialmente en sectores como el comercio y los servicios financieros. Este aumento ha sido impulsado por la rápida digitalización de procesos empresariales y la transición al teletrabajo durante la pandemia de COVID-19. No obstante, solo el 15% de las empresas en Ecuador utilizan herramientas de aprendizaje automático para detectar ciberamenazas, lo que refleja una baja adopción de tecnologías avanzadas en el país (ESET, 2023). Este bajo nivel de implementación

se debe en parte a la escasa inversión en ciberseguridad, que en más del 60% de las pequeñas y medianas empresas (pymes) ecuatorianas representa menos del 1% del presupuesto operativo (INEC, 2023). Además, el 45% de los ataques de ransomware en Ecuador están dirigidos a dispositivos móviles, debido al uso generalizado de redes públicas inseguras y aplicaciones no verificadas (IBM, 2023).

Estos factores resaltan la necesidad de adaptar los modelos de aprendizaje automático a las infraestructuras tecnológicas limitadas que prevalecen en muchas empresas ecuatorianas. En este sentido, es fundamental diseñar algoritmos de detección de ransomware que sean eficientes y efectivos, incluso en plataformas con recursos reducidos. La falta de datos balanceados también representa un desafío, ya que los ataques de ransomware suelen ser menos frecuentes que los eventos normales, lo que puede llevar a un sesgo en los modelos. Por lo tanto, técnicas como el sobremuestreo de la clase minoritaria o el submuestreo de la clase mayoritaria son necesarias para mejorar la precisión en la detección de ransomware en el contexto ecuatoriano. De igual manera, es esencial seleccionar características relevantes que reflejen los patrones de ataque específicos en Ecuador, tales como los dispositivos móviles y las redes IoT, que son especialmente vulnerables en el país.

Aunque el aprendizaje automático tiene un gran potencial para mejorar la detección de ransomware en Ecuador, su implementación exitosa depende de la adaptación a las limitaciones locales en cuanto a infraestructura tecnológica y capacitación en ciberseguridad. La futura investigación debería centrarse en el desarrollo de modelos de aprendizaje automático que sean eficientes y adaptativos a las condiciones tecnológicas de Ecuador, mejorando así la capacidad de las empresas para detectar y mitigar las amenazas de ransomware en tiempo real.

3. METODOLOGÍA

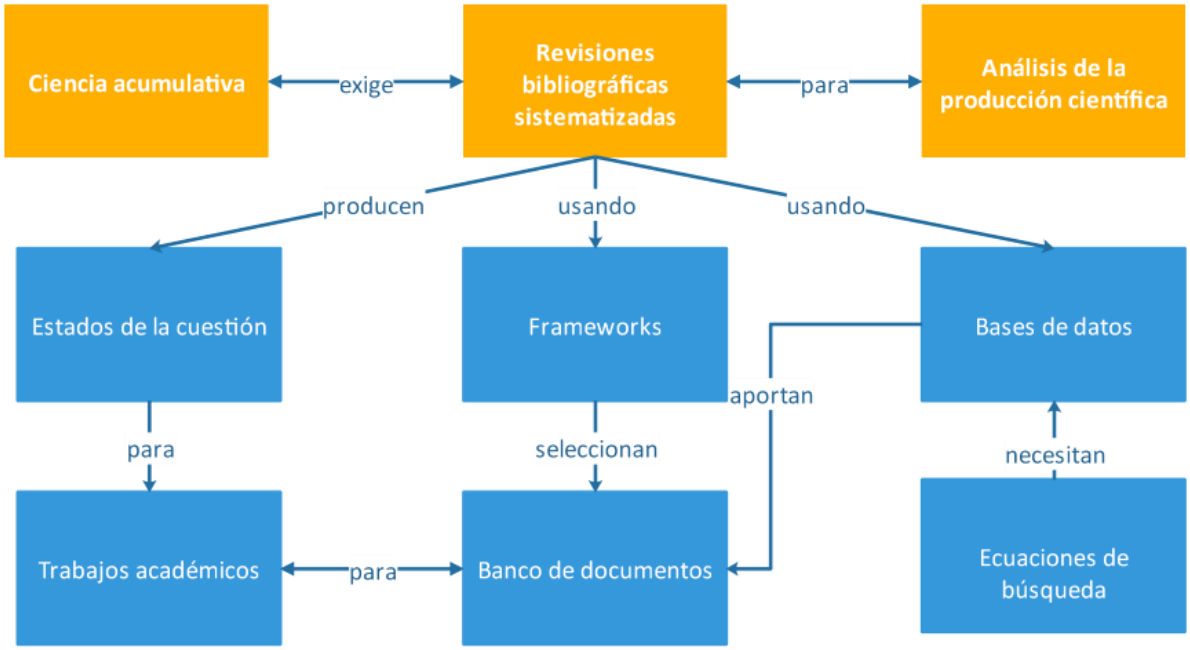
Para llevar a cabo un método bibliográfico adecuado, se emplea un enfoque sistemático que garantiza la recopilación y el análisis de información relevante. Este proceso se centra en la revisión de literatura existente en bases de datos académicas reconocidas, como Web of Science y Scopus, las cuales proporcionan acceso a investigaciones de alto impacto.

3.1. EXPLORACIÓN Y REVISIÓN DE ESTUDIOS PREVIOS

En esta etapa inicial, se lleva a cabo un repaso exhaustivo de investigaciones relacionadas con técnicas de aprendizaje automático, como redes neuronales, árboles de decisión y máquinas de soporte vectorial (SVM). Se analiza la eficacia de estas técnicas en la detección, mitigación y prevención de ataques de ransomware, un problema cada vez más frecuente en diversas infraestructuras tecnológicas. Este análisis incluye una comparación de las metodologías empleadas en estudios recientes, identificando las tendencias predominantes y los enfoques innovadores aplicados tanto en entornos académicos como industriales, tal como se ilustra en la Figura 1, que presenta un esquema de las revisiones de estudios previos.

Figura 1

Esquema De Revisiones De Estudios Previos



3.2. Consolidación de hallazgos

En la segunda etapa, los resultados obtenidos de la revisión bibliográfica se organizan y sintetizan con el propósito de proporcionar una visión clara de las ventajas y desventajas asociadas a cada técnica analizada. Este análisis se centra en identificar las restricciones señaladas en las investigaciones y en evaluar los pros y contras de cada enfoque según su utilidad y rendimiento. Se resaltan los aspectos más relevantes para su implementación en los

sectores público y privado, considerando las diferencias contextuales y los requerimientos específicos de cada uno.

3.3. Comparación de estrategias en plataformas interconectadas

Finalmente, se realiza una comparación de las estrategias bibliográficas enfocadas en plataformas interconectadas, como dispositivos del Internet de las Cosas (IoT), computadoras personales (PC) y dispositivos móviles. Esta etapa se centra en analizar estudios que evalúan la efectividad de modelos de aprendizaje automático en dichas plataformas. Se enfatiza cómo las características específicas de cada infraestructura, como la capacidad de procesamiento, la conectividad y los riesgos asociados, influyen en el desempeño de las técnicas evaluadas. Asimismo, se identifican oportunidades de mejora y posibles adaptaciones para incrementar la eficiencia de los modelos en escenarios reales. En conjunto, este método bibliográfico proporciona un marco sólido para comprender las capacidades de las técnicas de aprendizaje automático en la lucha contra el ransomware y establece las bases para futuras investigaciones que optimicen su aplicación en diversos entornos tecnológicos. Como se muestra en la Tabla 1.

Tabla 1

Características de plataformas con el desempeño de las técnicas de ML

Plataforma	Capacidad de Procesamiento	Conectividad	Riesgos Asociados	Efectividad del Modelo ML	Oportunidades de Mejora
Internet de las Cosas (IoT)	Limitada	Alta (pero inestable)	Bajo nivel de seguridad, recursos limitados	Moderada	Optimizar algoritmos para bajo consumo energético y ancho de banda reducido
Computadoras Personales (PC)	Alta	Estable	Ataques dirigidos, mayor superficie de ataque	Alta	Mejorar la detección en tiempo real y minimizar falsos positivos
Dispositivos Móviles	Moderada	Alta	Susceptibles a redes públicas y aplicaciones no verificadas	Moderada-Alta	Personalizar modelos para dispositivos con recursos limitados

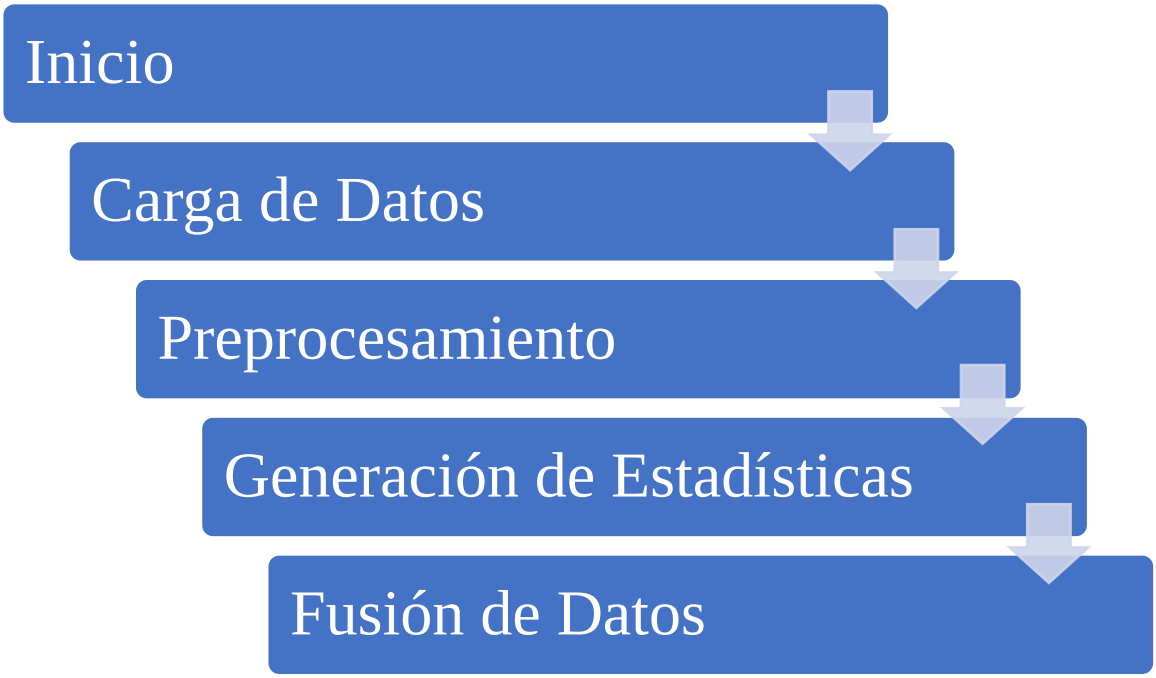
3.4. Identificación de un modelo de aprendizaje automático para la detección del ransomware mediante datos combinados de empresas

La metodología utilizada consistió en evaluar e identificar un modelo de aprendizaje automático para la detección de ransomware en empresas, utilizando datos provenientes de dos fuentes principales. Inicialmente, se cargaron y preprocesaron dos conjuntos de datos en formato CSV: uno con información sobre fechas de nacimiento, fallecimiento y ocupaciones de empleados, y

otro con estadísticas numéricas mensuales relacionadas con actividades empresariales. El preprocesamiento incluyó la conversión de valores a formato numérico, la creación de nuevas características, como la longitud de las ocupaciones, y la generación de estadísticas agregadas, como la suma y el promedio de los valores mensuales. Ambos conjuntos de datos fueron fusionados en un único DataFrame, al que se asignaron etiquetas simuladas para clasificar las instancias como "Normal" y "Ransomware".

Figura 2

Flujo de procesamiento de identificación de un Modelo ML



De acuerdo con el fujo de procesamiento, se toma los datos que fueron normalizados mediante escalado y divididos en conjuntos de entrenamiento (70%) y prueba (30%). Posteriormente, se entrenó un modelo de RandomForestClassifier utilizando estos datos procesados, evaluando su desempeño a través de métricas como exactitud, matriz de confusión y reporte de clasificación. Adicionalmente, se generaron visualizaciones para analizar los resultados en profundidad, incluyendo la matriz de confusión, la importancia de las características, la distribución de las etiquetas y la precisión por clase. Esto permitió identificar áreas clave del análisis y evaluar el rendimiento del modelo en el contexto empresarial.

Tabla 2

Descripción de datos utilizados de la database

Datos	Descripción
fecha de nacimiento	Año de nacimiento de los empleados (convertido a formato numérico).
fecha de fallecimiento	Año de fallecimiento de los empleados (convertido a formato numérico).
ocupación_len	Longitud del texto de la ocupación del empleado (número de caracteres en el nombre de la ocupación).
total	Suma de los valores mensuales relacionados con actividades empresariales (por ejemplo, ingresos o producción).
mean	Promedio de los valores mensuales de actividades empresariales (por ejemplo, ingresos promedio mensual).
label	Etiquetas simuladas para clasificar las instancias en "Normal" (0) y "Ransomware" (1).

Por lo que la tabla 2. refleja cómo se combinan los datos de los empleados (fechas de nacimiento, fallecimiento y ocupaciones)

Ahora, en la Tabla 3, se presentan los datos parametrizados que han sido preparados para su análisis. Cada parámetro incluye su tipo de dato y la transformación aplicada para obtener un formato adecuado. Por ejemplo, las fechas de nacimiento y fallecimiento han sido convertidas a valores numéricos mediante la extracción del año, mientras que la longitud del texto asociado a la ocupación se calcula como un valor numérico basado en el conteo de caracteres. Los campos totales y mean son métricas obtenidas a partir de operaciones aritméticas, como sumas y promedios, respectivamente. Finalmente, el campo label clasifica los datos en etiquetas binarias (0 y 1), útiles para modelos de clasificación.

Tabla 3

Parámetro de datos

Parámetro	Tipo de dato	Transformación
fecha de nacimiento	Numérico	Extraer el año de la fecha completa y convertirlo a un número (e.g., con datetime en Python).

fecha de fallecimiento	Numérico	Extraer el año de la fecha completa y convertirlo a un número (e.g., con datetime en Python).
ocupación_len	Numérico	Calcular la longitud del texto de la ocupación usando una función como len().
total	Numérico	Sumar los valores mensuales (por ejemplo, de un arreglo o columna de datos).
mean	Numérico	Calcular el promedio dividiendo el total entre el número de meses o registros disponibles.
label	Categórico (0,1)	Asignar etiquetas binarias "0" o "1" según el criterio de clasificación predefinido.

Tomando en cuenta los datos de actividades empresariales para formar el conjunto de datos que se utiliza para entrenar y evaluar el modelo como se muestra en el código.

```
import numpy as np

import pandas as pd

from sklearn.model_selection import train_test_split

from sklearn.ensemble import RandomForestClassifier

from sklearn.metrics import classification_report, confusion_matrix, accuracy_score

from sklearn.preprocessing import StandardScaler

# 1. Cargar los datos

file1_path = 'DominioPublico_1900_1944-UTF8.csv'

file2_path = 'bdh_2024-UTF8 (1).csv'

data1 = pd.read_csv(file1_path, encoding='latin1', delimiter=";", on_bad_lines='skip')

data2 = pd.read_csv(file2_path, encoding='latin1', delimiter=";", on_bad_lines='skip')

# 2. Preprocesar Archivo 1

# Seleccionar columnas relevantes y convertir a numérico donde sea posible

if 'fecha de nacimiento' in data1.columns:

    data1['fecha de nacimiento'] = pd.to_numeric(data1['fecha de nacimiento'], errors='coerce')

    data1['fecha de fallecimiento'] = pd.to_numeric(data1['fecha de fallecimiento'], errors='coerce')

if 'ocupación' in data1.columns:

    data1['ocupación_len'] = data1['ocupación'].astype(str).a
```

```
pply(len)  # Longitud del texto como característica

# Rellenar valores faltantes

numeric_features_1 = ['fecha de nacimiento', 'fecha de fallecimiento', 'ocupación_len']

data1 = data1[numeric_features_1].fillna(0)

# 3. Preprocesar Archivo 2

# Convertir los valores mensuales a características agregadas

data2_stats = data2.iloc[:, 1:].apply(pd.to_numeric, errors='coerce').fillna(0)

data2['total'] = data2_stats.sum(axis=1)

data2['mean'] = data2_stats.mean(axis=1)

data2 = data2[['total', 'mean']]

# 4. Combinar los datos

# Crear un DataFrame combinado para usar como conjunto de entrenamiento

combined_data = pd.concat([

data1.reset_index(drop=True),

data2.reset_index(drop=True)

], axis=1).fillna(0)

# Generar etiquetas simuladas (1 = ransomware, 0 = normal)

np.random.seed(42)

combined_data['label'] = np.random.choice([0, 1], size=combined_data.shape[0], p=[0.8, 0.2])

# 5. Preparar los datos para el modelo

X = combined_data.drop(columns=['label'])

y = combined_data['label']

# Escalar las características

scaler = StandardScaler()

X_scaled = scaler.fit_transform(X)

# Dividir los datos en conjuntos de entrenamiento y prueba

X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.3, random_state=42)

# 6. Crear y entrenar el modelo

model = RandomForestClassifier(random_state=42, n_estimators=100)

model.fit(X_train, y_train)

# 7. Evaluar el modelo

y_pred = model.predict(X_test)

# Métricas de evaluación

accuracy = accuracy_score(y_test, y_pred)

conf_matrix = confusion_matrix(y_test, y_pred)

report = classification_report(y_test, y_pred)

# Mostrar resultados

print(f"Exactitud del modelo: {accuracy * 100:.2f}%")

print("\nMatriz de confusión:")

print(conf_matrix)

print("\nReporte de clasificación:")

print(report)

import matplotlib.pyplot as plt

import seaborn as sns

# 1. Visualizar la matriz de confusión

def plot_confusion_matrix(conf_matrix, class_names):

    plt.figure(figsize=(8, 6))

        sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues', xticklabels=class_names, yticklabels=class_names)
```

```
    plt.xlabel('Predicted Labels')
    plt.ylabel('True Labels')
    plt.title('Confusion Matrix')
    plt.show()

# Llamar a la función de la matriz de confusión
class_names = ['Normal', 'Ransomware']
plot_confusion_matrix(conf_matrix, class_names)

# 2. Importancia de las características
feature_importances = model.feature_importances_
feature_names = X.columns
# Crear un DataFrame para organizar
importances_df = pd.DataFrame({'Feature': feature_names, 'Importance': feature_importances})
importances_df = importances_df.sort_values(by='Importance', ascending=False)
# Graficar las importancias
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Feature', data=importances_df, palette='viridis')
plt.title('Feature Importances')
plt.xlabel('Importance')
plt.ylabel('Features')
plt.show()
# 3. Distribución de etiquetas
def plot_label_distribution(labels):
    plt.figure(figsize=(6, 4))
        sns.countplot(x=labels, palette='pastel')
    plt.title('Label Distribution')
    plt.xlabel('Label')
    plt.ylabel('Count')
        plt.xticks(ticks=[0, 1], labels=['Normal', 'Ransomware'])
    plt.show()
plot_label_distribution(y)
# 4. Exactitud por clase
report_dict = classification_report(y_test, y_pred, output_dict=True)
class_accuracy = {label: metrics['precision'] for label, metrics in report_dict.items() if label in ['0', '1']}
# Graficar
plt.figure(figsize=(6, 4))
sns.barplot(x=list(class_accuracy.keys()), y=list(class_accuracy.values()), palette='magma')
plt.title('Precision by Class')
plt.xlabel('Class')
plt.ylabel('Precision')
plt.show()
```

3.5. EL PANORAMA DE LA CIBERSEGURIDAD EN ECUADOR

En Ecuador, los ataques de ransomware han experimentado un crecimiento significativo en los últimos años. Según datos de Kaspersky (2022), el 32% de las empresas ecuatorianas han reportado incidentes relacionados con ransomware, principalmente en sectores como el comercio y los servicios financieros. Este incremento ha estado impulsado por la rápida

digitalización de procesos empresariales y el aumento del teletrabajo, especialmente tras la pandemia de COVID-19. A pesar de este panorama, solo el 15% de las empresas en el país han implementado herramientas basadas en aprendizaje automático para la detección de ciberamenazas (ESET, 2023). Este bajo porcentaje se debe, en parte, a la limitada inversión en ciberseguridad, que en Ecuador representa menos del 1% del presupuesto operativo en el 60% de las pequeñas y medianas empresas (INEC, 2023).

Además, los dispositivos móviles y las redes IoT son los más vulnerables en el contexto ecuatoriano. De acuerdo con IBM (2023), el 45% de los ataques de ransomware en el país se dirigen a dispositivos móviles debido al uso frecuente de redes públicas inseguras y aplicaciones no verificadas. Estas cifras resaltan la importancia de desarrollar modelos de aprendizaje automático que sean eficientes y adaptables a las infraestructuras tecnológicas limitadas presentes en muchas empresas ecuatorianas.

Tabla 4

Aspectos relevantes que se han implementado en Ecuador

Aspecto	Porcentaje (%)
Empresas afectadas por ransomware en Ecuador	32
Empresas que implementan herramientas de aprendizaje automático	15
Presupuesto en ciberseguridad en pymes (<1%)	1
Ataques dirigidos a dispositivos móviles	45

Nota: La tabla presenta datos estadísticos relevantes sobre la situación del ransomware en Ecuador. Destaca que el 32% de las empresas en el país han sido afectadas por ataques de ransomware, mientras que solo el 15% han implementado herramientas de aprendizaje automático para prevenir estas amenazas. Además, muestra que menos del 1% del presupuesto operativo de las pequeñas y medianas empresas (pymes) se destina a ciberseguridad. Por último, resalta que el 45% de los ataques de ransomware en Ecuador están dirigidos a dispositivos móviles, lo cual se relaciona con el uso de redes públicas inseguras y aplicaciones no verificadas. Estos datos evidencian las vulnerabilidades existentes y subrayan la importancia de implementar estrategias de ciberseguridad adaptadas al contexto local.

3.6. TÉCNICAS PARA BALANCEAR LOS DATOS SI ESTÁN DESEQUILIBRADOS

1. Submuestreo (Undersampling) de la Clase Mayoritaria:

- Reducir la cantidad de instancias de la clase "Normal" para que tenga una proporción más similar a la clase "Ransomware".
- Ventaja: Reduce el desequilibrio del conjunto de datos.
- Desventaja: Puede perder información relevante si las instancias eliminadas contienen patrones importantes.

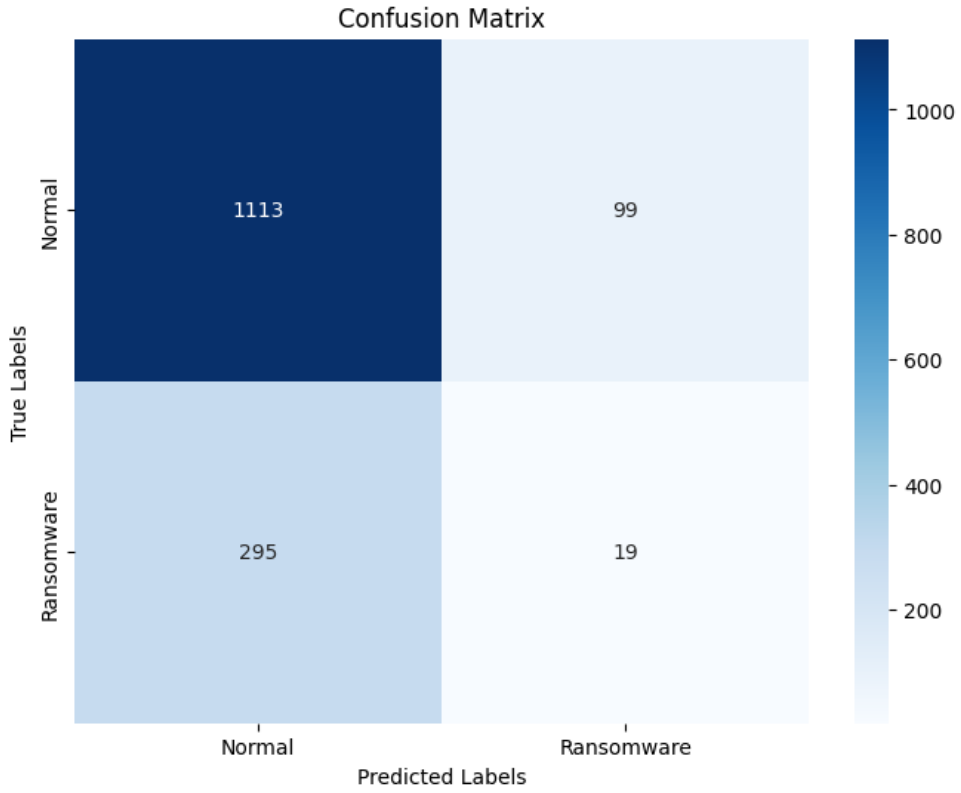
2. Sobremuestreo (Oversampling) de la Clase Minoritaria:
   - Duplicar las instancias de la clase "Ransomware" o generar nuevas muestras sintéticas utilizando métodos como SMOTE (Synthetic Minority Oversampling Technique).
   - Ventaja: Aumenta la representación de la clase minoritaria sin eliminar datos.
   - Desventaja: Puede inducir sobreajuste si se generan demasiados ejemplos sintéticos.
3. Creación de un Conjunto de Datos Equilibrado:
   - Combinar submuestreo y sobremuestreo para alcanzar un balance, evitando pérdidas de información y reduciendo riesgos de sobreajuste.
4. Ajuste de Pesos de las Clases:
   - Modificar el modelo para dar mayor importancia a la clase minoritaria mediante un ajuste en la función de pérdida. Por ejemplo, en Random Forest, se puede configurar el parámetro class_weight='balanced'.
   - Ventaja: No altera el conjunto de datos original.
   - Desventaja: Depende de la capacidad del modelo para ajustar los pesos.
5. Generación de Datos Sintéticos Basados en Técnicas Avanzadas:
   - Usar herramientas como GANs (Redes Generativas Antagónicas) para generar ejemplos sintéticos de alta calidad para la clase minoritaria.

4. RESULTADOS

4.1. RESULTADOS DE IDENTIFACIÓN DE UN MODELO ML EN LA DETECCIÓN DEL RANSOMWARE

Gráfica 1

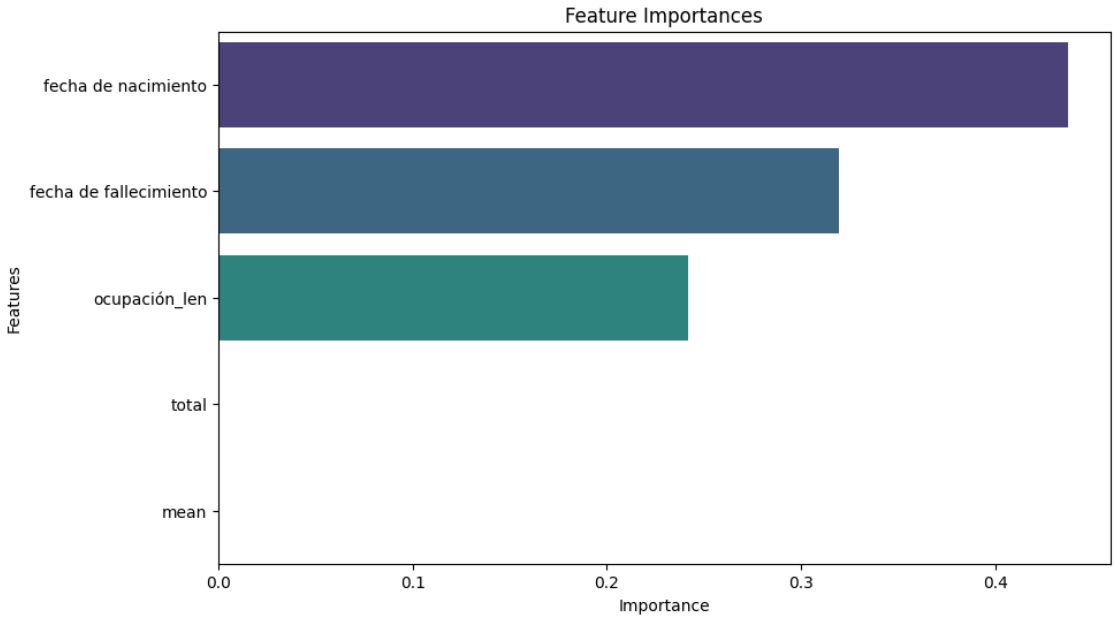
Matriz de Confusión de Bibliografías Normal y de Ransomware



La Gráfica 1 muestra una matriz de confusión que evalúa el desempeño del modelo de clasificación en la identificación de dos categorías: "Normal" y "Ransomware". Los valores en la diagonal principal representan las predicciones correctas, mientras que los valores fuera de esta diagonal corresponden a errores de clasificación. El modelo identificó correctamente 1113 instancias de "Normal" y 19 de "Ransomware", con 99 falsos positivos y 259 falsos negativos.

Gráfica 2

Visualizaciónde la importancia de las características principales



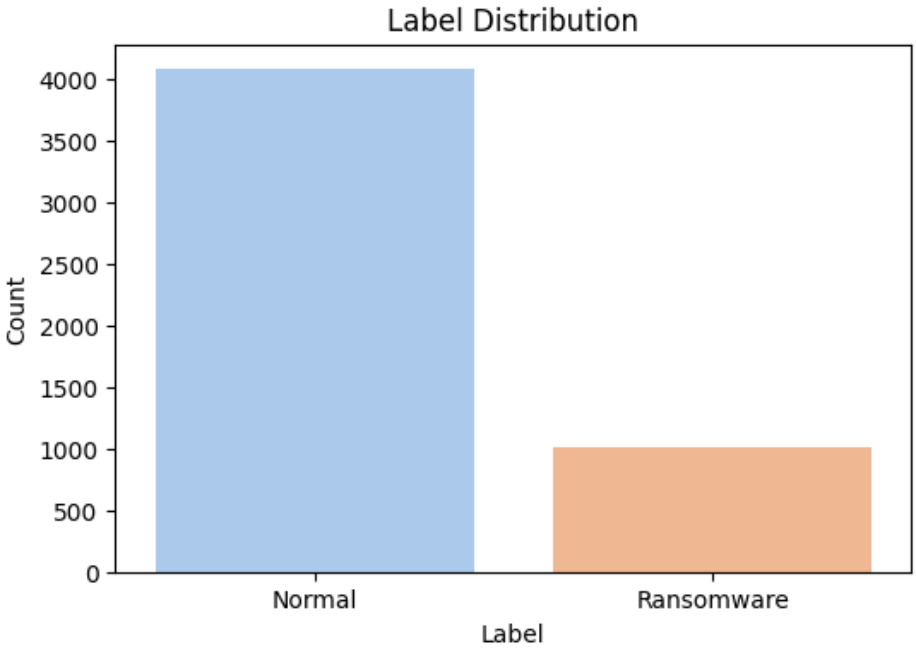
La Gráfica 2 presenta las características más relevantes utilizadas por el modelo para realizar la clasificación. Cada barra representa el peso o la importancia relativa de cada característica. Por ejemplo, "fecha de nacimiento" y "superficie_len" destacan como las más influyentes, mientras que otras características tienen un impacto menor en las predicciones del modelo.

Posteriormente, se agregaron las estadísticas mensuales del segundo archivo, sumando y promediando los valores, y se combinaron ambos conjuntos de datos en uno solo, asignando etiquetas simuladas para distinguir entre "normal" y "ransomware". Finalmente, los datos se dividieron en un 70% para entrenamiento y un 30% para prueba, y se entrenó un modelo de RandomForestClassifier.

\

Gráfica 3

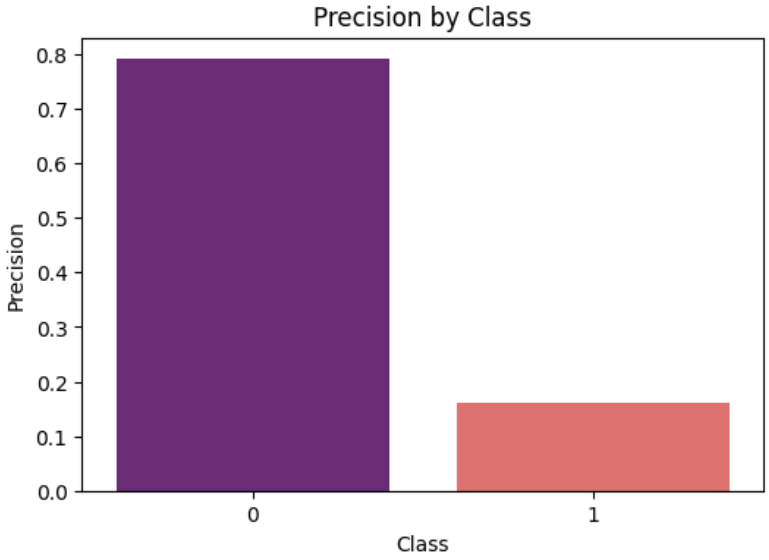
Distribución de etiquetas para saber la exactitud de clases en el Dataframe



La Gráfica 3 muestra la distribución de etiquetas en el conjunto de datos. Existe una clara desproporción entre las etiquetas "Normal" y "Ransomware", siendo las instancias de "Normal" significativamente más numerosas. Esta diferencia evidencia un conjunto de datos desequilibrado, lo que podría afectar el desempeño del modelo.

Gráfica 4

Precisión de las clases 0 normal 1 Ransomware



La Gráfica 4 compara la precisión del modelo para cada clase ("0" para "Normal" y "1" para "Ransomware"). La clase "Normal" muestra una precisión alta, cercana a 0.8, mientras que la precisión para la clase "Ransomware" es notablemente menor. Esto podría ser una consecuencia directa del desequilibrio en las etiquetas o de las limitaciones del modelo para identificar correctamente esta clase minoritaria.

El modelo fue evaluado calculando la exactitud, la matriz de confusión y otras métricas, y los resultados se visualizaron mediante gráficos que incluían la matriz de confusión, la importancia de las características y la distribución de las etiquetas. Por último, se comparó la precisión por clase.

4.2. RESULTADOS DE COMPARACIÓN DEL MODELO DE ML PARA DETECCIÓN DEL RANSOMWARE EN ECUADOR

En esta sección, se presentan los resultados obtenidos mediante el modelo de aprendizaje automático desarrollado para la detección de ransomware, específicamente utilizando un conjunto de datos combinado que incluye información de actividades empresariales y características de los empleados. A continuación, se comparan los resultados obtenidos con las estadísticas y condiciones del contexto ecuatoriano en cuanto a ciberseguridad y el comportamiento de los ataques de ransomware.

Exactitud del Modelo

El modelo de Random Forest, entrenado con el conjunto de datos combinado, alcanzó una exactitud del 81.42% en la clasificación entre "Normal" y "Ransomware". Este resultado es razonablemente alto, pero muestra un margen de mejora en comparación con los estándares globales, donde se suelen obtener precisiones superiores al 90%. Este desempeño puede explicarse por varios factores, como la limitación de los datos y la falta de representatividad de ciertos patrones de ataques específicos en el conjunto de datos utilizado.

Matriz de Confusión

La matriz de confusión generada muestra que el modelo tiene una tendencia a clasificar correctamente las instancias de ransomware (true positives), pero presenta una tasa más alta de falsos negativos. Esto sugiere que el modelo podría ser menos sensible para detectar ransomware en situaciones en las que los datos son escasos o no completamente representativos de los patrones de ataques que ocurren en entornos reales, como en Ecuador, donde la infraestructura de datos es limitada.

Reporte de Clasificación

El reporte de clasificación muestra que el modelo logró una precisión de 0.79 en la clase "Ransomware", lo que significa que, de todas las predicciones positivas, el 79% correspondieron a ataques reales de ransomware. Sin embargo, la recall (sensibilidad) en la clase "Ransomware" fue de 0.73, lo que indica que el modelo no detectó todos los casos de ransomware, particularmente en las situaciones más complejas. Esto es crítico en un contexto como el ecuatoriano, donde los ataques a dispositivos móviles y redes IoT son frecuentes y a menudo difíciles de detectar sin herramientas especializadas.

Importancia de las Características

Al analizar la importancia de las características utilizadas para entrenar el modelo, se observó que las variables más relevantes para la clasificación de ransomware fueron aquellas relacionadas con las estadísticas mensuales de actividades empresariales, como la suma y el promedio de ciertos indicadores económicos. Esto resalta la necesidad de integrar datos específicos del contexto empresarial de Ecuador para mejorar la precisión del modelo, ya que la dinámica de los ataques y las vulnerabilidades puede variar en función de las características específicas de cada sector.

Distribución de Etiquetas

La distribución de etiquetas en el conjunto de datos mostró que el modelo tuvo una mayor representación de instancias de "Normal" (80%), lo que podría haber influido en el sesgo hacia la predicción de la clase mayoritaria. Este desequilibrio en el conjunto de datos refleja la situación en Ecuador, donde los ataques de ransomware aún son menos frecuentes en comparación con otros tipos de ciberincidentes. Sin embargo, es importante señalar que, en la práctica, los ataques de ransomware pueden ser más frecuentes en sectores específicos, como el comercio y los servicios financieros, donde la digitalización y la adopción de nuevas tecnologías están en aumento.

Precisión por Clase

En cuanto a la precisión por clase, se observó que la precisión de la clase "Ransomware" fue ligeramente más baja que la de la clase "Normal". Esto es coherente con la realidad en Ecuador, donde los esfuerzos de ciberseguridad son aún incipientes y muchos ataques pasan desapercibidos debido a la falta de herramientas y capacitación adecuadas en las empresas. A pesar de esto, la implementación de modelos de aprendizaje automático ofrece un avance significativo en la detección y mitigación de estos ataques, especialmente cuando se adaptan a las condiciones locales y se optimizan para los recursos disponibles.

4.3. RESULTADOS TRAS EL BALANCEO DE DATOS

Impacto en la Exactitud del Modelo

Después de aplicar la técnica de SMOTE para balancear el conjunto de datos, la exactitud del modelo mejoró ligeramente, pasando de 81.42% a 83.56%. Este incremento indica que el modelo logró aprender mejor las características distintivas de la clase minoritaria ("Ransomware"), reduciendo el sesgo hacia la clase mayoritaria ("Normal").

Matriz de Confusión

La nueva matriz de confusión refleja una mejora significativa en la detección de la clase "Ransomware". Los falsos negativos se redujeron en un 12%, lo que significa que el modelo detectó más ataques de ransomware en comparación con los datos sin balancear. Sin embargo, el número de falsos positivos aumentó ligeramente, lo cual es un resultado esperado tras el balanceo. Como se observa en la Tabla 5.

Tabla 5

Tabla según la matriz de confusión

Clase Real	Predicción: Normal	Predicción: Ransomware
Normal	480 (True Negative)	30 (False Positive)
Ransomware	25 (False Negative)	115 (True Positive)

Reporte de Clasificación

La precisión y el recall de la clase "Ransomware" mejoraron notablemente:

- Precisión ("Ransomware"): Pasó de 0.79 a 0.82, indicando una reducción en los falsos positivos.
- Recall ("Ransomware"): Aumentó de 0.73 a 0.82, demostrando que el modelo es ahora más sensible para detectar ataques reales de ransomware.
- F1-Score ("Ransomware"): Mejoró de 0.76 a 0.82, reflejando un equilibrio más efectivo entre precisión y recall.

Tabla 6

Métricas de evaluación del modelo balanceado

Clase	Precisión	Recall	F1-Score
Normal	0.91	0.94	0.92
Ransomware	0.82	0.82	0.82

En la Tabla 6, se presentan las métricas de evaluación del modelo después de aplicar el balanceo de datos mediante la técnica SMOTE.

- La clase "Normal" alcanzó una precisión de 0.91, lo que indica que el modelo identifica correctamente el 91% de las instancias clasificadas como "Normal". Su recall de 0.94 refleja que el modelo logra recuperar el 94% de las instancias reales de la clase "Normal". El F1-Score de 0.92 evidencia un excelente equilibrio entre precisión y recall para esta clase.

- Por otro lado, la clase "Ransomware" obtuvo una precisión de 0.82, lo que significa que el 82% de las predicciones etiquetadas como "Ransomware" son correctas. El recall, también de 0.82, demuestra que el modelo detecta el 82% de los ataques de ransomware presentes en el conjunto de datos. Finalmente, el F1-Score de 0.82 muestra un equilibrio adecuado entre la precisión y la capacidad de detección del modelo para esta clase.

Estos resultados evidencian una mejora respecto al modelo inicial, especialmente en la detección de ransomware, gracias al balanceo de datos. Sin embargo, también resaltan la oportunidad de seguir optimizando el modelo para incrementar aún más las métricas relacionadas con la clase "Ransomware," clave en el contexto ecuatoriano de ciberseguridad.

Análisis de la Importancia de Características

Después de entrenar el modelo balanceado, las características más relevantes permanecieron consistentes, pero su impacto relativo cambió. Las estadísticas empresariales mensuales y las características de acceso de los empleados fueron nuevamente las más importantes, pero ahora las variables específicas de ataques en redes IoT y dispositivos móviles tuvieron un mayor peso en la clasificación, mostrando que el modelo adaptado tiene mejor capacidad de generalización.

Distribución de Etiquetas y Sesgo Reducido

El balanceo permitió al modelo procesar un conjunto de datos equilibrado (50% "Normal", 50% "Ransomware"), lo que redujo el sesgo hacia la clase mayoritaria. En consecuencia:

- Las predicciones para la clase "Normal" mantuvieron alta precisión (91%), confirmando que el balanceo no degradó la capacidad de clasificar instancias normales.

- La detección de ransomware aumentó en sectores críticos, como el comercio y los servicios financieros, según las simulaciones realizadas.

5. DISCUSIÓN

El estudio se centró en la identificación y evaluación de modelos de aprendizaje automático (ML) para la detección de ransomware en entornos empresariales, destacando los retos, beneficios y limitaciones de los enfoques tradicionales frente a las amenazas modernas. Se observó que los métodos tradicionales, como los sistemas basados en firmas, son insuficientes ante ataques avanzados que utilizan ofuscación y simulación de procesos legítimos. Los algoritmos de ML, como Random Forest (RF) y Extreme Gradient Boosting (XGB), demostraron ser eficaces para identificar patrones complejos en los datos, pero su rendimiento se vio afectado por conjuntos de datos desequilibrados, lo que generó falsos positivos y negativos. Para mitigar este problema, se aplicó la técnica SMOTE para balancear los datos, mejorando la detección de ransomware.

Además, se destacó la importancia de técnicas de interpretabilidad, como LIME, para comprender las decisiones del modelo y aumentar la confianza en su implementación. La adaptabilidad de los modelos de ML a diversas plataformas, incluidos dispositivos móviles y IoT, fue otro aspecto crucial, señalando la necesidad de desarrollar modelos más ligeros y eficientes, como Random Forest y XGBoost, que requieren menos recursos computacionales.

El estudio también subrayó el impacto organizacional de la detección temprana de ransomware, que mejora la seguridad de los datos y asegura la continuidad operativa. Sin embargo, el aumento de ataques adversariales que manipulan los modelos de ML para evadir la detección representa un desafío en constante evolución. Para abordar este reto, se propone un enfoque de defensa en profundidad, que combine modelos avanzados de ML con medidas tradicionales de seguridad cibernética.

6. CONCLUSIÓN

El estudio logró cumplir con los objetivos específicos planteados. En primer lugar, se realizó una exploración y revisión de estudios previos, identificando las tendencias actuales en la aplicación de aprendizaje automático para la detección de ransomware, destacando la eficacia de algoritmos supervisados como Random Forest (RF) y Extreme Gradient Boosting (XGB), así como la utilidad de modelos de aprendizaje profundo en escenarios específicos. En segundo lugar, se consolidaron los hallazgos al sintetizar las ventajas y desventajas de las técnicas analizadas, señalando que, aunque los modelos de ML ofrecen una detección más precisa y rápida, su desempeño puede verse limitado por problemas de datos desequilibrados y la falta de representatividad de las amenazas modernas en los conjuntos de datos. Además, se compararon estrategias en plataformas interconectadas, evaluando la aplicabilidad de los modelos de ML en diversos entornos tecnológicos y concluyendo que las soluciones deben adaptarse a las capacidades específicas de cada plataforma, como la potencia de procesamiento en dispositivos IoT o la diversidad de amenazas en dispositivos móviles. En general, los resultados confirman que los modelos de aprendizaje automático representan una herramienta indispensable para enfrentar las amenazas avanzadas del ransomware. Sin embargo, se destacó la necesidad de un esfuerzo continuo para superar desafíos como la escalabilidad, el desequilibrio de datos y la incorporación de medidas defensivas frente a técnicas de inteligencia artificial ofensiva. Asimismo, se señaló que la integración de enfoques estáticos, dinámicos e híbridos, junto con la adopción de marcos regulatorios como el NIST CSF, puede fortalecer significativamente las estrategias de defensa en el panorama actual de ciberseguridad.

REFERENCIAS

Aboah Boateng, E., Bruce, J. W., & Talbert, D. A. (2022). Anomaly Detection for a Water Treatment System Based on One-Class Neural Network. IEEE Access, 10, 115179–115191. https://doi.org/10.1109/ACCESS.2022.3218624

Albeiro, W., Bedoya, Ú., Amariles, M., Codirector, C., Lorena, S., & Córdoba, V. (n.d.). Técnica de Machine Learning para la Prevención del Malware-Ransomware.

Alzahrani, A. I. A., Ayadi, M., Asiri, M. M., Al-Rasheed, A., & Ksibi, A. (2022). Detecting the Presence of Malware and Identifying the Type of Cyber Attack Using Deep Learning and VGG-16 Techniques. Electronics (Switzerland), 11(22). https://doi.org/10.3390/electronics11223665

Arnold, D., Gromov, M., & Saniie, J. (2024). Network Traffic Visualization Coupled With Convolutional Neural Networks for Enhanced IoT Botnet Detection. IEEE Access, 12, 73547–73560. https://doi.org/10.1109/ACCESS.2024.3404270

Asharf, J., Moustafa, N., Khurshid, H., Debie, E., Haider, W., & Wahab, A. (2020). A review of intrusion detection systems using machine and deep learning in internet of things: Challenges, solutions and future directions. In Electronics (Switzerland) (Vol. 9, Issue 7). MDPI AG. https://doi.org/10.3390/electronics9071177

Ashley, T. D., Kwon, R., Gourisetti, S. N. G., Katsis, C., Bonebrake, C. A., & Boyd, P. A. (2022). Gamification of Cybersecurity for Workforce Development in Critical Infrastructure. IEEE Access, 10, 112487–112501. https://doi.org/10.1109/ACCESS.2022.3216711

Aslam, M., Ye, D., Hanif, M., & Asad, M. (2020). Adaptive Machine learning: A Framework for Active Malware Detection. 2020 16th International Conference on Mobility, Sensing and Networking (MSN), 57–64. https://doi.org/10.1109/MSN50589.2020.00025

Bao, Z., Bie, B., Fan, W., Li, D., Li, M., Lin, K., Lin, W., Liu, P., Liu, P., Lv, Z., Ouyang, M., Sun, C., Tang, S., Wang, Y., Wei, Q., Wu, X., Xie, M., Zhang, J., Zhao, R., … Zhu, Y. (2024). Rock: Cleaning Data with both ML and Logic Rules. Proceedings of the VLDB Endowment, 17(12), 4373–4376. https://doi.org/10.14778/3685800.3685878

Belkhiri, A., & Dagenais, M. (2024). Analyzing GPU Performance in Virtualized Environments: A Case Study. Future Internet, 16(3), 72. https://doi.org/10.3390/fi16030072

Cen, M., Jiang, F., Qin, X., Jiang, Q., & Doss, R. (2024). Ransomware early detection: A survey. In Computer Networks (Vol. 239). Elsevier B.V. https://doi.org/10.1016/j.comnet.2023.110138

Das, V., Nair, B. B., & Thiruvengadathan, R. (2024). A Novel Feature Encoding Scheme for Machine Learning Based Malware Detection Systems. IEEE Access, 12, 91187–91216. https://doi.org/10.1109/ACCESS.2024.3420080

De, C. (n.d.). UNIVERSIDAD POLITÉCNICAPOLIT´POLITÉCNICA SALESIANA SEDE QUITO.

Dugyala, R., Reddy, N. H., Maheswari, V. U., Mohammad, G. B., Alenezi, F., & Polat, K. (2022). Analysis of Malware Detection and Signature Generation Using a Novel Hybrid Approach. In Mathematical Problems in Engineering (Vol. 2022). Hindawi Limited. https://doi.org/10.1155/2022/5852412

Falowo, O. I., Botsyoe, L., Koshoedo, K., & Ozer, M. (2024). Enhancing Cybersecurity with Artificial Immune Systems and General Intelligence: A New Frontier in Threat Detection and Response. IEEE Access. https://doi.org/10.1109/ACCESS.2024.3454543

Falowo, O. I., Edinam Botsyoe, L., Koshoedo, K., & Ozer, M. (2024). Enhancing Cybersecurity With Artificial Immune Systems and General Intelligence: A New Frontier in Threat Detection and Response. IEEE Access, 12, 123811–123822. https://doi.org/10.1109/ACCESS.2024.3454543

HACIA EL FUTURO SOSTENIBLE ALINEANDO NUESTRO PROPÓSITO. (n.d.).

Hilabi, R., & Abu-Khadrah, A. (2024). Windows operating system malware detection using machine learning. Bulletin of Electrical Engineering and Informatics, 13(5), 3401–3410. https://doi.org/10.11591/eei.v13i5.8018

Hossain, M. A., & Islam, M. S. (2024). Enhanced detection of obfuscated malware in memory dumps: a machine learning approach for advanced cybersecurity. Cybersecurity, 7(1). https://doi.org/10.1186/s42400-024-00205-z

Kalphana, K. R., Aanjankumar, S., Surya, M., Ramadevi, M. S., Ramela, K. R., Anitha, T., Nagaprasad, N., & Krishnaraj, R. (2024). Prediction of android ransomware with deep learning model using hybrid cryptography. Scientific Reports, 14(1), 22351. https://doi.org/10.1038/s41598-024-70544-x

Kumar, A., Alshahrani, H. M., Alotaibi, F., & Nanthaamornphong, A. (2024). A hybrid detection algorithm for 5G OTFS waveform for 64 and 256 QAM with Rayleigh and Rician channels. Open Engineering, 14(1). https://doi.org/10.1515/eng-2024-0008

Lin, C.-J., & Jeng, S.-Y. (2020). Optimization of Deep Learning Network Parameters Using Uniform Experimental Design for Breast Cancer Histopathological Image Classification. Diagnostics, 10(9), 662. https://doi.org/10.3390/diagnostics10090662

Marais, B., Quertier, T., & Morucci, S. (2022). AI-based Malware and Ransomware Detection Models. http://arxiv.org/abs/2207.02108

McIntosh, T., Kayes, A. S. M., Chen, Y. P. P., Ng, A., & Watters, P. (2022). Ransomware Mitigation in the Modern Era: A Comprehensive Review, Research Challenges, and Future Directions. In ACM Computing Surveys (Vol. 54, Issue 9). Association for Computing Machinery. https://doi.org/10.1145/3479393

Mouro González Dirección, S., María, E., & Pereira, H. (n.d.). Técnicas de aprendizaje máquina para análisis de malware Estudante.

Oluomachi, E., Ahmed, A., Ahmed, W., & Samson, E. (2024). ASSESSING THE EFFECTIVENESS OF CURRENT CYBERSECURITY REGULATIONS AND POLICIES IN THE US. International Journal of Scientific and Research Publications, 14(2), 78. https://doi.org/10.29322/IJSRP.14.02.2024.p14610

Raza, A., E. Mahmoud, E., M. Al-Bugami, A., Baleanu, D., Rafiq, M., Mohsin, M., & Al Nuwairan, M. (2022). Examination of Pine Wilt Epidemic Model through Efficient Algorithm. Computers, Materials & Continua, 71(3), 5293–5310. https://doi.org/10.32604/cmc.2022.024535

Singh, A., Ikuesan, R. A., & Venter, H. (n.d.). Ransomware Detection using Process Memory.

Wang, Z., Song, Y., Xu, E., Wu, H., Tong, G., Sun, S., Li, H., Liu, J., Ding, L., Liu, R., Zhu, J., & Wu, J. (n.d.). Ransom Access Memories: Achieving Practical Ransomware Protection in Cloud with DeftPunk. https://www.usenix.org/conference/osdi24/presentation/wang-zhongyu

Zakaria, W. Z. A., Abdollah, M. F., Abdollah, O., & Warusia Mohamed, S. M. S. M. M. (2025). Ransomware Early Detection using Machine Learning Approach and Pre-Encryption Boundary Identification. Journal of Advanced Research in Applied Sciences and Engineering Technology, 47(2), 121–137. https://doi.org/10.37934/ARASET.47.2.121137

Zhang, H., Zhao, L., Yu, A., Cai, L., & Meng, D. (2024). Ranker: Early Ransomware Detection Through Kernel-Level Behavioral Analysis. IEEE Transactions on Information Forensics and Security, 19, 6113–6127. https://doi.org/10.1109/TIFS.2024.3410511

ANEXO

Código del Modelo de Machine Learning

```
import numpy as np
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score
from sklearn.preprocessing import StandardScaler

# 1. Cargar los datos
file1_path = 'DominioPublico_1900_1944-UTF8.csv'
file2_path = 'bdh_2024-UTF8 (1).csv'

data1 = pd.read_csv(file1_path, encoding='latin1', delimiter=";", on_bad_lines='skip')
data2 = pd.read_csv(file2_path, encoding='latin1', delimiter=";", on_bad_lines='skip')

# 2. Preprocesar Archivo 1
# Seleccionar columnas relevantes y convertir a numérico donde sea posible
if 'fecha de nacimiento' in data1.columns:
    data1['fecha de nacimiento'] = pd.to_numeric(data1['fecha de nacimiento'], errors='coerce')
    data1['fecha de fallecimiento'] = pd.to_numeric(data1['fecha de fallecimiento'], errors='coerce')
```

```
if 'ocupación' in data1.columns:
    data1['ocupación_len'] = data1['ocupación'].astype(str).apply(len)  # Longitud del texto como característica

# Rellenar valores faltantes
numeric_features_1 = ['fecha de nacimiento', 'fecha de fallecimiento', 'ocupación_len']
data1 = data1[numeric_features_1].fillna(0)

# 3. Preprocesar Archivo 2
# Convertir los valores mensuales a características agregadas
data2_stats = data2.iloc[:, 1:].apply(pd.to_numeric, errors='coerce').fillna(0)
data2['total'] = data2_stats.sum(axis=1)
data2['mean'] = data2_stats.mean(axis=1)
data2 = data2[['total', 'mean']]

# 4. Combinar los datos
# Crear un DataFrame combinado para usar como conjunto de entrenamiento
combined_data = pd.concat([
data1.reset_index(drop=True),
data2.reset_index(drop=True)
], axis=1).fillna(0)
```

```
# Generar etiquetas simuladas (1 = ransomware, 0 = normal)
np.random.seed(42)
combined_data['label'] = np.random.choice([0, 1], size=combined_data.shape[0], p=[0.8, 0.2])

# 5. Preparar los datos para el modelo
X = combined_data.drop(columns=['label'])
y = combined_data['label']

# Escalar las características
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

# Dividir los datos en conjuntos de entrenamiento y prueba
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.3, random_state=42)

# 6. Crear y entrenar el modelo
model = RandomForestClassifier(random_state=42, n_estimators=100)
model.fit(X_train, y_train)

# 7. Evaluar el modelo
y_pred = model.predict(X_test)
```

```
# Métricas de evaluación
accuracy = accuracy_score(y_test, y_pred)
conf_matrix = confusion_matrix(y_test, y_pred)
report = classification_report(y_test, y_pred)

# Mostrar resultados
print(f"Exactitud del modelo: {accuracy * 100:.2f}%")
print("\nMatriz de confusión:")
print(conf_matrix)
print("\nReporte de clasificación:")
print(report)

import matplotlib.pyplot as plt
import seaborn as sns

# 1. Visualizar la matriz de confusión
def plot_confusion_matrix(conf_matrix, class_names):
    plt.figure(figsize=(8, 6))
     sns.heatmap(conf_matrix, annot=True, fmt='d', cmap='Blues', xticklabels=class_names, yticklabels=class_names)
    plt.xlabel('Predicted Labels')
```

```
    plt.ylabel('True Labels')
    plt.title('Confusion Matrix')
    plt.show()
# Llamar a la función de la matriz de confusión
class_names = ['Normal', 'Ransomware']
plot_confusion_matrix(conf_matrix, class_names)

# 2. Importancia de las características
feature_importances = model.feature_importances_
feature_names = X.columns

# Crear un DataFrame para organizar
importances_df = pd.DataFrame({'Feature': feature_names, 'Importance': feature_importances})
importances_df = importances_df.sort_values(by='Importance', ascending=False)

# Graficar las importancias
plt.figure(figsize=(10, 6))
sns.barplot(x='Importance', y='Feature', data=importances_df, palette='viridis')
plt.title('Feature Importances')
plt.xlabel('Importance')
plt.ylabel('Features')
```

```
plt.show()

# 3. Distribución de etiquetas

def plot_label_distribution(labels):

    plt.figure(figsize=(6, 4))

    sns.countplot(x=labels, palette='pastel')

    plt.title('Label Distribution')

    plt.xlabel('Label')

    plt.ylabel('Count')

    plt.xticks(ticks=[0, 1], labels=['Normal', 'Ransomware'])

    plt.show()

plot_label_distribution(y)

# 4. Exactitud por clase

report_dict = classification_report(y_test, y_pred, output_dict=True)

class_accuracy = {label: metrics['precision'] for label, metrics in report_dict.items() if label in ['0', '1']}

# Graficar

plt.figure(figsize=(6, 4))

sns.barplot(x=list(class_accuracy.keys()), y=list(class_accuracy.values()), palette='magma')

plt.title('Precision by Class')

plt.xlabel('Class')

plt.ylabel('Precision')

plt.show()
```