UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA
FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

TÍTULO DEL TRABAJO DE TITULACIÓN

TÉCNICAS DE MINERÍA DE DATOS PARA LA DETECCIÓN DE AMENAZAS PERSISTENTES AVANZADOS EN PÁGINAS WEB

AUTOR

BALÓN GONZÁLEZ NIDIA ANAHI

MODALIDAD DE TITULACIÓN

PROYECTO DE UNIDAD DE INTEGRACIÓN CURRICULAR

Previo a la obtención del grado académico en
INGENIERO EN TECNOLOGÍAS DE LA INFORMACIÓN

TUTOR

MSIA. DANIEL QUIRUMBAY

Santa Elena, Ecuador

Año 2025

UNIVERSIDAD ESTATAL PENÍNSULA

DE SANTA ELENA

FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

TRIBUNAL DE SUSTENTACIÓN

Ing. José Sánchez Aquino, Mgt.

DIRECTOR DE LA CARRERA

Lsi. Daniel Quirumbay Yagual, Msia.

TUTOR

Ing. Iván Coronel Suárez, Mgt.

DOCENTE ESPECIALISTA

Ing. Marjorie Coronel Suárez, Mgtí.

DOCENTE GUÍA UIC

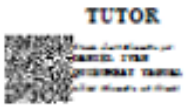
UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA
FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

CERTIFICACIÓN

Certifico que luego de haber dirigido científica y técnicamente el desarrollo y estructura final del trabajo, este cumple y se ajusta a los estándares académicos, razón por el cual apruebo en todas sus partes el presente trabajo de titulación que fue realizado en su totalidad por Srta Nidia Anahi Balón González, como requerimiento para la obtención del título de Ingeniero en Tecnologías de la Información.

La Libertad, a los 10 días del mes de diciembre del año 2025

TUTOR



Msia. Daniel Quirumbay

UNIVERSIDAD ESTATAL PENÍNSULA

DE SANTA ELENA

FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

DECLARACIÓN DE RESPONSABILIDAD

Yo, **Balón González Nidia Anahi**

DECLARO QUE:

El trabajo de Titulación**, Técnicas de Minería de Datos para la Detección de Amenazas Persistentes Avanzadas en Páginas Web**, previo a la obtención del título en Ingeniero en Tecnologías de la Información, ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

La Libertad, a los 10 días del mes de diciembre del año 2025

EL AUTOR

Balón González Nidia Anahi

UNIVERSIDAD ESTATAL PENÍNSULA

DE SANTA ELENA

FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

CERTIFICACIÓN DE ANTIPLAGIO

Certifico que después de revisar el documento final del trabajo de titulación denominado **Técnicas de Minería de Datos para la Detección de Amenazas Persistentes Avanzadas en Páginas Web** presentado por la estudiante **Balón González Nidia Anahi** fue enviado al Sistema Antiplagio, presentando un porcentaje de similitud correspondiente al 8%, por lo que se aprueba el trabajo para que continúe con el proceso de titulación.


CERTIFICADO DE ANÁLISIS
magister

TI_Nidia_Anahi_Balon_Gonzalez

8%
Textos sospechosos

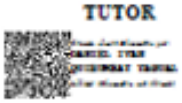
4% Similitudes
< 1 % similitudes entre comillas
0 % entre las fuentes mencionadas
5% Idiomas no reconocidos
20% Textos potencialmente generados por la IA (ignorado)

Nombre del documento: TI_Nidia_Anahi_Balon_Gonzalez.docx
ID del documento: 7f963000b5c24dd5585b9a7f23e3d4b6bf68703b
Tamaño del documento original: 4,41 MB

Depositante: DANIEL IVAN QUIRUMBAY YAGUAL
Fecha de depósito: 9/12/2025
Tipo de carga: interface
fecha de fin de análisis: 9/12/2025

Número de palabras: 23.706
Número de caracteres: 165.738

TUTOR



Msia. Daniel Quirumbay

UNIVERSIDAD ESTATAL PENÍNSULA

DE SANTA ELENA

FACULTAD DE SISTEMAS Y TELECOMUNICACIONES

AUTORIZACIÓN

Yo, **Balón González Nidia Anahi**

Autorizo a la Universidad Estatal Península de Santa Elena, para que haga de este trabajo de titulación o parte de él, un documento disponible para su lectura consulta y procesos de investigación, según las normas de la Institución.

Cedo los derechos en línea patrimoniales del trabajo de titulación con fines de difusión pública, dentro de las regulaciones de la Universidad, siempre y cuando esta reproducción no suponga una ganancia económica y se realice respetando mis derechos de autor

Santa Elena, a los 10 días del mes de diciembre del año 2025

AUTOR

Balón González Nidia Anahi

AGRADECIMIENTO

Doy gracias a Jehová, quien me dio fuerzas, iluminación y la oportunidad de llegar hasta este momento sin su guía y protección, este logro no habría sido posible. A mi mamá, **Alba González** y hermano **Bryan Balón** por su apoyo incondicional y por ser mi mayor impulso para seguir adelante. A mis abuelitos paternos, **Maximina Guale y Justo Balón**, por su cariño, por sus palabras sabias y por estar siempre presentes.

A mi tutor, el **Msia. Daniel Quirumbay**, por su paciencia, su guía y el tiempo que dedicó a acompañarme en este proceso. Su apoyo fue clave para que pudiera culminar esta etapa.

A mis amigos y, sobre todo, a mi mejor amigo, Anghelo, por hacer este camino más llevadero con su compañía, ánimo y amistad sincera. Y a mis **amigos Nicole y Aaron**, que desde el primer día me brindaron apoyo sin dudarlo.

Y de manera muy especial a mi novio **Christian Alejandro** quien también formó parte importante de este trabajo y de este proceso. Gracias por tu apoyo, tu paciencia, tu amor y por estar a mi lado en los momentos de estrés, cansancio y esfuerzo. Tu presencia fue una fuerza que me impulsó a seguir adelante.

Nidia Anahi Balón Gonzalez.

DEDICATORIA

A mi mamá, por su amor infinito, por estar a mi lado en cada desafío y por enseñarme que con esfuerzo y constancia todo sueño puede alcanzarse. Gracias por ser mi mayor apoyo y mi guía en cada etapa.

A mis abuelitos paternos, por su ternura, sus consejos llenos de experiencia y por brindarme siempre un hogar de confianza y fortaleza. Su presencia ha sido un sostén invaluable en mi vida.

A mi novio, por su cariño, su paciencia y por acompañarme en cada momento de esta trayectoria. Gracias por animarme, por tu comprensión y por creer firmemente en mi capacidad de llegar hasta aquí.

Les dedico este trabajo a ustedes, que siempre confiaron en mí, que me alentaron con sus palabras y que, con su amor y ejemplo, hicieron posible este logro académico.

Nidia Anahi Balón Gonzalez.

ÍNDICE GENERAL

ÍNDICE DE TABLA

ÍNDICE DE FIGURAS

RESUMEN

Este estudio utilizó técnicas avanzadas de minería de datos y aprendizaje automático para identificar Amenazas Persistentes Avanzadas (APTs) en el tráfico web de la UPSE. Su propósito fue crear un sistema de detección eficiente, desarrollar modelos predictivos, aplicar algoritmos de clasificación y visualizar los resultados. La metodología siguió la metodología CRISP-DM, desde la comprensión del problema hasta la explotación de los hallazgos, incluyendo un análisis detallado de un dataset con 394,771 registros de tráfico de red. Se evaluaron cuatro algoritmos: SVM, Random Forest, K-means e Isolation Forest, los resultados mostraron que los modelos supervisados son más factibles, mientras los no supervisados mostraron un rendimiento inferior. Se concluye que la integración de minería de datos con modelos de machine learning supervisados constituye una estrategia práctica y eficaz para identificar APTs en entornos de red institucionales. El sistema desarrollado, junto con su dashboard, ofrece un enfoque claro que refuerza la seguridad ayudando de manera proactiva a proteger las infraestructuras e información contra APTs.

Palabras claves: Mineria de Datos, APTs, Machine Learning

ABSTRACT

This study used advanced data mining and machine learning techniques to identify Advanced Persistent Threats (APTs) in UPSE's web traffic. Its purpose was to create an efficient detection system, develop predictive models, apply classification algorithms, and visualize the results. The methodology followed the CRISP-DM framework, from understanding the problem to exploiting the findings, including a detailed analysis of a dataset with 394,771 network traffic records. Four algorithms were evaluated: SVM, Random Forest, K-means, and Isolation Forest. The results showed that supervised models were more feasible, while unsupervised models performed worse. The study concludes that integrating data mining with supervised machine learning models constitutes a practical and effective strategy for identifying APTs in institutional network environments. The developed system, along with its dashboard, offers a clear approach that strengthens security by proactively helping to protect infrastructure and information against APTs.

Keywords: Data Mining, APTs, Machine Learning

INTRODUCCIÓN

Las Amenazas Persistentes Avanzadas (APTs) representan uno de los mayores desafíos en la ciberseguridad actual. A diferencia de las amenazas convencionales, las APTs son ataques específicos, discretos y de larga duración, diseñados para infiltrarse en redes, mantener un acceso prolongado y exfiltrar datos sensibles sin ser detectados. Su sofisticación técnica y su capacidad para evadir los mecanismos de seguridad tradicionales las convierten en un riesgo crucial para organizaciones que gestionan información valiosa, como instituciones educativas y gubernamentales, así como empresas. En un entorno digital cada vez más conectado, detectar estas amenazas de forma anticipada se ha vuelto una prioridad para asegurar la integridad y confidencialidad de la información.

La relevancia de esta investigación radica en la exposición de entornos académicos, como la Universidad Estatal Península de Santa Elena (UPSE), a este tipo de amenazas. La universidad gestiona una red compleja con múltiples dispositivos, servicios en línea y tráfico web constante, lo que la convierte en un área ideal para realizar diversos ataques. Actualmente, la falta de un sistema de monitoreo avanzado que analice patrones de comportamiento anómalo limita la capacidad de respuesta ante incidentes de seguridad. Implementar un enfoque basado en minería de datos y aprendizaje automático no solo permitiría detectar APTs de manera proactiva sino también optimizar los recursos de seguridad y reducir las falsas alarmas, fortaleciendo así la capacidad de detectar anomalías en la red de la institución.

El propósito principal de este trabajo es diseñar técnicas avanzadas de minería de datos para identificar de manera eficiente las Amenazas Persistentes Avanzadas en el tráfico web de la UPSE. Para lograrlo, se plantean tres objetivos específicos: desarrollar modelos predictivos que puedan detectar comportamientos anómalos y prever posibles ataques, emplear técnicas de análisis y minería de datos para reconocer patrones relacionados con las APTs, como beaconing, exfiltración y escaneo de puertos y crear un panel de control visual que permita monitorear los resultados y validar la efectividad de las técnicas implementadas.

La estructura de este trabajo de investigación consta de dos capítulos principales. El Capítulo I presenta las bases conceptuales y metodológicas del estudio, en este capítulo se analiza el contexto institucional, se revisa el estado del arte y se definen las variables y técnicas para la recolección de datos. Por otro lado, el Capítulo II desarrolla el marco teórico y conceptual, especifica los requisitos funcionales y técnicos, y describe en detalle la parte central de la investigación: la implementación de la metodología CRISP-DM.

Este estudio se basa en la metodología CRISP-DM, que asegura un proceso estructurado desde la comprensión del problema hasta la explotación de los resultados. Además, se alinea con las políticas nacionales de innovación tecnológica y seguridad digital, ayudando no solo a proteger la infraestructura universitaria, sino también a promover avances en conocimientos de ciberseguridad aplicada. Los resultados de esta investigación podrán adaptarse y replicarse en otras instituciones o incluso en empresas, contribuyendo a proteger de manera más efectiva los entornos digitales expuestos a riesgos.

CAPÍTULO I: FUNDAMENTACIÓN

1.1 ANTECEDENTES

El progreso tecnológico ha sido un motor clave en la transformación de nuestra sociedad contemporánea, permitiendo avances significativos en diversos ámbitos de la vida cotidiana. Desde el nacimiento de internet hasta la llegada de dispositivos móviles y la inteligencia artificial, hemos sido testigos de un crecimiento exponencial en las capacidades tecnológicas que han transformado profundamente nuestra manera de vivir, trabajar y relacionarnos [1].

En el entorno digital actual, la seguridad informática es fundamental para proteger la información y garantizar la confidencialidad de los datos. En este sentido, la aplicación de técnicas avanzadas, como la minería de datos, resulta esencial para identificar patrones y comportamientos anómalos que puedan indicar posibles amenazas [2].

La Universidad Estatal Península de Santa Elena obtuvo su aprobación en 1990, iniciando con 4 facultades actualmente ha experimentado un crecimiento y se han establecido 18 carreras distribuidas en 7 facultades, añadiendo que también tiene instalaciones de la sede en el cantón Playas, Provincia del Guayas [3].

Dentro de esta estructura, la Facultad de Sistemas y Telecomunicaciones (FACSISTEL), creada en 2010, alberga miles de estudiantes, además de los docentes, en la cual hace uso de sus instalaciones, que incluyen oficinas administrativas, salas para docentes y diversos laboratorios, entre ellos uno certificado por CISCO. En un recorrido de observación realizado por la facultad**,** se identificó la coexistencia de múltiples conexiones de red, tanto alámbricas como inalámbricas, algunas abiertas y otras privadas. Estas redes comprenden servicios proporcionados por la universidad y puntos de acceso generados por dispositivos externos. En la actualidad, no se dispone de un sistema de monitoreo que registre los sitios web visitados por los estudiantes, lo que incrementa el riesgo de que accedan a páginas web peligrosas que puedan comprometer la seguridad de la infraestructura tecnológica.

Desde hace aproximadamente ocho años, cuando comenzaron a documentarse estas amenazas de forma más específica se observó un crecimiento en el primer trimestre del 2017, un 28% del malware detectado en América Latina y el Caribe estaba dirigido a Android comparado al 20% registrado en el último trimestre de 2016, lo que demuestra una evolución más rápida en comparación con otras regiones [4].

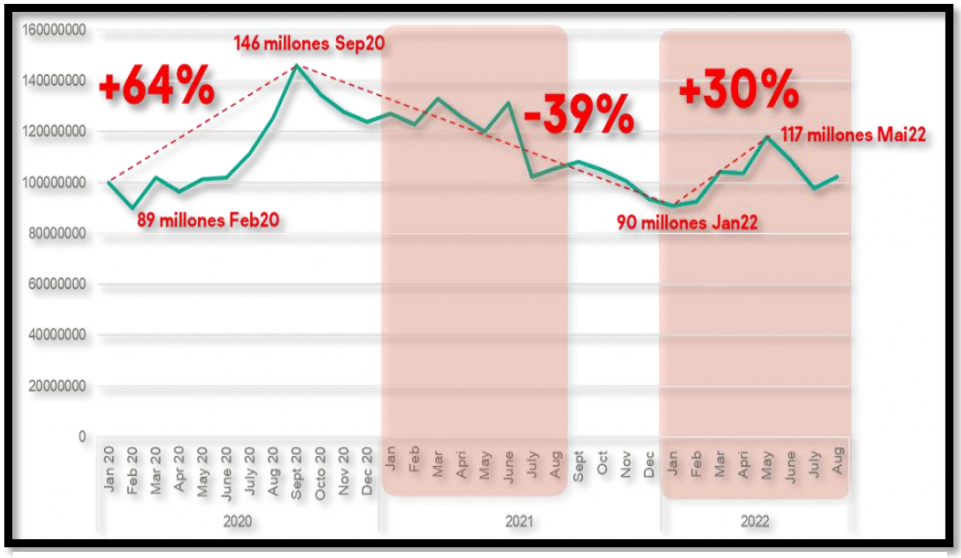


Ilustración 1:Panorama de Amenazas de Kaspersky 2021 – 2022 [93].

El trabajo titulado *"Una revisión sistemática de la literatura sobre comportamientos de amenazas persistentes avanzadas y su estrategia de detección",* analiza los desafíos que plantean las amenazas persistentes avanzadas (APT) debido a su naturaleza sofisticada y persistente, que amenaza la confidencialidad, la integridad y la disponibilidad de los sistemas organizacionales [5]. Para mejorar la precisión de la detección, los autores proponen integrar el análisis del comportamiento de ataques en múltiples etapas con técnicas de evaluación y visualización de vulnerabilidades. Este enfoque permite la identificación temprana de objetivos potenciales y respalda estrategias de defensa proactivas para fortalecer la ciberseguridad organizacional [5].

El trabajo de proyecto de maestría *"Uso de técnicas de web mining: aplicación empírica en el sector de la administración pública"* habla sobre la minería de datos como una herramienta poderosa para optimizar la experiencia del usuario y la estructura de sitios institucionales y el aspecto clave que es el preprocesamiento

de *logs*, donde se elimina ruido (spiders, errores) y se agrupan sesiones, permitiendo análisis más precisos. Sin embargo, añade que desafíos como la falta de datos demográficos y el sesgo temporal en los registros señalan áreas para futuras investigaciones en este campo [6].

En el trabajo de Say Valdez titulado *"Minería web como herramienta de análisis de ficheros log en servidores web"*, se plantea la importancia de aplicar técnicas de minería de datos en entornos web para descubrir patrones de comportamiento y extraer conocimiento útil a partir de grandes volúmenes de información almacenada en los registros de los servidores. Este estudio evidencia cómo la minería web constituye una extensión de la minería de datos que aprovecha la información contenida en los ficheros históricos para comprender las preferencias de los usuarios, optimizar servicios y apoyar la toma de decisiones organizacionales [7].

En el trabajo "El modelo teórico que combina el ciclo de vida de un ataque APT con técnicas de machine learning" [8], se tiene como objetivo facilitar la detección temprana de estas amenazas, reduciendo el tiempo de exposición y mitigando los posibles impactos en las organizaciones afectadas. A través de un análisis exhaustivo de las tácticas, técnicas y procedimientos (TTP) utilizados por los atacantes, así como de la evaluación de algoritmos de ML, busca contribuir al avance de la ciberseguridad frente a una de las amenazas más persistentes y dañinas de la actualidad [9].

El estudio titulado " Mejora del modelo de Defensa en Profundidad de la ANH con el fin de aumentar la protección ante las Amenazas Persistentes Avanzadas" [10] analiza la eficacia del modelo de defensa en capas en una entidad gubernamental colombiana frente a amenazas avanzadas persistentes (APT) [10]. Este trabajo destaca la necesidad de complementar la defensa en profundidad con mecanismos de detección proactiva, correlación de eventos y gestión continua de incidentes, ofreciendo un marco relevante para organizaciones que enfrentan ciberamenazas sofisticadas.

La prevención de ataques informáticos como el ransomware a entidades públicas y privadas del Ecuador. El ransomware es conocido como un programa malicioso y

dañino que perjudica los dispositivos como laptops, computadores, entre otros con la finalidad de encriptar los archivos. Las instituciones involucradas por parte de entidades públicas se vieron afectadas: la embajada, el Banco Central, el SRI y la Presidencia. Todo este conflicto se vio relacionado con los problemas diplomáticos relacionados con Julian Assange. Este margen informático puso en evidencia la insuficiente preparación y el casi escaso sobre el tema de ciberseguridad en el país [11] .

Las fuerzas armadas de contar con un equipamiento de elite en fuerza y sobre todo de mantener una infraestructura informática implacable sobre sus unidades navales y militares, mantienen servicios automatizados y así mismo de sistemas corporativos. Los errores presentes se ven relacionado por el bajo nivel de control de seguridad de la información lo que permite que aumente los ataques informáticos desde el exterior. Las Fuerzas Armadas del Ecuador (FF. AA.) es la encargada de gestionar los datos militares confidenciales y secretos, es claro que la falta de capacitación, el cumplimiento de políticas de seguridad ha provocado la reducción de preservar la confiabilidad, disponibilidad y la integridad de la información [12].

Todo el enfoque investigativo que abarca al tema central de amenazas persistentes avanzadas, que evaden los sistemas tradicionales, lo que requiere enfoques innovadores basados en minería de datos e inteligencia artificial para la detección temprana sobre estas situaciones críticas que provocan un estado emergente a organizaciones que cuentan con activos de información sensibles. La parte involucrada se sumerge en las páginas web, aquella con cuenta con una infraestructura de cliente-servidor en donde receptas las peticiones de usuarios, y no abordar las soluciones de brechas de seguridad, el activo de información se encontrará en peligro.

Por ende, el trabajo aporta un fortalecimiento sobre la seguridad organizacional, también en desarrollo seguro y la innovación de técnicas que permitan la identificación de patrones de ataque, como también una evaluación de vulnerabilidades, permitiendo así una protección sobre información crucial y la implementación de medidas preventivas y proactivas.

1.2 DESCRIPCIÓN DEL PROYECTO

El proyecto de investigación cuenta con el desarrollo de técnicas avanzadas de minería de datos y análisis de comportamiento para detectar eficazmente las Amenazas Persistentes Avanzadas (APTs) en el tráfico web. El objetivo principal es determinar actividades maliciosas y garantizar la seguridad cibernética en entornos digitales mediante el uso de modelos predictivos y análisis de patrones anómalos, garantizando la calidad y coherencia de los datos analizados.

El proyecto está basado en la de metodología que son la CRISP-DM (Proceso Estándar para Minería de Datos entre Industrias).

Fase 1: Comprensión del problema

- Definir problema y objetivos del modelo

Fase 2: Análisis de los datos

- Se recoge información que son los dataset preprocesados
- Se verifica la calidad de los datos
- Generar gráficos exploratorios

Fase 3: Preparación de los datos

- Se investiga profundamente y se escoge los algoritmos adecuados
- Limpieza de datos optimizando su calidad y transformación de datos
- Identificar patrones preliminares y filtrar ruido

Fase 4: Modelado

- Se selecciona un modelo adecuado y específico.
- Se mide las KPI de cada algoritmo para poder determinar el mejor modelo

Fase 5: Evaluación

- Se emplean el análisis e interpretación de los resultados

Fase 6: Explotación

- Dashboard de visualización

Esta investigación proporciona una base metodología sólida para para la implementación de técnicas avanzadas de minería de datos y aprendizaje automático, creando un marco flexible que puede adaptarse a diferentes escenarios de análisis de tráfico de red. Este trabajo no solo busca mejorar la capacidad de detección de ataques cibernéticos complejos, sino que también contribuye al desarrollo de soluciones innovadoras en el campo de la seguridad informática. Del mismo modo, los resultados crearán materiales valiosos para la comunidad académica y profesional, contribuyendo al desarrollo de sistemas intelectuales y más estables frente a las amenazas en el entorno digital.

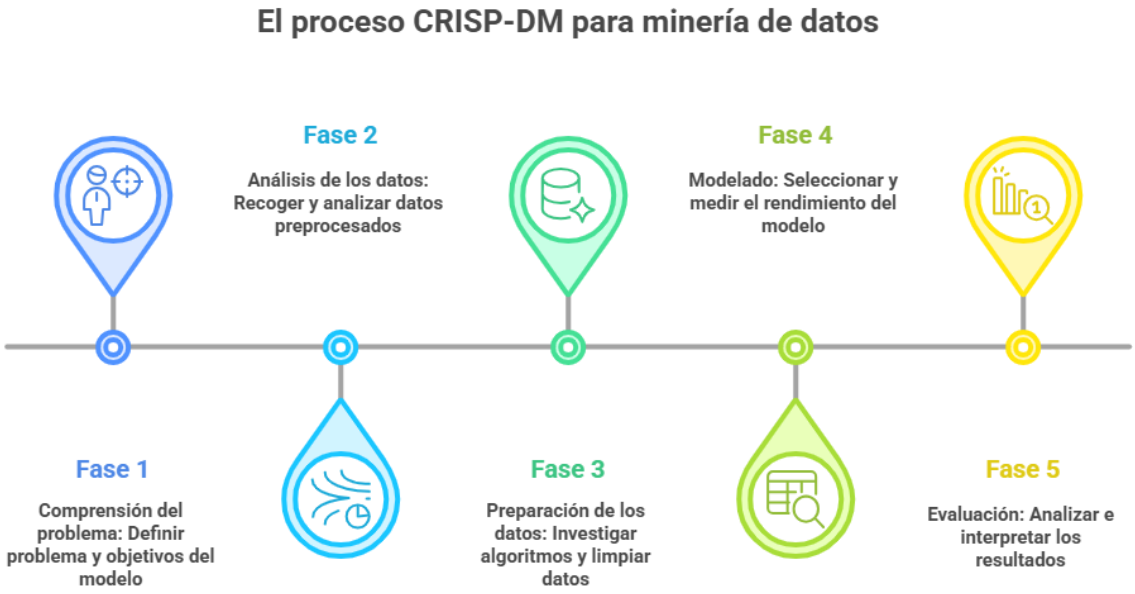


Ilustración 2:Proceso de metodología

1.3 OBJETIVOS DEL PROYECTO

1.3.1 OBJETIVO GENERAL

Desarrollar técnicas avanzadas de minería de datos para la detección eficaz de amenazas persistentes avanzadas (APTs) en páginas web, con el fin de identificar y mitigar actividades maliciosas, mejorando así la seguridad cibernética.

1.3.2 OBJETIVOS ESPECÍFICOS

- Aplicar modelos predictivos que permitan identificar comportamientos anómalos y anticipar posibles ataques, garantizando la calidad y coherencia de los datos utilizados para el análisis de tráfico web.

- Aplicar técnicas de análisis y minería de datos para detectar patrones anómalos en el tráfico web.
- Visualizar en un dashboard los resultados obtenidos con métodos de detección de amenazas para demostrar las ventajas y mejoras aportadas por las técnicas de análisis de datos y minería de datos desarrolladas.

1.4 JUSTIFICACIÓN DEL PROYECTO

Actualmente el creciente riesgo de amenazas cibernéticas, especialmente las Amenazas Persistentes Avanzadas (APT), que se han convertido en un desafío crítico para la seguridad informática en entornos académicos y corporativos. Los sistemas tradicionales de detección de intrusiones basados en firmas y reglas estáticas no son suficientes para abordar las amenazas emergentes que evolucionan rápidamente. Pang-Ning y otros colaboradores mencionan el potencial de las técnicas de minería de datos para identificar patrones anómalos en grandes volúmenes de datos [13]. Sin embargo, la detección tradicional no ha sido capaz de adaptarse a la necesidad para detectar las amenazas persistentes, lo que hace aún más necesario el uso de algoritmos de aprendizaje automático en la seguridad cibernética.

El alto uso de la tecnología y la exposición a amenazas que enfrentan las instituciones académicas y empresariales, como la UPSE, que gestionan datos sensibles y servicios en línea, son el tema central de esta investigación. La aplicación de minería de datos en el análisis de tráfico web ha sido investigada en trabajos previos, demostrando que el uso de modelos predictivos basados en datos puede mejorar la precisión de la detección de amenazas cibernéticas [14].

Para la UPSE, el proyecto contribuirá a fortalecer la seguridad de la red y proteger la infraestructura académica frente a posibles ataques a la red de la institución. Al aplicar técnicas de minería de datos y aprendizaje automático al tráfico de red es posible disminuir la cantidad de falsos positivos y negativos incrementando la precisión en la detección de amenazas. Para la comunidad académica, este estudio proporciona un marco de referencia sobre el uso de la minería de datos para la

ciberseguridad, abriendo posibilidades para futuras investigaciones sobre cómo aplicar estos métodos a otros contextos.

En la Ilustración 3 muestra cómo los principales servicios en la red están concentrados en puertos estándar como HTTPS y HTTP, pero también revela la presencia de puertos no estándar que requieren un análisis más exhaustivo. Estos datos reflejan la necesidad de garantizar una detección precisa de amenazas. Esto resalta la importancia de las técnicas de minería de datos con las metodologías de seguridad informática para abordar las vulnerabilidades de la red de forma eficaz.

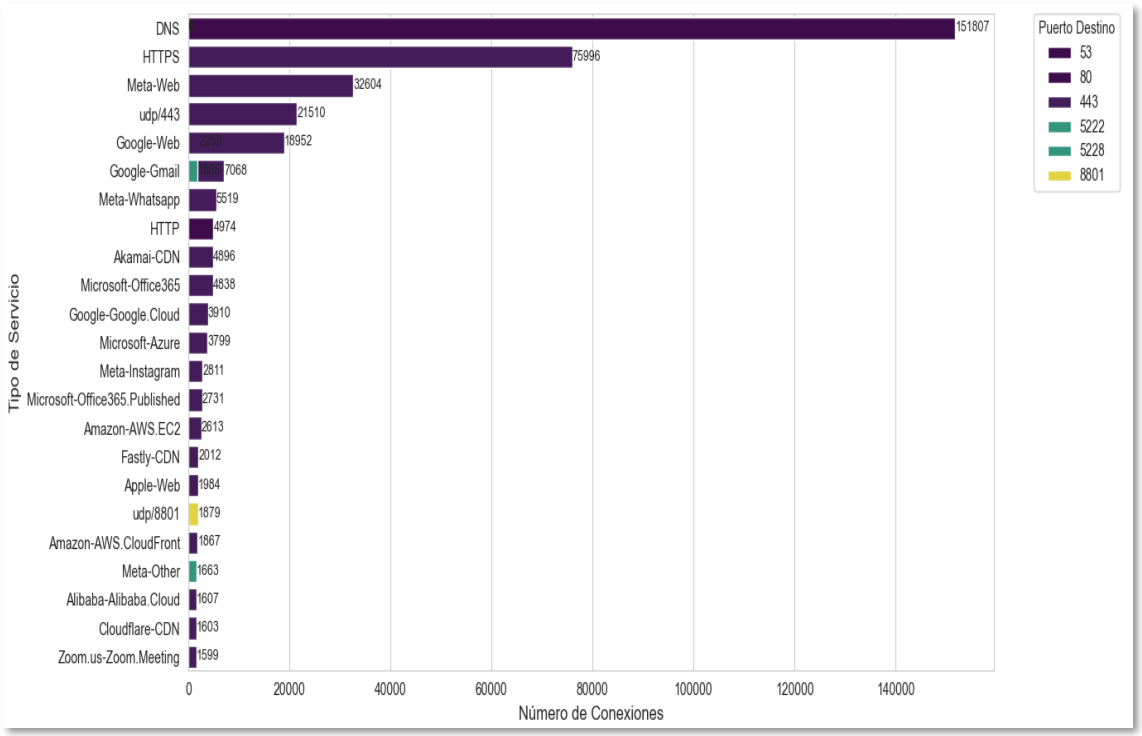


Ilustración 3:Servicios por Puerto Destino y Número de Conexiones: Elaboración Propia

El impacto esperado de esta investigación es tanto teórico como práctico, el proyecto contribuirá al conocimiento sobre cómo las técnicas de minería de datos y sobre los APTs ayudando así a mejorar la detección de amenazas esperando que los resultados permitan a la institución implementar un modelo de detección de amenazas mejorando la seguridad de los estudiantes y de la red. Además, este modelo podría servir como base para la adaptación de sistemas de detección en

otras universidades o instituciones impulsando el avance en el ámbito de la ciberseguridad.

El aprendizaje automático, las técnicas de minería de datos en los APTs ya se han aplicado en otros trabajos de investigación y así también en el campo empresarial para mejorar su seguridad web. Este proyecto se construye sobre esos trabajos previos y busca extender su aplicación a un entorno académico, donde las amenazas cibernéticas son un desafío creciente.

El presente proyecto de investigación se acopla al Plan de creación de oportunidades del Eje social.

Objetivo 7.- "Potenciar las capacidades de los ciudadanos y contribuir al desarrollo de una educación innovadora, equitativa y de excelencia para todas las personas" [15].

Política 7.1.- "Promover la modernización y eficiencia del modelo educativo por medio de la innovación y el uso de herramientas tecnológicas" [15].

1.5 ALCANCE DEL PROYECTO

El alcance del proyecto comprende el diseño de un modelo predictivo para el monitoreo, recopilación, gestión y análisis del tráfico de red de la facultad, orientado a la identificación de patrones y anomalías a través de técnicas de minería de datos aplicadas a los registros generados por los distintos estudiantes y administrativos; este modelo se plantea como una herramienta complementaria a las soluciones de seguridad ya existentes, con el fin de generar información útil para la detección temprana de incidentes, la optimización del rendimiento de la infraestructura tecnológica y la toma de decisiones estratégicas que aseguren la continuidad y la protección de la información académica y administrativa de la Universidad.

Fase 1: Compresión del Problema

En esta fase, se busca identificar y entender el contexto del problema, el reto específico que se enfrenta, y cómo la minería de datos o técnicas similares pueden ayudar a abordarlo. En el caso de la detección de APTs, en esta fase se define claramente el problema a resolver: identificar las Amenazas Persistentes Avanzadas

en la infraestructura de la red. Esto implica comprender cómo funcionan las APTs, qué comportamientos o patrones en los datos indican una posible amenaza, y qué objetivos se quieren alcanzar con la detección

En esta fase se establece el enfoque para la aplicación de técnicas de minería de datos, considerando los algoritmos que podrían ser útiles según los estudios previos y cómo estas técnicas podrán ayudar a detectar patrones de comportamiento anómalos en el tráfico de red.

Fase 2: Análisis de los Datos

En esta etapa se recolectará la información proveniente de los dispositivos y servicios de la red institucional, la cual se almacenará en datasets preprocesados para su estudio. Durante este proceso, se realizará un preprocesamiento de datos que incluye la ingeniería de características para extraer las variables más relevantes que ayuden a identificar patrones de comportamiento anómalos y amenazas persistentes avanzadas (APTs). Posteriormente, se verificará la calidad de los datos a fin de asegurar su completitud y se generarán gráficos exploratorios que permitirán una primera visualización de los volúmenes de tráfico y las posibles tendencias, con el propósito de comprender de manera preliminar el comportamiento de la red y detectar anomalías potenciales.

Fase 3: Preparación de los Datos

Durante esta fase se llevará a cabo una limpieza exhaustiva de los datos para eliminar registros incompletos, duplicados o inconsistentes, optimizando así su calidad. Se aplicarán técnicas de transformación de datos para estandarizar la información, lo que facilitará la aplicación de algoritmos adecuados para el análisis. Asimismo, se identificarán patrones preliminares y se filtrará el ruido presente en los registros, asegurando que los datos estén listos para ser utilizados en la fase de modelado de manera confiable y eficiente.

Fase 4: Modelado

En esta fase se seleccionará el modelo de minería de datos más adecuado para el análisis del tráfico de red, tomando en cuenta la esencia de los datos y las metas definidas en la etapa inicial. Se aplicarán diferentes algoritmos de clasificación y

detección de anomalías, evaluando el desempeño de cada uno mediante indicadores clave de rendimiento (KPI), como la precisión, la sensibilidad y la tasa de falsos positivos. El propósito será determinar cuál es el modelo más eficiente y efectivo para la detección temprana de patrones anómalos.

Fase 5: Evaluación

En esta etapa se emplearán diversos gráficos estadísticas y analíticas para interpretar los resultados obtenidos del modelo. Se compararán los resultados esperados con los alcanzados, verificando si los patrones detectados cumplen con los objetivos planteados. Además, se identificarán posibles limitaciones del modelo y se propondrán mejoras en su desempeño, con el fin de garantizar que la solución final sea robusta, confiable y aplicable al entorno real de la red de la facultad.

Fase 6: Explotación

En esta fase se pondrá en práctica el modelo desarrollado mediante la construcción de un dashboard de visualización que permita a los responsables de la red monitorear de manera continua y monitorear posteriores dataset para ver el comportamiento del tráfico. Este panel facilitará la interpretación de los resultados, brindará alertas tempranas sobre posibles anomalías y proporcionará un soporte visual para la toma de decisiones estratégicas. De esta forma, se garantizará que el conocimiento generado a partir del modelo tenga una aplicación práctica y un impacto directo en la seguridad y eficiencia de la red.

1.6 METODOLOGÍA DEL PROYECTO

1.6.1 METODOLOGÍA DE INVESTIGACIÓN

La metodología exploratoria [16] se centra en encontrar la mayor y relevante información posible para entender con más claridad lo incomprensible, se toma en cuenta esta metodología porque a base de estudios previos se entenderá y recopilará información sobre la minería de datos en APTs.

La investigación diagnóstica tiene como objetivo identificar y entender problemas específicos [17], observando el comportamiento de los distintos dispositivos y servicios conectados. Se identificarán los factores que influyen en el desempeño de la red, las variables relevantes a medir, y se evaluará la situación actual del control

y monitoreo del tráfico, especialmente en relación con las solicitudes y peticiones en línea.

1.6.2 BENEFICIARIOS DEL PROYECTO

Los beneficiarios de este proyecto incluyen al Departamento de Tecnologías de la Información de la Universidad Estatal Península de Santa Elena (UPSE), así como al personal administrativo y docente de la Facultad de Sistemas y Telecomunicaciones (FACSISTEL), quienes dispondrán de una herramienta para analizar y monitorear el tráfico de red con mayor eficiencia. De manera complementaria, la comunidad académica de la UPSE también se ve beneficiada, ya que el sistema contribuye a fortalecer la seguridad informática institucional, mejorar la estabilidad de la navegación y ofrecer un entorno digital más confiable para el desarrollo de actividades académicas y administrativas.

1.6.3 VARIABLES DEL ESTUDIO

Variable Independiente: Técnicas de Minería de datos empleadas en el análisis de páginas web.

Variable Dependiente: Capacidad de detección de Amenazas Persistentes Avanzadas (ATP) en páginas web.

1.6.4 ANÁLISIS DE RECOLECCIÓN DE INFORMACIÓN

- **Técnicas**

  Estado del arte, análisis de logs, revisión documental y análisis de registros de tráfico de red.

- **Instrumentos**

  En este proyecto, la principal fuente de información proviene de los registros del firewall (logs), los cuales contienen detalles sobre conexiones, protocolos utilizados, direcciones IP de origen y destino, cantidad de paquetes enviados y recibidos, eventos sospechosos y sus respectivas marcas de tiempo.

  Estos logs fueron extraídos directamente desde el dispositivo de seguridad Fortigate 2500E posteriormente preprocesados a limpieza, filtrado y

normalización, eliminando datos incompletos, duplicados o inconsistentes, y estandarizando los formatos temporales para facilitar su análisis.

Una vez depurada, la información fue organizada en un dataset estructurado, adecuado para el uso de técnicas de minería de datos y modelos de aprendizaje profundo, lo que permite identificar patrones normales, anomalías y posibles signos de Amenazas Persistentes Avanzadas (APT) en páginas web.

- **Población**

  La población determinada de esta investigación es de estudiantes de la Facultad de Sistemas y Telecomunicaciones, así como del personal administrativo y docente de la Facultad de Sistemas y Telecomunicaciones, de los que les viene el tráfico de red a estudiar.

1.6.5 ANÁLISIS DE RECOLECCIÓN DE DATOS

Para el desarrollo del proyecto de tesis se realizó una recolección de datos los registros del firewall los cuales proporcionan información detallada sobre conexiones, protocolos, paquetes enviados y recibidos, así como marcas de tiempo de los eventos, que son los llamados log. Posteriormente, los datos recolectados fueron sometidos a procesos de limpieza y normalización, eliminando registros incompletos o duplicados y garantizando la uniformidad en los formatos utilizados para representar el tiempo. Esta preparación asegura que el dataset resultante sea eficiente para entrenar modelos de aprendizaje profundo y aplicar técnicas de minería de datos facilitando la identificación de patrones normales y anómalos en el tráfico de red. De esta manera, la información obtenida es importante para mejorar la detección de APT respetando el cumplimiento de normas de seguridad y privacidad institucionales.

Además, los datos procesados permitieron realizar análisis exploratorios mediante estadísticas descriptivas y visualización de patrones de tráfico en la herramienta Rstudio, esto facilitó la identificación de tendencias, picos de actividad y comportamientos atípicos que podrían estar relacionados con amenazas persistentes avanzadas. La limpieza y organización del dataset permitió dividir la información

en conjuntos de entrenamiento, validación y prueba, asegurando un proceso riguroso para la construcción y evaluación de los modelos de detección. Gracias a este enfoque, la recolección de datos no solo aportó información técnica, sino que también fortaleció la calidad y confiabilidad del análisis realizado en el proyecto.

1.7 METODOLOGÍA DE DESARROLLO DEL PROYECTO

El proyecto se basa en la metodología CRISP-DM desarrollada por un consorcio de empresas en 1997 [18], que integra fases estructuradas de análisis de datos para obtener conocimiento útil y aplicable.

Basado en los lineamientos de la metodología, el proyecto estará enfocado en cinco fases que se detallan a continuación:

Fase 1: Compresión del Problema

Esta fase se centra en entender a profundidad el contexto y el alcance del proyecto, definiendo los objetivos del modelo y las necesidades del análisis. En la práctica, se documenta el problema, se identifican los posibles desafíos y se establece un marco que guiará todo el proceso de minería de datos.

Fase 2: Análisis de los datos

Esta fase busca comprender la naturaleza de los datos y sus posibles patrones se realizan acciones como revisar los datasets preprocesados, verificar inconsistencias y generar gráficos exploratorios que permitan visualizar tendencias y relaciones entre variables, estudiando así la información disponible para evaluar su calidad y relevancia.

Fase 3: Preparación de los datos

Esta fase consiste en transformar la información cruda en un formato adecuado para el análisis. Se investigan los algoritmos más apropiados, se limpian los datos, se optimiza su calidad y se filtra el ruido. Además, se identifican patrones preliminares que facilitan el posterior modelado y aseguran que los datos sean confiables.

Fase 4: Modelado

Esta fase implica conceptualizar y construir representaciones matemáticas o de aprendizaje automático que permitan capturar los patrones presentes en los datos.

En la práctica, se seleccionan modelos específicos y se miden indicadores clave de desempeño (KPI) para determinar cuál se ajusta mejor a los objetivos del proyecto.

Fase 5: Evaluación

En esta fase se analiza la efectividad y confiabilidad del modelo construido. Se interpretan los resultados utilizando diversas herramientas y métricas, verificando que el modelo cumpla con los objetivos planteados y que sus conclusiones sean válidas y generalizables.

Fase 6: Explotación

La fase de explotación busca aplicar el conocimiento obtenido para facilitar la toma de decisiones. Conceptualmente, implica convertir los resultados en información útil y accesible. En la práctica, se desarrollan dashboard y visualizaciones que permiten a los usuarios finales interactuar con los hallazgos y utilizarlos de manera estratégica.

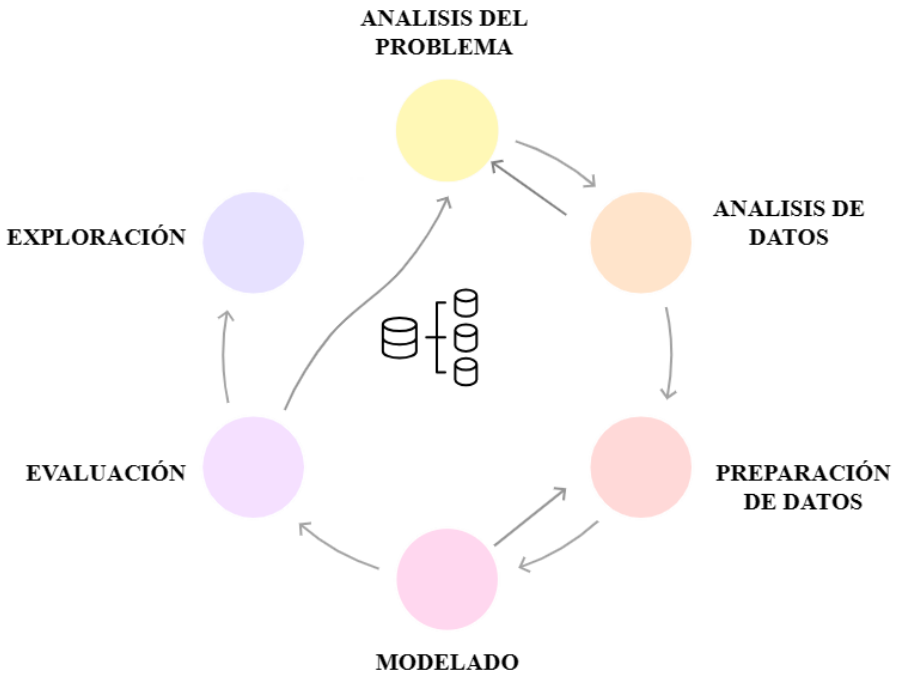


Ilustración 4: Fases de la metodología CRISP-DM: Elaboración Propia

CAPÍTULO II: PROPUESTA

2.1 MARCO CONTEXTUAL

La Universidad Estatal Península de Santa Elena (UPSE), situada en la provincia de Santa Elena, Ecuador, es una institución educativa de reconocida trayectoria en la formación de profesionales en diversas disciplinas. Fundada el 2 de julio de 1998 por medio de la Ley N° 110, su creación fue oficializada con la publicación en el suplemento del Registro Oficial N° 366 el 22 de julio de 1998. Su sede se encuentra en la avenida principal de La Libertad-Santa Elena, en el cantón La Libertad. La universidad dispone de una infraestructura moderna y tecnológica, que facilita a estudiantes y docentes realizar sus actividades académicas con eficiencia y efectividad. [19] .

Misión

Formar profesionales que aportan al desarrollo sostenible, contribuye a la solución de los problemas de la comunidad y promueve la cultura [20].

Visión

Ser reconocida por su calidad académica, el impacto de sus investigaciones y su aporte al desarrollo de la sociedad [20].

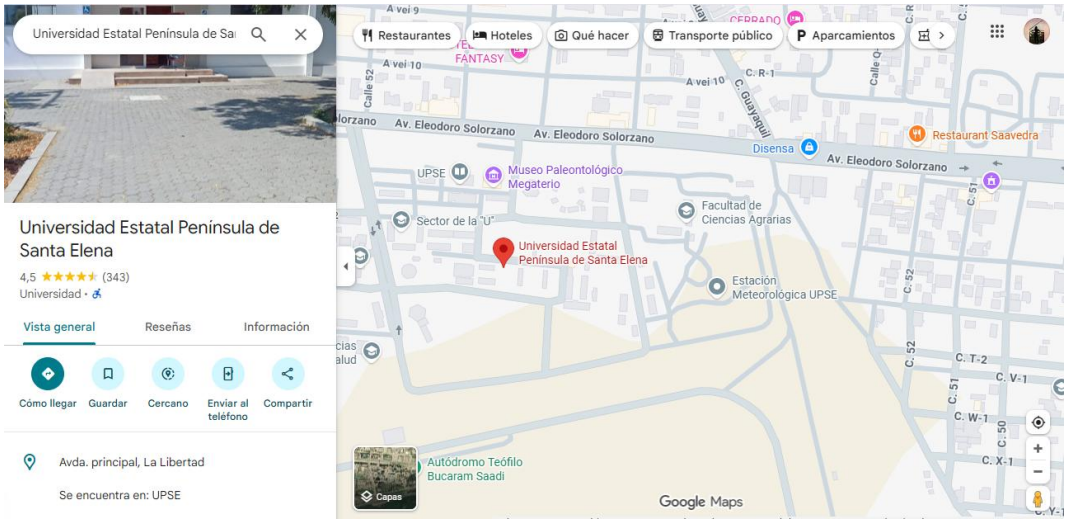


Ilustración 5: Vista del mapa de UPSE

En el mundo digital actual, las páginas web se han solidificado como el principal medio de interacción entre usuarios y organizaciones, lo que permite convertirlo en

el blanco fácil y crítico para los ciberatacantes. Los riesgos más cruciales incluyen las amenazas persistentes avanzadas (APT), que se caracterizan por su sofisticación, sigilo y permanencia en los sistemas afectados. Estas amenazas no solo buscan interrumpir los servicios, sino también infiltrar y mantener dentro de la infraestructura web al delincuente informático con el fin de robar información sensible, manipular datos o comprometer la reputación de la organización [21].

En consonancia con el contexto actual, se registra un aumento significativo en los incidentes relacionados con Amenazas Persistentes Avanzadas (APT) impulsado por la expansión de la digitalización, el comercio electrónico, la banca en línea y las plataformas de servicios web. El panorama pone de manifiesto que los mecanismos tradicionales de detección, basados en firmas o reglas estáticas, resultan insuficientes, dado que las amenazas persistentes avanzadas emplean técnicas como el cifrado y patrones de comportamiento dinámico lo cual complica sus mecanismos de identificación.

El desafío de la minería de datos proviene de una opción tecnológica fundamental; esta disciplina se encarga de extraer patrones ocultos, correlaciones y anomalías en grandes volúmenes de información, como los datos generados por el tráfico web, registros de acceso y, especialmente, las actividades de los usuarios. Mediante algoritmos de clasificación, detección temprana de ataques y la predicción de comportamientos maliciosos, se apoya la toma de decisiones en ciberseguridad [22].

Por consiguiente, tanto en el ámbito regional como local, donde el crecimiento de entornos digitales es cada vez mayor, la aplicación de técnicas de minería de datos en la detección de ATP en páginas web juega un papel crucial. No solo contribuye a la protección de la información, sino que también fortalece la confianza de los usuarios frente a los servicios en línea. De esta manera, se presenta un marco contextual que conecta la importante necesidad de proteger la información en la web activa con el potencial de la minería de datos como una herramienta analítica avanzada, ayudándonos a enfrentar mejor las amenazas sofisticadas.

2.1.1 BASE LEGAL

Constitución de la Republica del Ecuador

Artículo 66.- Derecho a la protección de datos de carácter personal Garantiza a todas las personas el derecho a la protección de sus datos personales. Este derecho implica que el individuo tiene el control y la decisión sobre la información y los datos que le conciernen, así como su correspondiente protección. Cualquier recolección, archivo, procesamiento, distribución o difusión de estos datos o información requerirán la autorización del titular o el mandato de la ley [23].

Ley Orgánica de Protección de Datos Personales

Art. 10.- Principios

Seguridad de la información personal: Las personas encargadas de gestionar la información personal deben implementar todas las medidas de seguridad necesarias para protegerla, de acuerdo con el estado de la técnica, para proteger los datos personales frente a cualquier riesgo, amenaza o vulnerabilidad. Estas medidas pueden ser de tipo organizativo, técnico, y deben ser adecuadas considerando la naturaleza de los datos y el contexto en el que se encuentran [24].

Capítulo VI: Seguridad de datos personales

Art. 37.- Seguridad de datos personales

Los encargados o responsables del tratamiento de datos personales deben cumplir con el principio de seguridad, adoptando medidas fundamentadas en las prácticas más destacadas de seguridad integral y evaluando los costos de aplicación en función de la naturaleza, el alcance, el contexto y los objetivos del tratamiento. Además, deben tener en cuenta los riesgos probables que puedan presentarse [24].

Art. 40.- Análisis de riesgo, amenazas y vulnerabilidades

Para evaluar los riesgos, amenazas y vulnerabilidades, el responsable del tratamiento de datos debe emplear un método de análisis que considere aspectos

como los detalles específicos del tratamiento, las particularidades de los participantes involucrados, y las categorías y la cantidad de información personal tratada [24].

Art. 41.- Determinación de medidas de seguridad aplicables

Las medidas de seguridad deben ser definidas tomando en cuenta, entre otros factores, los resultados del análisis de peligros, amenazas y vulnerabilidades, el tipo de información personal, las características de los participantes, así como los historiales de acceso no autorizado o modificaciones accidentales o deliberadas a los datos. Estas medidas deben ser autorizadas por el estado de la técnica y ser aplicables al contexto específico del tratamiento de la información [24].

Código Orgánico Integral Penal (COIP)

Sección Tercera: Delitos relacionados con la seguridad de los sistemas de información y comunicación

Art. 229.- Acceso ilícito a la base de datos

Quien divulgue datos almacenados en documentos, ficheros o bases de datos, a través de sistemas informáticos o de telecomunicaciones, vulnerando la privacidad de los individuos, será sancionado con una pena privativa de libertad de 1 a 3 años [25].

Art. 230.- Intercepción ilegal de datos

Se sancionará con pena privativa de libertad de 3 a 5 años a [25]:

- Quien, sin orden judicial previa, intercepte, desvíe, escuche o registre información digital en su origen, destino o dentro de un sistema digital [25].
- Quien elabore, comercialice, o transmita enlaces maliciosos, certificados o páginas web fraudulentas [25].
- Quien fabrique o posea dispositivos electrónicos destinados a la realización de estas acciones [25].

Art. 232.- Ataque a la integridad de sistemas informáticos

La persona que dañe, destruya, altere o deteriore datos informáticos, correos

electrónicos, o sistemas de telecomunicaciones será sancionada con pena privativa de libertad de 3 a 5 años [25].

Art. 234.- Acceso no autorizado a un sistema informático o de telecomunicaciones

La persona que, sin autorización, acceda total o parcialmente a un sistema informático o de telecomunicaciones, o permanezca en él sin el consentimiento de quien lo habilitó, con el fin de explotar o modificar el sistema de manera ilegítima, será sancionada con una pena privativa de libertad de 3 a 5 años [25].

2.2 MARCO TEÓRICO

2.2.1 Desarrollo de modelos para la detección de phishing

El estudio se enfoca en el desarrollo de un modelo de ciberseguridad que pueda identificar ataques de phishing a tiempo utilizando métodos de procesamiento del lenguaje natural e inteligencia artificial. El modelo tiene como propósito examinar las cabeceras de los correos electrónicos para detectar patrones relacionados con mensajes maliciosos, dado que el phishing es una de las amenazas más comunes y dañinas en el espacio digital. Para estructurar y sistematizar el proceso, se utiliza la metodología ASUM-DM de IBM, que incluye desde la identificación del problema y la recopilación de información hasta la preparación de los datos y la creación y evaluación de modelos predictivos [26].

Después de procesar los datos, el estudio analiza la evolución del phishing y las técnicas actuales para detectarlo, luego entrena el modelo utilizando redes neuronales y métodos de aprendizaje automático. Los resultados muestran que se tiene una precisión elevada en la identificación de correos electrónicos fraudulentos, lo cual confirma la eficacia del enfoque y la importancia de la inteligencia artificial para evitar las amenazas digitales.

2.2.2 Análisis exploratorio de ataques informáticos aplicando minería de datos

El estudio doctoral titulado "Análisis exploratorio de ataques cibernéticos mediante el uso de herramientas de minería de datos para gestionar la seguridad de redes inalámbricas en universidades arequipeñas" pone en evidencia la relevancia del

empleo de métodos de minería de datos para salvaguardar las infraestructuras tecnológicas en entidades educativas superiores. Siguiendo esta misma línea, el estudio presente utiliza la metodología CRISP-DM para estructurar el proceso de clasificación, asociación y segmentación de datos con el fin de capturar y examinar el tráfico de redes tanto alámbricas como inalámbricas en un campus universitario. Esta perspectiva posibilita distinguir los diferentes tipos de ataques, su duración, procedencia y grado de agresividad, así como analizar la eficacia de las medidas de seguridad que se han puesto en marcha [27].

El análisis indica que la minería de datos permite evaluar la eficacia de la seguridad en las redes universitarias y prever posibles ataques. Se detectó un incremento en la ejecución de código y la presencia de spyware, siendo REPTree y CHAID los algoritmos más confiables, la falta de auditorías sistemáticas hace que las instituciones sean más vulnerables, lo que subraya la importancia de usar estas técnicas para fortalecer la seguridad informática [27].

2.2.3 Nuevas perspectivas en amenazas persistentes avanzadas

La tesis *"Nuevas perspectivas en amenazas persistentes avanzadas"* destaca que este tipo de ataques son sofisticados, dirigidos y personalizados, representando un riesgo crítico para organizaciones que manejan información sensible o infraestructuras estratégicas. En los últimos años, su análisis ha despertado gran interés en la comunidad científica, impulsando la creación de modelos y herramientas para su detección temprana. El uso de inteligencia artificial y aprendizaje automático se plantea como una solución eficaz para identificar, alertar y predecir estos ataques, reduciendo el tiempo de permanencia de los intrusos en las redes [8].

El trabajo propone un modelo teórico basado en el ciclo de vida del ataque, apoyado en técnicas de aprendizaje automático, que se fundamenta en una revisión bibliográfica sobre las amenazas persistentes avanzadas y sus aplicaciones de detección en el ámbito de la ciberseguridad

2.3 MARCO CONCEPTUAL

2.3.1 Amenazas Persistentes Avanzadas(APT)

Las amenazas persistentes avanzadas (APT) son operaciones de ciberataque altamente sigilosas cuyo propósito es infiltrarse en sistemas críticos durante largos periodos para robar información sensible, llevar a cabo espionaje digital o comprometer infraestructuras clave. A diferencia de amenazas más visibles, como el ransomware, los actores APT buscan mantenerse ocultos mientras consolidan y expanden su control dentro de la red objetivo [28].

2.3.1.1 Persistencia Operacional

Una APT se distingue por su habilidad para mantener el acceso a un sistema o red comprometida durante meses o incluso años sin ser detectada. Esto implica que, tras la intrusión inicial, el atacante implanta puertas traseras, persistencias ocultas, cuentas encubiertas o malware diseñado para regenerarse, de modo que incluso si se detecta algún indicio y se remedia una entrada, otras permanecen activas. Esta persistencia permite al atacante moverse lateralmente, recolectar información con calma y mantener control a largo plazo [28].

2.3.1.2 Orientación Estratégica

Las APT suelen dirigirse a blancos de alto valor: organizaciones, instituciones o empresas con activos críticos como propiedad intelectual, secretos comerciales, infraestructura sensible o datos confidenciales. Los atacantes eligen cuidadosamente sus objetivos en función del valor de la información o de la repercusión estratégica no se trata de ataques masivos al azar, sino de operaciones dirigidas con un propósito específico [28].

2.3.1.3 Sofisticación Técnica

Para infiltrarse y mantenerse ocultas, estas amenazas emplean técnicas avanzadas: malware a medida, vulnerabilidades de día cero (zero-day), exploits desconocidos, ataques muy dirigidos (spear-phishing), uso de herramientas personalizadas, comunicaciones cifradas, métodos de evasión, y aprovechar incluso funcionalidades legítimas del sistema (“living-off-the-land”). Esta complejidad técnica las hace capaces de evadir controles tradicionales de seguridad [28].

2.3.1.4 Adaptabilidad Continua

Las APT no siguen siempre el mismo patrón: modifican sus tácticas, técnicas y procedimientos (TTPs) conforme las defensas del objetivo cambian. Esto significa que si se detecta una infección o una técnica usada es bloqueada, los atacantes pueden cambiar de estrategia nuevos vectores de entrada, diferentes malware, distintos métodos de persistencia para seguir operando sin ser detectados [28].

2.3.1.5 Ciclo de vida de los APTs:

2.3.1.6 Fase de reconocimiento

En esta etapa inicial, los atacantes recopilan información sobre el objetivo: infraestructura de red, empleados, sistemas operativos, servicios expuestos, posibles vulnerabilidades, y cualquier dato que pueda facilitar un acceso futuro. El reconocimiento puede implicar minería de información pública, ingeniería social, escaneo de puertos y recolección de huellas digitales (fingerprinting). Este paso es clave para planear un ataque dirigido, ya que permite definir el vector de entrada más efectivo [29].

2.3.1.7 Fase de Infiltración Inicial

Aquí el atacante ejecuta el ataque inicial para obtener acceso al sistema o red objetivo. Esto puede lograrse mediante phishing dirigido, explotación de vulnerabilidades, uso de credenciales robadas, ataques de día cero o malware, entre otros métodos. Si el ataque tiene éxito, el atacante compromete al menos un sistema de la red, lo que le permite comenzar la fase de compromiso persistente [29].

2.3.1.8 Fase de Establecimiento de Persistencia

Después del acceso inicial, el atacante instala mecanismos que le permitan mantener el control a largo plazo, incluso si las defensas detectan y eliminan parte del malware. Esto puede incluir backdoors, rootkits, cuentas ocultas, malware con recuperación automática o persistencia en múltiples sistemas. El objetivo es asegurar que el acceso no dependa de una única puerta y que se pueda reconectar tras periodos de limpieza parcial [29].

2.3.1.9 Fase de movimiento Lateral

Con persistencia garantizada, el atacante expande su acceso dentro de la red comprometida, moviéndose de un sistema inicial a otros sistemas internos, elevando privilegios, explorando recursos internos, identificando activos críticos y adquiriendo mayor control. Este movimiento lateral le permite llegar a servidores sensibles, bases de datos o sistemas de administración, extendiendo su huella sin ser detectado [29].

2.3.1.10 Fase de Exfiltración de Datos

Llegado este punto, el atacante recopila información valiosa —datos sensibles, propiedad intelectual, secretos comerciales, credenciales, registros de usuarios u otro tipo de datos críticos y los extrae de la red a servidores externos o medios controlados por el atacante. Esta fase puede implicar compresión, cifrado, fragmentación o técnicas de ofuscación para evitar la detección. En algunos casos, también puede incluir sabotaje, destrucción de datos o preparación para futuros ataques [29].

2.3.2 Modelos de detección de anomalías

La detección de anomalías resulta clave en sectores como finanzas, retail y ciberseguridad, pero cualquier organización debería adoptarla, ya que permite identificar automáticamente comportamientos atípicos que comprometen la integridad de los datos y la operación. En áreas como la banca, este proceso facilita descubrir transacciones irregulares y patrones anómalos asociados a fraude, fortaleciendo la protección de la información. Sin estos mecanismos, una empresa queda expuesta a pérdidas económicas, daños en la reputación, brechas de seguridad y filtración de datos sensibles, situaciones que pueden deteriorar la confianza del cliente de forma permanente [30].

2.3.2.1 Detección basada en firmas

Los métodos de detección basados en firmas identifican amenazas comparando el tráfico de red con un conjunto de patrones previamente conocidos, llamados *firmas de ataque*, que representan características específicas de malware o comportamientos maliciosos. Un IPS de este tipo utiliza una base de datos de firmas

para evaluar cada paquete y, si detecta una coincidencia, activa una respuesta de seguridad. Estas bases deben actualizarse de forma continua para incorporar nuevas amenazas, ya que los ataques emergentes o variantes no registradas pueden pasar desapercibidos y evadir este mecanismo de protección [31].

2.3.2.2 Detección basada en Heurística

La heurística es un método de detección proactiva que identifica código malicioso sin depender de firmas predefinidas. Este enfoque analiza el comportamiento de un archivo y lo compara con patrones asociados a actividades potencialmente dañinas, asignando puntajes a cada acción observada. Si la suma de estos puntajes supera un umbral establecido, el archivo se clasifica como una posible amenaza nueva. Su relevancia radica en el volumen creciente de malware que surge diariamente, lo que hace insuficiente basarse únicamente en firmas tradicionales. Por ello, la heurística complementa los métodos basados en firmas al permitir la detección temprana de amenazas recién creadas o desconocidas [32].

2.3.2.3 Detección basada en Deep Learning

La diferencia esencial entre machine learning y deep learning radica en la complejidad de sus arquitecturas neuronales. Los modelos de machine learning tradicional emplean redes con una o pocas capas computacionales, mientras que el deep learning utiliza arquitecturas profundas compuestas por decenas, cientos o incluso miles de capas para entrenar modelos altamente complejos. Además, el aprendizaje supervisado del machine learning requiere datos estructurados y etiquetados, mientras que el deep learning puede trabajar con datos no estructurados mediante aprendizaje no supervisado, extrayendo de forma autónoma patrones, características y relaciones a partir de la información en bruto. Estos modelos también pueden ajustar y mejorar sus propias predicciones de manera iterativa, lo que incrementa significativamente su precisión [33].

2.3.2.4 Detección basada en Listas

La detección basada en listas es un método de seguridad que compara eventos, direcciones IP, dominios, aplicaciones o comportamientos con conjuntos predefinidos de elementos permitidos (listas blancas) o bloqueados (listas negras).

Si un elemento coincide con una lista negra se considera una posible amenaza y se bloquea; si coincide con una lista blanca se permite su ejecución sin restricciones. Este enfoque es eficaz para filtrar tráfico conocido y controlar accesos, pero tiene la limitación de que no detecta amenazas nuevas o desconocidas que no estén previamente registradas en las listas [34].

2.3.3 Página web

Una página web es un documento digital accesible a través de la World Wide Web (WWW), que se muestra mediante un navegador de internet. Puede contener texto, imágenes, videos, enlaces y otros elementos interactivos, presentados en un formato diseñado para la visualización del usuario, estas páginas se desarrollan utilizando lenguajes de programación como HTML (HyperText Markup Language), CSS (Cascading Style Sheets) y JavaScript, que permiten estructurar el contenido y agregar funciones interactivas formando lo que se denomina un sitio web, que puede ser estático (con contenido fijo) o dinámico (que se adapta y cambia según la interacción del usuario) [35].

2.3.3.1 Seguridad de página web

¡El internet es un lugar riesgoso! Es común que nos enteremos de sitios web que han sido atacados y ya no están disponibles, o que muestran información alterada (y a menudo dañina) en sus páginas principales. En otros casos de alto nivel, se han filtrado millones de contraseñas, direcciones de correo electrónico y detalles de tarjetas de crédito al dominio público, lo que ha puesto en riesgo financiero y personal a los usuarios del sitio web. El objetivo de la seguridad web es evitar asaltos de esta (o cualquier otra) categoría. Más formalmente, la seguridad consiste en proteger los sitios web contra el acceso, uso, alteración, eliminación o interrupción no autorizados [36].

2.3.4 Análisis de tráfico de red

El Análisis de Tráfico de Red (NTA) consiste en la captura, procesamiento, interpretación y representación de los flujos de datos que transitan por una infraestructura de comunicaciones, con el propósito de garantizar la operación eficiente de la red, optimizar su desempeño, identificar comportamientos anómalos

o potenciales amenazas, y apoyar la detección y resolución de incidentes operativos o de seguridad [37].

De acuerdo con fabricantes como IBM y Cisco, este tipo de análisis permite obtener una visibilidad granular del comportamiento real del tráfico y de las interacciones entre los nodos de la red, proporcionando un nivel de detalle superior al que ofrecen exclusivamente los registros tradicionales de cortafuegos u otros dispositivos perimetrales [38].

2.3.4.1 Análisis por Flujo

El análisis de flujo implica revisar metadatos producidos por dispositivos de red como NetFlow, sFlow o IPFIX para describir y evaluar las comunicaciones entre hosts. Estos registros resumen información esencial del tráfico, como IPs, puertos, protocolos, volumen y duración, permitiendo identificar patrones, anomalías, uso de ancho de banda y posibles actividades maliciosas, sin necesidad de inspeccionar el contenido completo de los paquetes [39].

2.3.4.2 Análisis de Cabecera

El análisis de cabecera consiste en examinar exclusivamente los campos de metadatos que contiene la cabecera de los paquetes de red como dirección IP de origen y destino, puertos, protocolo, TTL, flags, longitud, etc. Este método facilita filtrar, enrutar o descartar paquetes según reglas establecidas (como políticas, filtros de firewall, etc.) siendo eficiente para manejar grandes volúmenes de tráfico con un bajo costo computacional. Es útil para detectar rápidamente patrones de comunicación, gestionar el control de acceso, bloquear por IP/puerto y aplicar reglas de seguridad, logrando un equilibrio entre visibilidad y rendimiento. [40].

2.3.4.3 Análisis de Paquete

El análisis de paquetes consiste en examinar el tráfico de red en su forma más básica para identificar fallas o anomalías que no se detectan con métodos generales; aunque los paquetes son solo las unidades mínimas del flujo de datos, su inspección detallada se vuelve esencial cuando surgen problemas complejos, como fugas de información sin señales visibles, bajo rendimiento de aplicaciones sin causa aparente, posibles compromisos en equipos o redes, explotación del WiFi por

terceros o cuellos de botella en servidores que no muestran alto volumen de tráfico [41].

2.3.4.4 Análisis de Anomalías

El análisis de anomalías, también conocido como detección de outliers, consiste en identificar patrones, eventos o registros que se desvían significativamente del comportamiento esperado en un conjunto de datos. Estas desviaciones pueden representar errores, fraudes, fallos en el sistema o incluso ataques maliciosos. En ciberseguridad, esta técnica es fundamental para detectar intrusiones desconocidas, ataques de día cero, tráfico inusual en la red y comportamientos atípicos de los usuarios que podrían indicar compromisos de seguridad. Su valor radica en la capacidad de identificar amenazas emergentes que no pueden detectarse mediante firmas o reglas predefinidas [42].

2.3.5 IP

Una dirección IP es un identificador numérico estructurado en cuatro octetos separados por puntos, como 192.158.1.38, donde cada octeto puede tomar valores entre 0 y 255. Este formato define un espacio de direccionamiento que abarca desde 0.0.0.0 hasta 255.255.255.255. La asignación de estas direcciones no es arbitraria: la Autoridad de Números Asignados de Internet (IANA), entidad operada por la ICANN, administra y distribuye de manera jerárquica los bloques de direcciones IP, siguiendo procedimientos y criterios técnicos establecidos para garantizar su organización y disponibilidad global [43].

2.3.6 Mineria de datos

La minería de datos es un procedimiento que se lleva a cabo con la ayuda de una computadora y sirve para examinar y procesar conjuntos extensos de información durante los análisis. Las organizaciones tienen la capacidad de revelar patrones y conexiones encubiertas en sus datos debido a las técnicas y herramientas de minería de datos. La minería de datos convierte los datos brutos en conocimientos aplicables. Las empresas emplean este conocimiento para solucionar problemas, examinar las repercusiones futuras de sus decisiones y ampliar su margen de ganancias [44].

La capacidad de detectar patrones y vínculos en grandes cantidades de datos que provienen de diversas fuentes es la mayor ventaja de la minería de datos. La minería de datos brinda las herramientas necesarias para aprovechar al máximo los big data y transformarlos en inteligencia que se puede accionar, considerando el aumento de la cantidad de información disponible (que proviene de fuentes tan diversas como redes sociales, sensores remotos o reportes con mayor detalle sobre la actividad del mercado y el movimiento de productos). De hecho, puede funcionar como un mecanismo para pensar fuera de la caja [45].

2.3.7 Minería de datos aplicada a la ciberseguridad

La minería de datos (o *data mining*) en ciberseguridad es el proceso de recolectar, procesar y analizar grandes volúmenes de datos generados por sistemas, redes y dispositivos, con el fin de descubrir patrones, anomalías, correlaciones y tendencias que permitan detectar amenazas, intrusiones, fraudes o comportamientos maliciosos, y mejorar la toma de decisiones de seguridad. Esto implica usar técnicas como aprendizaje automático, estadística, detección de anomalías, clustering, reglas de asociación, etc., aplicadas a datos tanto estructurados como no estructurados, para prevenir o mitigar riesgos de seguridad informática [46].

2.3.8 Machine Learning

El aprendizaje automático es una sección de la inteligencia artificial (IA) centrada en entrenar a computadoras y máquinas para imitar el modo en que aprenden los humanos,
llevar a cabo actividades de manera independiente y perfeccionar su eficiencia y e xactitud mediante la experiencia y la asimilación de un mayor volumen de datos [47].

2.3.8.1 Aprendizaje Supervisado

El aprendizaje supervisado entrena modelos usando datos previamente etiquetados, lo que permite aprender la relación entre las entradas y las salidas esperadas. Durante el entrenamiento, el algoritmo analiza grandes volúmenes de datos para identificar patrones y luego evalúa su desempeño con datos de prueba o mediante

validación cruzada para comprobar su capacidad de generalización. Para ajustar el modelo, se emplean métodos de optimización como el descenso de gradiente y su variante estocástica (SGD), los cuales minimizan la función de pérdida, es decir, la diferencia entre las predicciones generadas y los valores reales [48].

2.3.8.2 Aprendizaje no Supervisado

El aprendizaje no supervisado es un enfoque de la inteligencia artificial en el que los modelos procesan datos sin etiquetar y son capaces de identificar patrones, estructuras o relaciones sin intervención humana ni instrucciones explícitas. A diferencia del aprendizaje supervisado, este método permite que el sistema descubra por sí mismo agrupamientos, tendencias o características relevantes dentro de los datos. Actualmente, este tipo de aprendizaje es clave en aplicaciones que optimizan procesos y decisiones, como sistemas de recomendación, traducción automática o generación de contenido, contribuyendo a mejorar la eficiencia y reducir costos en diversos sectores [49].

2.3.9 Técnicas de minería de datos relevantes

2.3.9.1 Clustering

El Clustering, o agrupamiento, es una técnica de minería de datos no supervisada que consiste en dividir un conjunto de datos en grupos (clusters) de objetos similares entre sí, de manera que los elementos dentro de un mismo grupo presenten alta similitud y los elementos de diferentes grupos sean lo más disímiles posible. En ciberseguridad, el clustering se utiliza para descubrir patrones ocultos en grandes volúmenes de datos sin necesidad de que estos estén previamente etiquetados. Esto lo hace muy útil para identificar comportamientos anómalos, crear perfiles de tráfico normal en redes, detectar nuevos tipos de ataques o malware, y realizar segmentación de eventos de seguridad [50].

2.3.9.2 Clasificación

La clasificación es una técnica de minería de datos supervisada que consiste en asignar una etiqueta o categoría predefinida a cada instancia de un conjunto de datos [51], utilizando un modelo entrenado con datos históricos que ya poseen dichas etiquetas. Su objetivo es predecir la clase a la que pertenece un nuevo dato

basándose en patrones aprendidos. Para ello, se construye un modelo (por ejemplo, un árbol de decisión, un clasificador bayesiano o una red neuronal) a partir de un conjunto de entrenamiento y luego se utiliza para clasificar datos nuevos o desconocidos [52].

2.3.9.3 Reglas de Asociación

Las reglas de asociación son una técnica de minería de datos que busca descubrir relaciones frecuentes entre conjuntos de datos, generalmente expresadas como "si X ocurre, entonces Y tiene cierta probabilidad de ocurrir". Su meta es detectar patrones de co-ocurrencia que ofrezcan información valiosa para la toma de decisiones. En ciberseguridad, esta técnica se emplea para correlacionar alertas de distintos sistemas, detectar secuencias de eventos que preceden un ataque y identificar comportamientos sospechosos en grandes volúmenes de registros o tráfico de red, lo que ayuda a prever amenazas y reforzar las defensas del sistema [53].

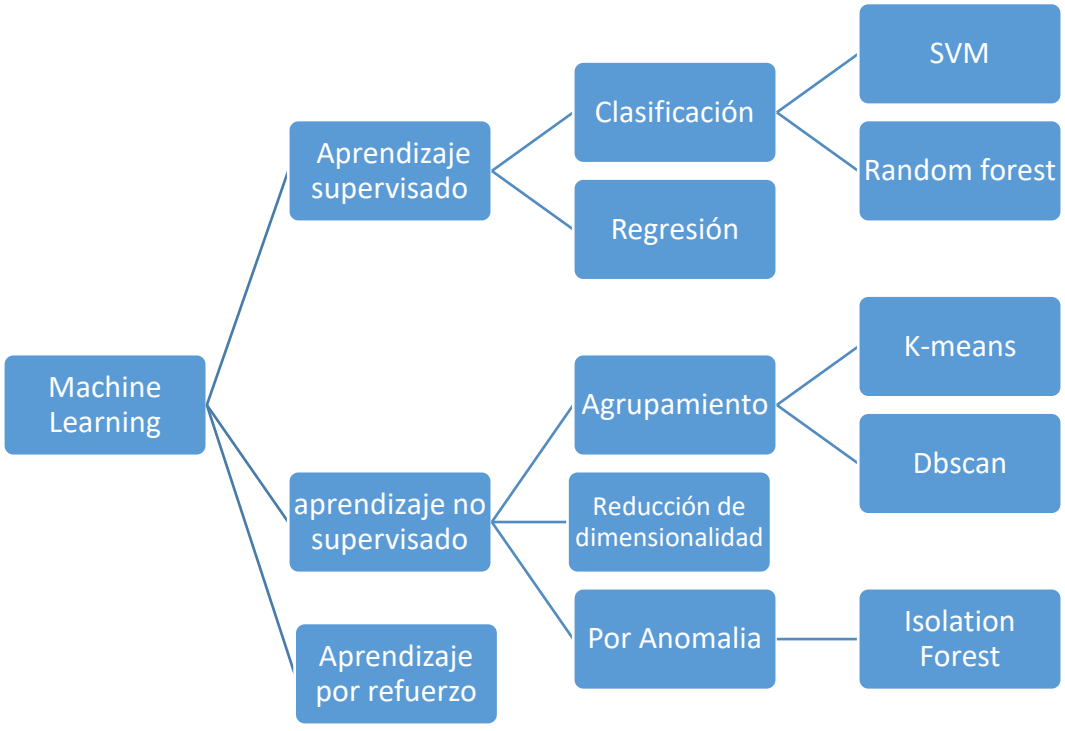


Ilustración 6:Tipo de Aprendizaje:

2.3.10 CRISP-DM – Proceso Estándar para Minería de datos entre Industrias

2.3.10.1 Fase de Análisis de los Datos.

La fase consiste en examinar los datos disponibles para comprender su estructura, calidad y distribución, incluye la identificación de valores faltantes, outliers, correlaciones y patrones iniciales. Esta fase permite validar la relevancia de los datos respecto al problema y determinar necesidades de transformación o limpieza antes del modelado [54].

2.3.10.2 Fase de Preparación de los Datos.

Implica seleccionar, limpiar, transformar, integrar los datos, la reducción de dimensionalidad y construcción de atributos necesarios para el modelado. Su objetivo es generar un conjunto de datos final óptimo para los algoritmos de minería de datos [54].

2.3.10.3 Fase de Modelado.

En esta etapa, se seleccionan y aplican técnicas de minería de datos ajustando los parámetros, probando diferentes algoritmos y generando modelos predictivos o descriptivos utilizando el dataset preparado. También se evalúan variaciones del modelo para identificar la mejor configuración [54].

2.3.10.4 Fase de Evaluación.

Se evalúa el rendimiento del modelo mediante métricas estadísticas, validaciones cruzadas y pruebas adicionales. La finalidad es determinar si el modelo cumple los objetivos del negocio y si es confiable para su implementación. También se revisan posibles riesgos o limitaciones [54].

2.3.10.5 Fase de Explotación

Consiste en implementar el modelo en un entorno real, automatizar su uso y monitorizar su desempeño. Incluye la generación de reportes, integración del modelo en procesos operativos y planificación del mantenimiento. Representa la fase donde el conocimiento generado aporta valor a la organización [54].

2.3.11 Herramientas

2.3.11.1 Base de datos

Es conocido como una recopilación organizada de información o también datos estructurados, que usualmente se almacena de manera digital en un sistema informático. Esencialmente, una base de datos esta precisamente ejecutada por un sistema de gestión de bases de datos (DBMS). Tanto el DBMS y los datos, en conjunto manifiesta conexión sobre la aplicación en plataformas que tengan el funcionamiento de generar procesos de consultas de datos [55].

2.3.11.2 Python

Python es un lenguaje de programación que se usa extensamente en el machine learning (ML), la ciencia de datos, el desarrollo de software y las aplicaciones web. Dado que es eficiente, sencillo de aprender y se puede ejecutar en diversas plataformas, los desarrolladores usan Python. El software Python, que es de descarga gratuita y se acopla correctamente a todo tipo de sistemas, acelera el desarrollo [56].

2.3.11.3 Jupyter Notebook

Esta aplicación favorece la elaboración de documentos que posibilitan la combinación de código con otros componentes, como texto enriquecido, imágenes, vínculos y demás. Esto ha llevado a que la comunidad de ciencia de datos lo use con frecuencia. Específicamente, Jupyter Notebook es un programa cliente/servidor que tiene la capacidad de ejecutarse de manera local en un navegador web sin requerir una conexión a Internet. No obstante, al ser una aplicación de red, también tiene la capacidad de ejecutarse a distancia mediante Internet [57].

2.3.11.4 Scikit-Learn

Scikit-learn es una de las bibliotecas de aprendizaje automático (ML) más empleadas. Este conjunto de herramientas de ciencia de datos, desarrollado en Python, acelera el modelado estadístico y la inteligencia artificial (IA) ML a través de una interfaz coherente. Contiene módulos fundamentales para la agrupación,

clasificación, regresión y reducción de dimensionalidad, todos desarrollados con las bibliotecas Matplotlib, NumPy y SciPy [58].

2.3.11.5 Matplotlib

La librería matplotlib permite generar y personalizar distintos tipos de gráficos en dos dimensiones, como líneas, histogramas, espectros de potencia, diagramas de barras, gráficos de error y dispersión. El módulo matplotlib.pyplot incluye funciones que controlan tanto el contenido como la apariencia de las gráficas, permitiendo crear figuras, definir áreas de dibujo, trazar líneas y añadir etiquetas. Además, este módulo administra la figura y los ejes activos, de manera que las funciones de trazado actúan directamente sobre el espacio gráfico vigente [59].

2.3.11.6 Rstudio

R es una herramienta ampliamente utilizada para el análisis estadístico y el aprendizaje automático, ya que permite gestionar datos y realizar pruebas, modelos, análisis y gráficos. RStudio, integrado en IBM watsonx.ai Studio, funciona como un entorno de desarrollo para trabajar con scripts en R, aunque sólo está disponible en las implementaciones de watsonx alojadas en IBM Cloud y no en AWS. Desde RStudio se puede acceder a los archivos almacenados en el bucket de IBM Cloud Object Storage asociado al proyecto y, aunque es posible crear aplicaciones Shiny, estas no pueden desplegarse directamente en IBM watsonx [60].

2.3.11.7 Excel

Excel es posible gestionar datos numéricos de manera eficiente, automatizando la inserción de información mediante la función de Autorrellenar. A partir de los conjuntos de datos, la herramienta genera recomendaciones de gráficos que pueden construirse de forma inmediata. Además, Excel simplifica el análisis visual de tendencias y patrones usando formatos condicionales como barras de datos, escalas de color e íconos indicadores [61].

2.3.11.8 Cvs

Un archivo CSV (Valores Separados por Comas) es un formato de datos que se puede crear o editar desde Excel y que estructura la información usando valores separados por comas en lugar de emplear columnas estructuradas. Al almacenar

texto y números en este formato, los datos pueden transferirse fácilmente entre diferentes aplicaciones. Por ejemplo, es posible exportar los contactos de Google en un archivo CSV e importarlos posteriormente en Outlook [62].

2.3.11.9 Dashboard

Un dashboard es una herramienta visual interactiva que muestra de manera clara y dinámica métricas clave como Recall, Precision, F1-score, Accuracy y la curva ROC, entre otras. Se emplea para supervisar el rendimiento de modelos de clasificación y ayuda a interpretar resultados complejos a través de gráficas y visualizaciones. [63]. Además, ofrece una visión completa de las métricas de evaluación, facilitando decisiones fundamentadas en datos. Su habilidad para presentar la información de forma ordenada y fácil de entender permite un análisis ágil y eficaz, optimizando el proceso de interpretación y monitoreo del rendimiento del modelo. [63].

Ilustración 7: Importancia de un Dashboard: Data School [64]

2.4 REQUERIMIENTOS

2.4.1 REQUERIMIENTOS FUNCIONALES

Carga de Archivos Código	
Código	**Descripción**
RF-1	Permitir subir archivos en formato .cvs para procesar y analizar los datos.
RF-2	Validar que sea archivo CSV antes de procesar
RF-3	Mensaje de confirmación tras la validación y carga del archivo

Tabla 1:Requerimiento funcional carga

Implementación Algoritmos ML	
Código	**Descripción**
RF-4	El sistema debe ser capaz de detectar y eliminar automáticamente las columnas que no aportan un valor importante al análisis.
RF-5	Se creará nuevas características para el proceso de los algoritmos

Tabla 2:Requerimiento funcional implementación

Selección Algoritmos ML	
Código	**Descripción**
RF-6	El sistema debe proporcionar al usuario dos algoritmos para procesar datos: SVM y Random Forest.
RF-7	La interfaz del sistema debe permitir que el usuario seleccione solo un algoritmo por sesión de análisis, evitando la ejecución simultánea de varios modelos. Esta restricción asegura un uso eficiente de los recursos y resultados coherentes.

Tabla 3:Requerimiento funcional algoritmo

Análisis de Datos	
Código	**Descripción**
RF-8	El sistema debe realizar todo el proceso de análisis usando únicamente el algoritmo que el usuario seleccionó previamente. Este método mejora el flujo de trabajo al ajustar el procesamiento a las particularidades del modelo.

RF-9	Como parte fundamental del análisis, el sistema debe dividir todo el tráfico de red en dos categorías distintas: normal y anómalo. Esta separación se realiza según los criterios técnicos específicos definidos por el algoritmo seleccionado para identificar posibles patrones.
RF-10	Una vez finalizado el análisis, el sistema debe calcular y generar automáticamente un informe con diferentes métricas de rendimiento. Estas métricas permiten una evaluación objetiva de la efectividad y la fiabilidad del modelo empleado.

Tabla 4:Requerimiento funcional análisis

Visualización de Resultados	
Código	**Descripción**
RF-11	El sistema debe contar con un panel de estadísticas completas que resuma de forma clara los resultados del análisis. Este panel debe mostrar métricas importantes como el total de registros procesados, las tasas de detección de anomalías y métricas concretas de rendimiento del algoritmo.

RF-12	El sistema debe mostrar un gráfico de barras especializado que represente de manera visual los datos de tráfico normal y anómalo dentro puertos y protocolos.

Tabla 5:Requerimiento funcional visualización

2.4.2 REQUERIMIENTOS TECNICOS

Código	Tipo	Descripción
RT-1	Procesador	Intel Core i7 de octava generación o Ryzen 7 3700H.
RT-2	Memoria RAM	16 GB de memoria RAM
RT-3	Disco Duro	500 GB
RT-4	Tarjeta de video	Envidia GeForce GTX 1650
RT-5	Tarjeta de red	802.11 b/g/n/ac 2.4 GHZ - 5 GHZ

Tabla 6:Requerimiento técnicos

2.5 COMPONENTE DE LA PROPUESTA

En esta sección se describe el proyecto de investigación, el cual se centra en el desarrollo de un sistema para la detección de Amenazas Persistentes Avanzadas (APTs) en el tráfico web mediante técnicas avanzadas de minería de datos y análisis de comportamiento. El desarrollo sigue una metodología que integra los procesos de CRISP-DM (Proceso Estándar para Minería de Datos)

CRISP-DM (Proceso Estándar para Minería de Datos entre Industrias)

- Fase de Comprensión del Problema
- Fase de Análisis de los Datos
- Fase de Preparación de los Datos
- Fase de Modelado
- Fase de Evaluación
- Fase de Explotación

2.5.1 FASE DE COMPRENSIÓN DEL PROBLEMA

Se establece el objetivo principal de identificar las Amenazas Persistentes Avanzadas (APTs) dentro de la infraestructura de red de la UPSE. Se identifican los retos específicos, como la dificultad de detectar APTs debido a su capacidad para eludir las medidas de seguridad tradicionales. Los objetivos particulares incluyen la mejora de la detección de APTs, la reducción de falsos positivos y el incremento de la precisión en la identificación de patrones anómalos.

Además, se plantea el fortalecimiento de los mecanismos de monitoreo continuo para anticipar comportamientos sospechosos antes de que comprometan la red. También se busca optimizar el uso de herramientas de análisis avanzado que permitan correlacionar eventos en tiempo real. Finalmente, se prioriza el desarrollo de estrategias proactivas que permitan mitigar rápidamente cualquier indicio de intrusión sofisticada.

Análisis y Comprensión de las Amenazas Persistentes Avanzadas (APTs)

La implementación de este sistema de detección se fundamenta en el análisis de seis métricas clave derivadas de logs de red que después se utilizara, diseñadas específicamente para identificar patrones de comportamiento asociados a Amenazas Persistentes Avanzadas (APTs). Cada métrica se encuentra técnicamente referenciada con el framework MITRE ATT&CK, proporcionando una base sólida para la identificación de tácticas de ataque documentadas en campañas reales.

Detección de Beaconing (Comando y Control)

T1071.001 - Application Layer Protocol: Web Protocols

La técnica **T1071.001** representa una de las amenazas más insidiosas y extendidas en el panorama actual de ciberseguridad. Lo que la hace extraordinariamente peligrosa es una paradoja fundamental: **utiliza los protocolos más comunes y necesarios para el negocio como vector de ataque**. Mientras las organizaciones se esfuerzan por mantener sus operaciones digitales, los atacantes aprovechan estos mismos canales legítimos para mantener presencia persistente en sus redes [65].

Matriz de Detección de Comando y Control mediante Protocolos Web (T1071.001)

Protocolo	Detección Clave	Mitigación Esencial	Impacto y Riesgo	Descripción Táctica
HTTP/HTTPS	Tráfico HTTP/S desde procesos no-navegador, User-Agents anómalos, patrones de beaconing.	Filtrado de tráfico web saliente, restricción de tráfico outbound HTTP no esencial.	Alto - Evasión de detección.	Uso de protocolos web estándar para mezclar el tráfico de C2 con el tráfico legítimo y evadir controles perimetrales.
WebSockets	Conexiones WebSocket anómalas, upgrade requests HTTP sospechosos.	Inspección profunda de protocolos, bloqueo de WebSocket no autorizados (uso de WAF/IPS).	Medio-Alto - Baja detección.	Protocolo WebSocket para C2 persistente y comunicación bidireccional en tiempo real, difícil de detectar si está cifrado.
HTTP APIs	Llamadas API REST anómalas, patrones de consumo irregular de APIs de servicios	Restricción de acceso a APIs externas, monitorización de uso (CASB).	Medio - Enmascaramiento.	Uso de APIs web públicas (Dropbox, Microsoft Graph, etc.) para camuflar el tráfico C2 como comunicación

	web conocidos.			legítima con la nube.
HTTP Encoded	Datos codificados en headers HTTP, parámetros URL o cookies de forma inusual o con alta entropía.	Inspección de contenido HTTP, análisis de comportamiento del tráfico.	Alto - Evasión DLP.	Codificación de comandos y datos en campos de protocolo HTTP (no en el cuerpo principal) para evadir la inspección tradicional.
Mixto HTTP/DNS	Combinación de tráfico HTTP con DNS para resiliencia; actividad multi-protocolo simultánea.	Defensa en profundidad multi-capas (NGFW + DNS Filtering) y correlación de eventos.	Alto - Persistencia.	Uso combinado de múltiples protocolos web y DNS para asegurar la persistencia y el éxito de las comunicaciones C2.

Tabla 7:Análisis de técnicas T1071.001: MITRE ATT&CK (2024)

Detección de Exfiltración por Volumen

T1041 - Exfiltration Over C2 Channel

La técnica **T1041** representa una de las amenazas más efectivas y difíciles de detectar en el panorama actual de exfiltración de datos. Su peligro radica en una premisa fundamental: **no crea nuevo tráfico, sino que aprovecha lo que ya existe**. Mientras las organizaciones monitorean conexiones sospechosas y transferencias

anómalas, los atacantes utilizan los mismos canales establecidos para comando y control para sacar los datos [66].

Matriz de Exfiltración sobre Canal de C2 (T1041)

Protocolo	Detección Clave	Mitigación Esencial	Impacto y Riesgo	Descripción Táctica
HTTP/HTTPS	Volumen de salida inusual proveniente de procesos no relacionados con navegadores, con patrones de transferencia masiva.	Prevención de pérdida de datos (DLP), intrusiones (IPS) y proxies web.	Crítico - Alta evasión.	Reaprovechamiento del canal de C2 ya presente para exfiltrar datos, combinando tráfico robado con tráfico web legítimo.
DNS	Consultas DNS masivas con datos codificados y alta entropía en los subdominios.	Filtrado DNS riguroso y bloqueo de dominios maliciosos.	Alto - Baja detección.	Codificación de datos robados mediante consultas DNS, aprovechando que el tráfico DNS saliente rara vez se inspecciona.
SMTP/Email	Emails automáticos con archivos adjuntos grandes, tráfico saliente desde servidores sin correo.	Restricción estricta en el SMTP saliente y gateways de seguridad de correo electrónico.	Medio - Límites de tamaño.	Uso del canal de correo electrónico como vector para enviar datos como adjuntos, camuflados en el tráfico.

Cloud APIs	Acceso a APIs en la nube desde sistemas no autorizados, transferencias anómalas a servicios de.	Control de acceso a las APIs, CASB (Broker de Seguridad de Acceso a la Nube), y supervisión de servicios en la nube.	Alto - Escalabilidad.	Abuso de APIs legítimas de servicios en la nube (como Dropbox, OneDrive, etc.) para una exfiltración masiva.
Protocolos Personalizados	Tráfico en puertos no estándar, patrones de *beaconing* inconsistentes, protocolos *custom* no identificados.	Segmentación de red estricta, Detección y Respuesta de Red (NDR), Filtrado de puertos.	Alto - Evasión avanzada.	Desarrollo de protocolos privados que combinan C2 y exfiltración en un único canal cifrado.

Tabla 8:Análisis de técnicas T1041: MITRE ATT&CK (2024).

Detección de Escaneo de Puertos

T1046 - Network Service Scanning

La técnica **T1046 de Descubrimiento de Servicios de Red** representa la fase fundamental que separa a los atacantes exitosos de los que fracasan. Mientras las organizaciones se concentran en prevenir la ejecución de malware o la exfiltración de datos, los adversarios inteligentes están realizando el trabajo de reconocimiento que hace posibles todos los demás ataques. Esta técnica permite a los atacantes identificar puertos abiertos, servicios vulnerables y configuraciones débiles que pueden ser explotadas posteriormente.

Además, el descubrimiento de servicios proporciona a los intrusos una visión detallada del entorno objetivo, permitiéndoles seleccionar las rutas de ataque más efectivas y minimizar las posibilidades de ser detectados. Es una etapa silenciosa

pero crítica, en la que pequeñas brechas de información pueden significar grandes compromisos de seguridad. Por ello, comprender y mitigar la técnica T1046 es esencial para fortalecer la postura defensiva, anticiparse a los movimientos del adversario y prevenir que un simple reconocimiento se convierta en una intrusión completa [67].

Matriz de Descubrimiento de Servicios de Red (T1046)

Protocolo	**Detección Clave**	**Mitigación Esencial**	**Impacto y Riesgo**	**Descripción Táctica**
SMB (445/TCP)	Escaneos secuenciales del puerto 445; múltiples conexiones SMB desde el mismo host en poco tiempo.	Desactivar SMBv1, la segmentación de red y la microsegmentación.	Alto - Movimiento lateral.	Escaneo del puerto 445 para identificar sistemas Windows, recursos compartidos accesibles y servicios vulnerables.
RDP (3389/TCP)	Conexiones RDP fallidas múltiples escaneo vertical rápido del puerto 3389.	Prevención de intrusiones (IPS), control estricto del acceso RDP externo y uso de Bastion Hosts.	Alto - Acceso remoto.	Detección de sistemas con Escritorio Remoto habilitado como objetivo para el movimiento lateral.

SSH (22/TCP)	Escaneos SSH provenientes de varias direcciones IP indicios de ataques de fuerza bruta.	Segmentar la red y aplicar autenticación mediante clave, eliminando el uso de contraseñas.	Medio-Alto - Acceso a entornos Linux.	Identificación de servidores SSH como punto de entrada para ataques de fuerza bruta o para la explotación de vulnerabilidad es en entornos Linux/Unix.
HTTP/HTTP S (80/443)	Escaneos de puertos web (sondeos) o solicitudes inusuales a servicios web y APIs.	Detección de escaneo web (anomalías de volumen), WAF (Firewall de Aplicaciones Web) para protección..	Medio - Reconocimien to de aplicaciones.	Detección de servicios web, aplicaciones y APIs en la red para identificar posibles puntos de interés para explotación.
mDNS/Bonjo ur (5353/UDP)	Consultas mDNS en masa y descubrimien to de servicios Bonjour en la red local.	Deshabilitar el protocolo Bonjour/mDNS cuando no sea estrictamente necesario.	Medio - Entorno Apple.	Uso del protocolo mDNS para descubrir automáticamen te servicios y hosts en redes Apple (macOS, iOS).

Tabla 9:Análisis de técnicas T1046: MITRE ATT&CK (2024).

Actividad en Horario No Laboral

T1021.001 - Remote Desktop Protocol

La técnica T1021.001 representa una de las amenazas más devastadoras en el panorama actual de ciberseguridad. Lo que la hace excepcionalmente peligrosa es su doble naturaleza: es a la vez una herramienta legítima de administración y un vector de ataque de máxima eficacia. Mientras las organizaciones dependen del Escritorio Remoto para sus operaciones diarias, los atacantes aprovechan esta dependencia para moverse silenciosamente a través de las redes [68].

Matriz de Movimiento Lateral via RDP (T1021.001*)*

Protocolo/Puerto	**Detección Clave**	**Mitigación Esencial**	**Impacto y Riesgo**	**Descripción Táctica**
RDP (3389/TCP)	Login RDP seguido de ejecución de procesos inusuales o acceso a archivos sensibles.	MFA, Segmentación de red y RDP Gateways.	Crítico - Acceso interactivo y control.	Utilización de credenciales válidas para movimientos laterales mediante sesiones interactivas completas.
RDP Tunneling (No Estándar)	Conexiones RDP en puertos elevados o encapsulados tráfico en puertos no convencionales.	Gateways RDP, análisis de comportamiento y detección de anomalías en tráfico de red.	Alto - Evasión de detección.	Encapsulamiento del tráfico RDP usando puertos no convencionales para evadir firewalls perimetrales.
RDP sobre Web Shells	Tráfico RDP inusual pasando a través de servicios web, patrones HTTP anómalos.	WAF, Deshabilitar RDP innecesario y Honeypots.	Alto - Bypass de perímetro y persistencia.	Uso de una puerta trasera web (web shell) como proxy para establecer una sesión RDP.

RDP Herramientas LOTL	Ejecución de mstsc.exe desde sistemas que no son administrativos ; conexiones múltiples.	EDR, UEBA, Auditoría de usuarios RDP.	Medio-Alto - Baja visibilidad (uso de binarios legítimos).	Abuso de la herramienta nativa mstsc.exe (Living Off The Land) para movimiento lateral sin malware.
RDP para Persistencia	Habilitación de RDP o cambios de configuración en sistemas donde estaba deshabilitado.	GPOs (Políticas de Grupo), Controles de configuración y timeout de sesiones.	Alto - Reacceso garantizado y control persistente.	Modificación de configuraciones del sistema operativo para garantizar el acceso continuo.

Tabla 10: Análisis de técnicas T1021.001: MITRE ATT&CK (2024).

Detección de Exfiltración por Asimetría

T1048.003 - Exfiltration Over Unencrypted Non-C2 Protocol

La técnica T1048.003 representa una de las amenazas más sigilosas y efectivas en el panorama actual de ciberseguridad. Lo que la hace particularmente peligrosa es su simplicidad operacional: los atacantes no necesitan desarrollar complejos mecanismos de cifrado o evasión, sino que aprovechan protocolos legítimos que existen en prácticamente todas las redes empresariales [69].

Matriz de detección *y mitigación de exfiltración mediante protocolos alternativos (T1048.003)*

Protocolo	**Detección Clave**	**Mitigación Esencial**	**Impacto y Riesgo**	**Descripción Táctica**
DNS	Consultas masivas, alta entropía en subdominios.	Filtrado DNS, bloqueo de dominios maliciosos (por ejemplo,	Alto - Evasión de DLP.	Datos codificados y fragmentados en subdominios

		Cisco Umbrella).		DNS para evadir controles.
FTP	Transferencias desde estaciones de trabajo en horarios anómalos.	Bloqueo de tráfico FTP saliente (puerto 21), segmentación de red.	Alto - Transferencia de gran volumen de datos, independiente del canal C2.	Utilización de un canal FTP para transferir grandes volúmenes de datos.
HTTP/HTTPS	Ratio upload/download desequilibrado (>10:1), son HTTP POST grandes.	DLP (Data Loss Prevention), inspección SSL/TLS (si es posible), proxies web avanzados.	Medio-Alto - Encriptado.	Datos ofuscados en el tráfico web normal el uso de HTTPS requiere una inspección profunda.
SMTP	Correos electrónicos automáticos, archivos adjuntos grandes y utilización de servidores no-mail.	Restricción en SMTP y gateways de correo electrónico (Proofpoint, Mimecast).	Medio - Límite por el tamaño del email.	Información enviada mediante adjuntos o notificaciones por correo electrónico.
WebDAV	Conexiones continuas a servicios de nube personales y autenticación externa.	Bloqueo del acceso externo a WebDAV mediante el uso de CASB (Cloud Access Security Broker).	Medio - Acceso remoto.	Uso de servicios de nube personales (p. ej., OneDrive/Google Drive) para almacenamiento

				temporal de datos.
TFTP	Transferencias desde equipos de infraestructura de red.	Bloqueo de protocolo TFTP, endurecimiento (hardening) de dispositivos de red.	Crítico - Infraestructura.	Exfiltración de archivos de configuración y tablas de enrutamiento críticos.
Múltiples	Actividad multi-protocolo simultánea, correlación de patrones.	Defensa en profundidad, sistemas de correlación de eventos (SIEM/EDR).	Alto - Persistencia.	Uso de varios protocolos combinados para asegurar la exfiltración exitosa y redundante.

Tabla 11:Análisis de técnicas T1048.003: MITRE ATT&CK (2024).

Sesiones de Larga Duración (C2 Persistente)

T1071 - Application Layer Protocol

La técnica T1071 representa uno de los cambios más significativos en la ciberseguridad de la última década: la migración masiva de protocolos de C2 customizados hacia protocolos de aplicación legítimos. Lo que comenzó como una técnica de evasión se ha convertido en la metodología estándar para el comando y control moderno [70].

Matriz de Comando y Control sobre Protocolos de Capa de Aplicación (T1071)

Protocolo	**Detección Clave**	**Mitigación Esencial**	**Impacto y Riesgo**	**Descripción Táctica**
HTTP/HTTPS (Web)	Tráfico HTTP/S desde procesos no-	Filtrado de tráfico web, Proxies avanzados y DLP	Alto - Evasión de detección.	Abuso de protocolos web estándar (puertos

	navegador, alto volumen saliente anómalo.	(Prevención de Pérdida de Datos).		80/443) para camuflar el tráfico C2 dentro de flujos legítimos.
DNS	Consultas DNS inusuales y alta entropía en subdominios, lo que indica posible codificación de datos.	Filtrado DNS riguroso con bloqueo de dominios maliciosos conocidos (como Cisco Umbrella, etc.).	Medio-Alto - Baja detección.	Utilización de consultas DNS para crear canales C2 (tunneling), aprovechando la ausencia de inspección profunda.
SMB	Tráfico SMB entre segmentos de red no relacionados detectar patrones anómalos de conexión o uso de shares.	Microsegmentación y restricción del tráfico SMB entre estaciones de trabajo.	Alto - Movimiento lateral.	Uso del protocolo SMB para comunicaciones C2 y desplazamiento en la red interna, fusionándose con tráfico de archivos legítimo.
SSH	Conexiones SSH inversas, creación de túneles persistentes y	Restricción rigurosa del SSH saliente, utilización de Bastion Hosts y	Alto - Evasión perimetral.	Configuración de túneles SSH, incluyendo conexiones inversas, para

	tráfico saliente inusual.	supervisión de la red.		mantener el C2 persistente y evadir mecanismos de control firewall.
IRC	Conexiones IRC desde sistemas empresariales, uso de puertos no estándar.	Detección de protocolos no empresariales, Application Control (NGFW, IPS).	Medio - Protocolo legado.	Uso de protocolo IRC para comunicaciones C2; aunque antiguo, su simplicidad y baja frecuencia de uso lo hacen pasar desapercibido.
Protocolos Personalizados	Tráfico en puertos no estándar, patrones de beaconing inconsistentes, protocolos custom no identificados.	Firmas personalizadas, Análisis de Comportamiento (Behavioral Analysis), Detección de Anomalías de Red (NDR).	Alto - Detección difícil.	Desarrollo de protocolos privados sobre TCP/UDP para evadir la detección basada en firmas y patrones conocidos.

Tabla 12:Análisis de técnicas T1071: MITRE ATT&CK (2024).

Análisis Comparativo de Algoritmos de Machine Learning en la Detección de APTs

La selección de algoritmos para este estudio se basó en una revisión exhaustiva de artículos y tesis científicos sobre detección de anomalías en seguridad informática.

Se optó por una combinación estratégica que incluye tanto enfoques supervisados como no supervisados, permitiendo una comparación comprehensiva de diferentes paradigmas de aprendizaje automático.

Support Vector Machine (SVM)

El Support Vector Machine (SVM) es un algoritmo de clasificación supervisada que se utiliza comúnmente en problemas de detección de intrusiones, incluidos los ataques de APT. SVM es eficaz debido a su capacidad para crear un hiperplano en un espacio multidimensional que separa los datos de una manera que maximiza el margen entre clases [71]. La investigación compara el desempeño de los modelos SVM con otros enfoques tradicionales en la detección de APTs, mostrando que SVM es una herramienta poderosa para detectar patrones anómalos en grandes conjuntos de datos de tráfico de red, lo que es esencial para la identificación de actividades maliciosas a largo plazo [72].

K-Means

El K-Means es un algoritmo de agrupamiento no supervisado que se agrupa los datos en "k" clústeres basados en características comunes de los datos de entrada [73]. K-Means ayuda a detectar patrones inusuales relacionados con APTs mediante la agrupación de datos con comportamientos similares. El artículo sugiere que, mediante la aplicación de K-Means, es posible detectar ataques avanzados que presentan patrones complejos en el tráfico de red, facilitando la identificación temprana de amenazas [74]. Además, existe EFMS-KMeans que se utiliza para segmentar flujos normales y anómalos mediante la estimación de centros densos y el cálculo de distancias a los centroides, permitiendo una preetiquetación automática y robusta que permite entrenamientos posteriores con otros algoritmos mostrados en el estudio [75].

Isolation Forest

El algoritmo Isolation Forest detecta anomalías de forma no supervisada mediante árboles de aislamiento que separan observaciones aleatoriamente. Las anomalías se identifican por requerir menos divisiones para ser aisladas, reflejando su naturaleza atípica en la estructura del bosque. La profundidad promedio de aislamiento en los

árboles funciona como medida de anomalía [76]. Este algoritmo es particularmente útil para detectar comportamientos atípicos en grandes volúmenes de datos, como los que se generan durante los ataques de APT. El artículo demuestra cómo el aislamiento de patrones anómalos facilita la identificación de amenazas ocultas dentro del tráfico de red, y cómo la eficiencia de este método permite su implementación en sistemas de detección en tiempo real [77].

Random Forest

El Bosque Aleatorio es un algoritmo de aprendizaje supervisado que combina múltiples árboles de decisión mediante consenso o promediado [78]. Cada árbol se entrena con subconjuntos aleatorios de datos y características, proporcionando robustez y evitando sobreajuste, Random Forest ofrece robustez frente a sobreajustes y es efectivo para manejar conjuntos de datos con valores faltantes o ruidosos. [78]. Se aplica el algoritmo Random Forest en el contexto de la detección de fraudes, lo cual es análogo a la detección de APTs, el artículo muestra que este enfoque tiene la capacidad de manejar grandes volúmenes de datos con múltiples características y detectar patrones complejos, lo que lo hace adecuado para la identificación de amenazas cibernéticas persistentes y avanzadas [79].

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) es un algoritmo eficiente para identificar agrupamientos no esféricos y detectar valores atípicos pero su alto costo computacional limita su aplicación en conjuntos de datos grandes, ya que requiere calcular distancias entre todos los puntos, lo que hace que DBSCAN sea muy lento a medida que aumenta la cantidad de puntos, complicando su uso en bases de datos extensas masivas. Aunque se implementen optimizaciones, como estructuras de datos eficientes la característica intrínseca del algoritmo sigue siendo un desafío al manejar grandes cantidades de datos. Este alto costo computacional puede hacer que DBSCAN no sea adecuado con grandes conjuntos de datos [80].

Además, su rendimiento puede verse afectado cuando la distribución de los datos es muy densa o irregular, aumentando aún más el tiempo necesario para el procesamiento. En contextos donde la escalabilidad es crítica, suelen preferirse

métodos alternativos o versiones modificadas que reduzcan el impacto computacional.

Parámetro	K-Means	SVM	Isolation Forest	Random Forest
Tipo de Aprendizaje	No supervisado (Clustering)	Supervisado (Clasificación)	No supervisado (Detección de Anomalías)	Supervisado (Ensemble)
Fundamento Teórico	Particionamiento basado en distancias euclidianas	Maximización del margen de separación entre clases	Aislamiento de anomalías mediante árboles de decisión	Combinación de múltiples árboles de decisión
Complejidad Computacional	$O(n \times k \times i \times d)$	$O(n^2)$ a $O(n^3)$	$O(n \times \log n)$	$O(m \times n \times \log n)$
Escalabilidad	Alta para datasets grandes	Media-Baja	Alta	Media-Alta
Manejo de Ruido	Sensible a outliers	Robustez media mediante márgenes	Especializado en detección de outliers	Alta robustez mediante promediado

Interpretabilidad	Media (visualización de clusters)	Baja (kernel trick)	Media (profundidad de aislamiento)	Alta (importancia de características)
Requerimiento de Preprocesamiento	Estandarización obligatoria	Normalización recomendada	Mínimo requerimiento	Manejo directo de mixed data
Parámetros Críticos	Número de clusters (k), Inicialización	Kernel, C, γ	Contaminación, n_estimators	n_estimators, max_depth, min_samples_split
Rendimiento en Datos Desbalanceados	Bajo (sensibilidad a distribuciones)	Medio (con class_weight)	Alto (especializado en minorías)	Alto (balanceo automático)
Métricas de Evaluación	Silhouette Score: Inertia:	Precisión: AUC	Precisión AUC	Precisión AUC
Aplicación Específica en APTs	Detección de grupos de comportamiento anómalo	Clasificación binaria de tráfico malicioso	Identificación de puntos extremos en flujos de red	Clasificación multi-categoría de técnicas APT
Ventajas en Contexto de Seguridad	• No requiere etiquetas previas	• Efectivo en alta dimensionali	• Detección específica de anomalías	• Alta precisión en clasificación

	Descubrimiento de patrones ocultos Agrupamiento por similitud	dad Generalización robusta Múltiples kernels disponibles	Eficiencia computacional Menor requerimiento de tuning	Resistencia a overfitting Ranking de características importantes
Limitaciones Identificadas	Sensibilidad a inicialización Asunción de clusters esféricos Dificultad en determinar k óptimo	Costo computacional elevado Dificultad interpretativa Sensibilidad a parámetros	Bajo rendimiento en datos complejos AUC subóptimo (0.74) Falsos positivos elevados	Consumo de memoria Tiempo de entrenamiento Complejidad de implementación

Tabla 13: Análisis Comparativo de Algoritmos de Minería de Datos para Detección de APTs

2.5.2 FASE DE ANÁLISIS DE LOS DATOS

2.5.2.1 Fuente de los Datos

El dataset en análisis proviene de un dispositivo de seguridad Fortigate 2500E, un firewall de próxima generación enterprise-grade desplegado en un entorno corporativo real. Este equipo forma parte de la infraestructura crítica de red y ha sido configurado para capturar tráfico de red exhaustivo como parte del proyecto de investigación titulado "Clasificación de tráfico web orientado a la identificación oportuna de ataques utilizando técnicas de Deep Learning" [81]. El dataset contiene 394,771 registros exhaustivamente preprocesados, representando una muestra significativa del tráfico de red corporativo **Ilustración 8.**

Date	Time	Eventtime	Tz	Logid	Type	Level	Srcip	Srcname	Srcintf	Dstip	Dstintf	Dstintfrole	Sessionid
2025-02-26	15:35:24	1740602125038897	-500	13	traffic	0	172.29.0.9	13686944-6497-	Archivos	186.5.11.4	LACP_WAN_1G_	wan	256955200
2025-02-26	15:35:24	1740602125038886	-500	13	traffic	0	172.17.64.128		ESTUDIANTES	163.70.152.1	LACP_WAN_1G_	wan	256994172
2025-02-26	15:35:24	1740602125038884	-500	13	traffic	0	172.23.0.29	DESKTOP-UEI5OS	VLAN126	186.5.11.4	LACP_WAN_1G_	wan	256955210
2025-02-26	15:35:24	1740602125028892	-500	13	traffic	0	172.17.10.177		VLAN_150	186.5.11.4	LACP_WAN_1G_	wan	256955185
2025-02-26	15:35:24	1740602125018897	-500	13	traffic	0	172.17.68.180	DESKTOP-G9TIQ3	ESTUDIANTES	142.250.189.13	LACP_WAN_1G_	wan	256955009
2025-02-26	15:35:24	1740602125018890	-500	13	traffic	0	172.17.10.231	Hornos	VLAN_150	23.227.60.200	LACP_WAN_1G_	wan	256955175
2025-02-26	15:35:24	1740602125018886	-500	13	traffic	0	172.17.10.231	Hornos	VLAN_150	186.5.11.4	LACP_WAN_1G_	wan	256955170
2025-02-26	15:35:24	1740602125008890	-500	13	traffic	0	172.17.66.246		ESTUDIANTES	163.70.152.1	LACP_WAN_1G_	wan	256990016
2025-02-26	15:35:24	1740602124998888	-500	13	traffic	0	172.17.65.115		ESTUDIANTES	163.70.152.1	LACP_WAN_1G_	wan	256952088
2025-02-26	15:35:24	1740602124998886	-500	13	traffic	0	172.29.0.9	13686944-6497-	Archivos	186.5.11.4	LACP_WAN_1G_	wan	256955145
2025-02-26	15:35:24	1740602124988897	-500	13	traffic	0	172.17.10.194	LAPTOP-523F1AE	VLAN_150	20.69.137.228	LACP_WAN_1G_	wan	256981528
2025-02-26	15:35:24	1740602124988889	-500	13	traffic	0	172.20.0.16	DESKTOP-QC090	VLAN_123	186.5.11.4	LACP_WAN_1G_	wan	256955150
2025-02-26	15:35:24	1740602124988888	-500	13	traffic	0	172.17.4.69	LAPTOP-F1KOKKL	ADMINISTRATIVO	186.5.56.4	LACP_WAN_1G_	wan	257032979
2025-02-26	15:35:24	1740602124988886	-500	13	traffic	0	172.17.64.229	Android-3	ESTUDIANTES	163.70.152.1	LACP_WAN_1G_	wan	256955155
2025-02-26	15:35:24	1740602124968885	-500	13	traffic	0	172.17.10.177		VLAN_150	186.5.11.4	LACP_WAN_1G_	wan	256955119
2025-02-26	15:35:24	1740602124968881	-500	13	traffic	0	172.17.67.0		ESTUDIANTES	31.13.67.63	LACP_WAN_1G_	wan	256955074
2025-02-26	15:35:24	1740602124888893	-500	13	traffic	0	172.25.0.5	DESKTOP-Q7TPS(	VLAN_128	186.5.11.4	LACP_WAN_1G_	wan	256955084
2025-02-26	15:35:24	1740602124888892	-500	13	traffic	0	172.17.10.231	Hornos	VLAN_150	142.250.218.13	LACP_WAN_1G_	wan	256955014
2025-02-26	15:35:24	1740602124888890	-500	13	traffic	0	172.17.64.21	Android-7	ESTUDIANTES	163.70.152.33	LACP_WAN_1G_	wan	257032000
2025-02-26	15:35:24	1740602124873753	-500	20	traffic	0	172.17.10.194	LAPTOP-523F1AE	VLAN_150	13.107.22.239	LACP_WAN_1G_	wan	256924600
2025-02-26	15:35:24	1740602124808885	-500	13	traffic	0	172.17.65.126	9d08c229-39ee-4	ESTUDIANTES	186.5.11.4	LACP_WAN_1G_	wan	256955049
2025-02-26	15:35:24	1740602124748886	-500	13	traffic	0	192.168.14.91	DESKTOP-58HAKI	VLAN_30	172.179.183.12	LACP_WAN_1G_	wan	257030052
2025-02-26	15:35:24	1740602124748884	-500	13	traffic	0	172.17.65.203	Android-16	ESTUDIANTES	186.5.56.4	LACP_WAN_1G_	wan	256954803
2025-02-26	15:35:24	1740602124718889	-500	13	traffic	0	172.17.68.158		ESTUDIANTES	163.70.152.1	LACP_WAN_1G_	wan	256997610
2025-02-26	15:35:24	1740602124706698	-500	20	traffic	0	172.17.5.114	COORDINACI?N	ADMINISTRATIVO	74.125.134.188	LACP_WAN_1G_	wan	254829642
2025-02-26	15:35:24	1740602124698892	-500	13	traffic	0	172.17.67.39	Redmi-10-2022	ESTUDIANTES	142.251.135.17	LACP_WAN_1G_	wan	256994624

Ilustración 8:Dataset original

2.5.2.2 Descripción de los Datos

El dataset abarca 44 campos meticulosamente registrados que proporcionan una visión multidimensional del tráfico de red, desde aspectos básicos de conectividad hasta métricas avanzadas de seguridad.

La integridad general de los datos es notable, con 27 campos manteniendo una completitud del 100% mostrado en la Tabla 14, lo que asegura una base sólida para análisis estadísticos y modelos predictivos. Los campos clave para detectar amenazas, como direcciones IP de origen y destino, puertos, duración de sesiones y volúmenes de datos, están completos, lo que permite realizar reconstrucciones precisas de las actividades en la red y de los patrones de comportamiento de la comunicación.

El conjunto de datos abarca detalles básicos de la capa de red, como protocolos, servicios y métricas de rendimiento, incluye información de seguridad avanzada, clasificando aplicaciones, evaluando riesgos y estableciendo acciones de políticas. Esta combinación facilita tanto el análisis técnico profundo como la evaluación del comportamiento a nivel general.

La presencia de campos de tiempo incluyendo timestamps en nanosegundos, facilita el análisis de patrones temporales complejos como beaconing y actividades coordinadas. Al mismo tiempo, las etiquetas de anomalías validadas ofrecen una

referencia de verdad necesaria para entrenar y validar modelos de machine learning en escenarios reales de detección de amenazas.

Aunque ciertos campos especializados, como fabricantes de hardware y servicios específicos de internet, muestran menor completitud, los componentes esenciales para la detección de APTs y el análisis de seguridad se encuentran claramente establecidos. El conjunto de datos en su totalidad es una base de datos excepcionalmente bien organizada y completa ideal para investigación en ciberseguridad, análisis forense digital y desarrollo de sistemas de detección de intrusiones.

Campo	***Tipo Dato***	***No Nulos***	***Total***	***% Completitud***	***Descripción***
Date	object	394,771	394,771	100%	Fecha del evento
Time	object	394,771	394,771	100%	Hora del evento
Eventtime	int64	394,771	394,771	100%	Timestamp en nanosegundos
Tz	int64	394,771	394,771	100%	Zona horaria UTC
Logid	int64	394,771	394,771	100%	ID único de log
Type	object	394,771	394,771	100%	Tipo de evento
Level	int64	394,771	394,771	100%	Nivel de severidad (0-7)
Srcip	object	394,771	394,771	100%	IP origen
Srcname	object	264,471	394,771	67%	Nombre dispositivo origen

Srcintf	object	394,771	394,771	100%	Interfaz origen
Dstip	object	394,771	394,771	100%	IP destino
Dstintf	object	394,771	394,771	100%	Interfaz destino
Dstintfrole	object	394,771	394,771	100%	Rol interfaz destino
Sessionid	int64	394,771	394,771	100%	ID sesión única
Proto	int64	394,771	394,771	100%	Protocolo (6=TCP, 17=UDP)
Policyid	int64	394,771	394,771	100%	ID política firewall
Poluuid	object	393,344	394,771	99.6%	UUID política
Policyname	object	393,344	394,771	99.6%	Nombre política
Service	object	394,771	394,771	100%	Servicio detectado
Transip	object	393,344	394,771	99.6%	IP después de NAT
Transport	float64	393,344	394,771	99.6%	Puerto después de NAT
Appcat	object	394,771	394,771	100%	Categoría aplicación
Duration	int64	394,771	394,771	100%	Duración sesión (segundos)
Sentbyte	int64	394,771	394,771	100%	Bytes enviados
Rcvdbyte	int64	394,771	394,771	100%	Bytes recibidos
Sentpkt	int64	394,771	394,771	100%	Paquetes enviados

Rcvdpkt	int64	394,771	394,771	100%	Paquetes recibidos
Srchwvendor	object	153,130	394,771	39%	Fabricante hardware origen
Osname	object	392,964	394,771	99.5%	SO host origen
Mastersrcmac	object	393,344	394,771	99.6%	MAC principal origen
Srcmac	object	393,344	394,771	99.6%	MAC interfaz origen
Srcport	int64	394,771	394,771	100%	Puerto origen
Dstport	int64	394,771	394,771	100%	Puerto destino
Dstinetsvc	object	127,232	394,771	32%	Servicio internet destino
Appid	float64	267,191	394,771	68%	ID aplicación Fortigate
App	object	267,191	394,771	68%	Nombre aplicación
Apprisk	int64	394,771	394,771	100%	Riesgo aplicación (1-5)
Applist	object	274,650	394,771	70%	Lista aplicaciones
Lanout	float64	131,675	394,771	33%	Métrica tráfico salida LAN
Utmaction	int64	394,771	394,771	100%	Acción UTM (0=deny,1=allow)
Utmref	object	134,413	394,771	34%	Referencia UTM
Authserver	float64	0	394,771	0%	Servidor autenticación
Countwaf	float64	0	394,771	0%	Contador eventos WAF

Anomaly	int64	394,771	394,771	100%	Etiqueta (0=normal,1=anómalo)

Tabla 14: Esquema del Dataset: Elaboración Propia

2.5.2.3 Exploración de los datos

En la fase de exploración de datos, se busca obtener una comprensión profunda de la estructura y las características del dataset, identificando patrones y relaciones clave que puedan ser útiles para el modelado posterior. Las gráficas presentadas ofrecen la distribución de las variables y sus relaciones, facilitando la identificación de características relevantes y anomalías.

En la Ilustración 9 se presenta una matriz de correlación entre las variables del dataset, donde círculos de diferentes tamaños y colores representan la intensidad de las relaciones, las casillas con "?" señalan que las variables tienen valores iguales o vacías Las variables más adecuadas para el modelado son aquellas que muestran una correlación significativa con la variable Anomaly, sin una redundancia excesiva. Los círculos mayores y de colores más intensos, ya sean azules o rojos, indican una correlación fuerte y son las más relevantes para diferenciar entre comportamientos normales y anómalos.

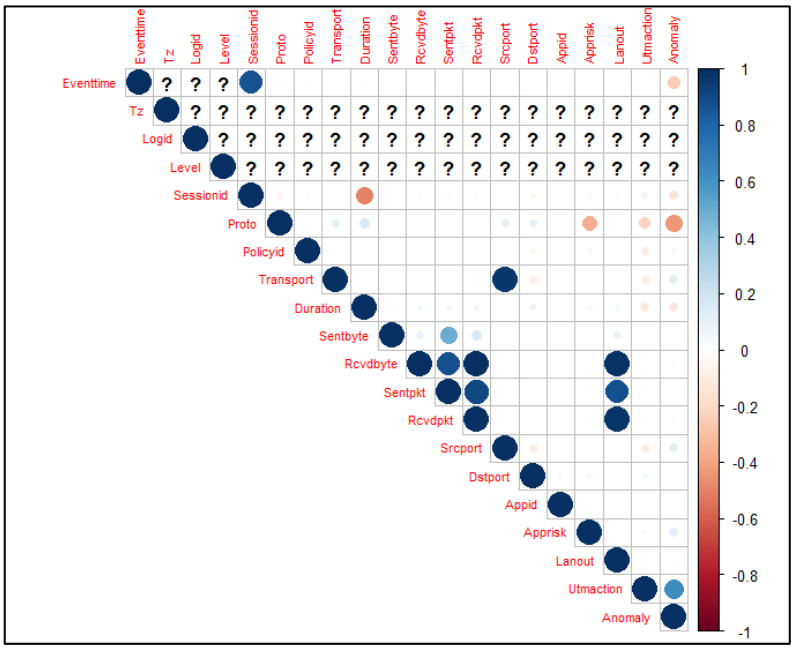


Ilustración 9: Matriz de correlación de características numéricas: Elaboración propia

En la Ilustración 10 se presenta la distribución general de la variable "Anomaly" en el conjunto de datos. Se nota una diferencia significativa entre los registros normales (categoría 0) y los registros anómalos (categoría 1), siendo mucho más comunes los normales, lo que refleja un desequilibrio en las clases.

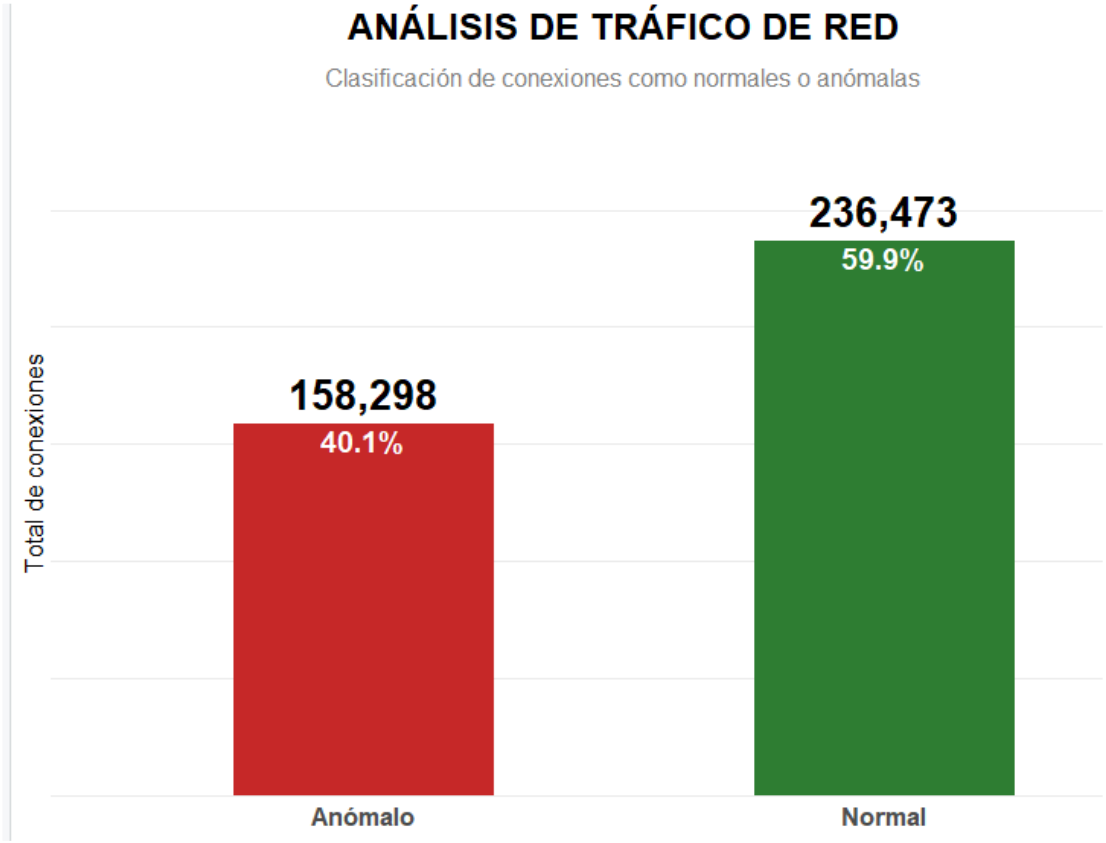


Ilustración 10:Distribución de anomalías: Elaboración propia

En la Ilustración 11 se muestra un histograma de frecuencia que representa las categorías de aplicaciones (Appcat) asociadas a eventos anómalos, clasificados según la etiqueta Anomaly. Se ha notado que las aplicaciones de redes sociales y las no escaneadas muestran una frecuencia mucho mayor de eventos anómalos, lo que indica que están más relacionadas con ciertos comportamientos anómalos. Las aplicaciones como Correo Electrónico, Colaboración y Intereses Generales se utilizan con menor frecuencia en comparación con otras. Este gráfico ayuda a identificar las categorías de aplicaciones que podrían estar vinculadas a ataques o comportamientos anómalos en la red.

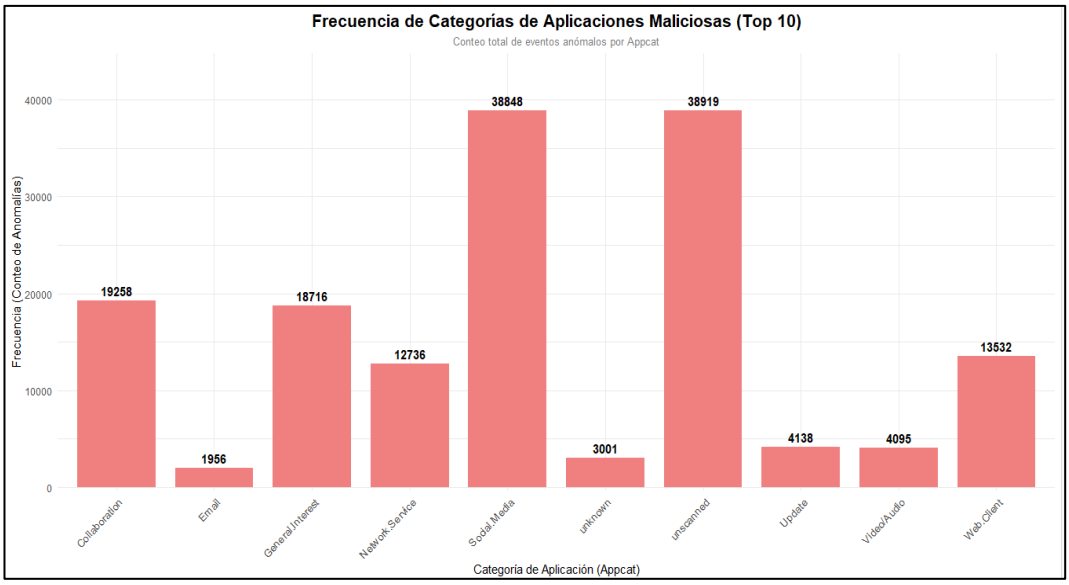


Ilustración 11:Top 10 de los más frecuentes categorías de app por anomalía: Elaboración propia

La Ilustración 12 revela que la principal preocupación es el tráfico NA (no identificado), que representa el 40% de todas las anomalías y constituye un punto ciego importante para la seguridad. Además, la alta concentración en Facebook, HTTPS.BROWSER y TikTok indica que se debe ajustar el modelo para reducir falsos positivos o investigar si hay un uso indebido masivo de estas plataformas canales.

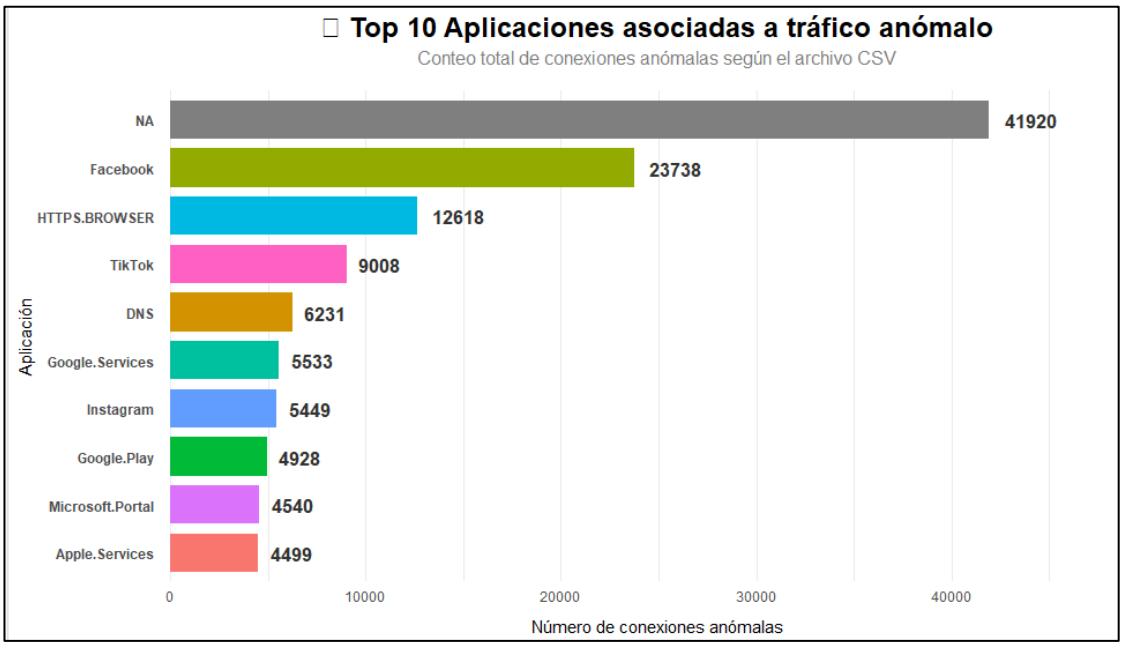


Ilustración 12:Aplicaciones asociadas a tráfico malicioso

La Ilustración 13 presenta un alto riesgo, ya que clasifica la mayor parte del tráfico, 302,673 conexiones, en el Nivel 3 (Elevado y Crítico), la categoría máxima según la escala personalizada. Esto indica que la mayoría de las aplicaciones en uso tienen un riesgo intrínseco muy alto, requiriendo una investigación urgente para identificar las aplicaciones específicas que generan esta amplia superficie de riesgo y aplicar mitigaciones prioritarias. Por otro lado, el tráfico de Riesgo Nivel 1 (Medium y Low) suma 91,123 conexiones, mientras que el Riesgo Nivel 2 (High) es casi inexistente en la clasificación aplicaciones.

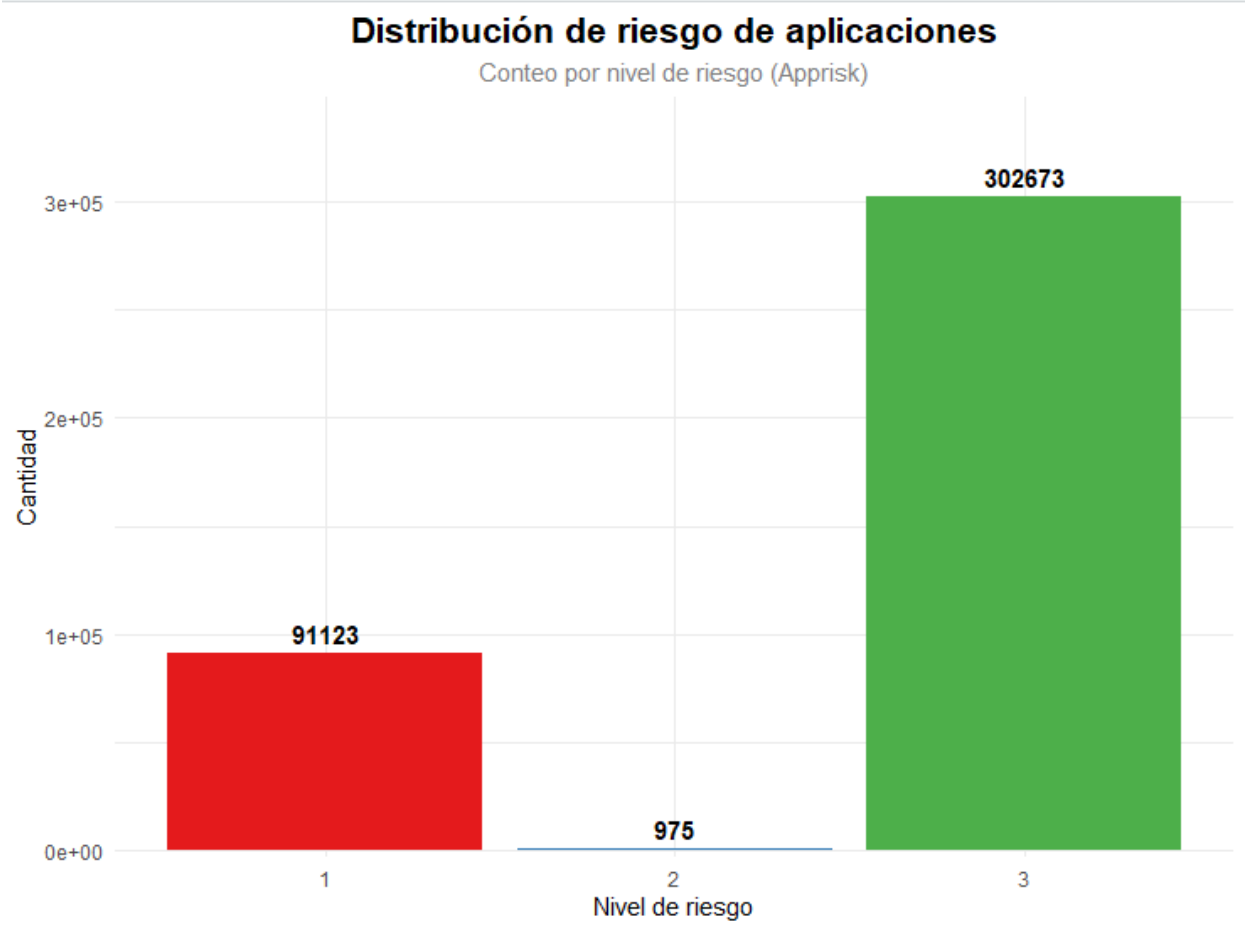


Ilustración 13:Distribución de riesgo de aplicaciones

La Ilustración 14 es importante porque muestra que la mayor amenaza está en el Puerto 443 (HTTPS), donde 131,958 conexiones anómalas superan ampliamente las normales, lo que indica un posible ataque masivo oculto en tráfico cifrado. El puerto 53 (DNS) presenta 11,945 conexiones anómalas, señal de un alto riesgo en

DNS, las anomalías en el Puerto 80 (HTTP) y en puertos no estándar confirman que los atacantes explotan tanto los protocolos web habituales como los protocolos personalizados para evadir la detección.

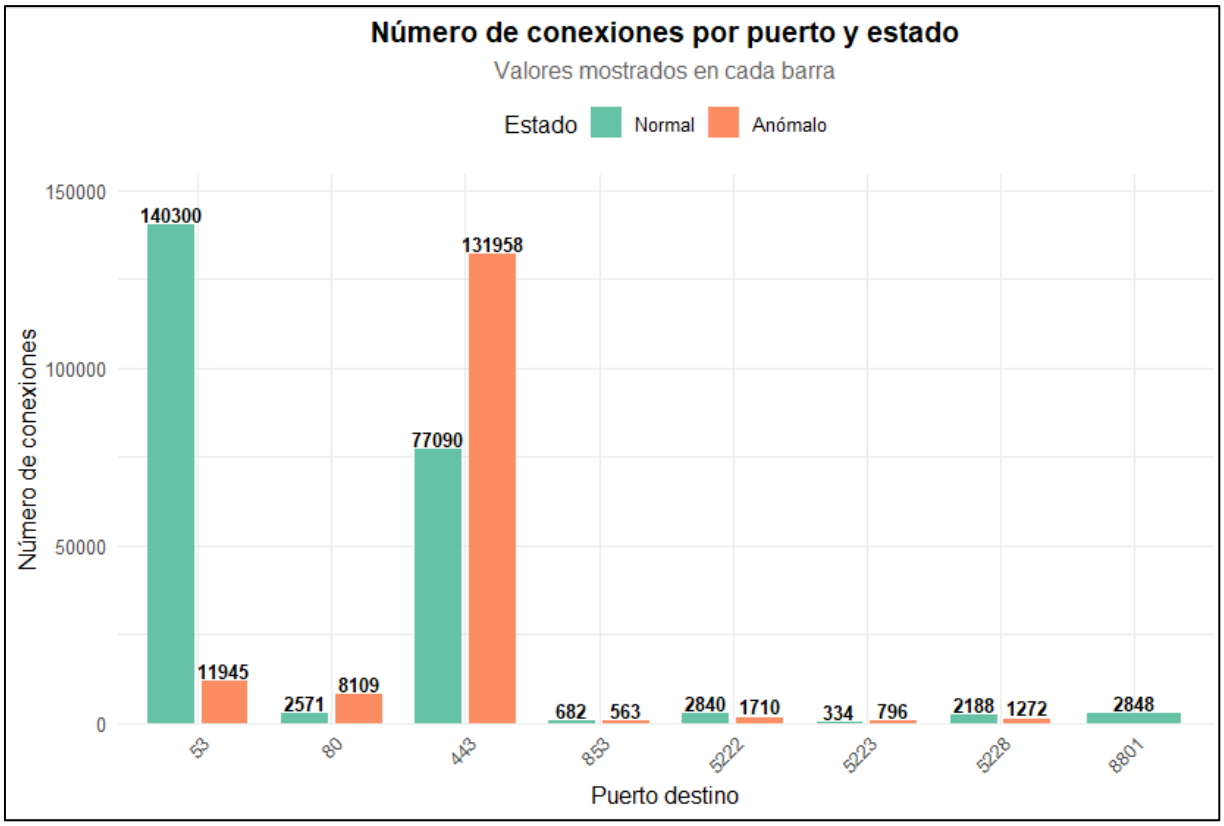


Ilustración 14: Número de conexiones por puerto

La Ilustración 15 muestra la capa de transporte del tráfico y el conteo total de conexiones por protocolo. El TCP alcanza su nivel más alto con 208.979 conexiones, típico en redes que usan tráfico confiable como HTTP/S, SSH y transferencias de archivos en cambio, UDP presenta un volumen considerable con 185.792 conexiones. Este elevado uso de UDP, empleado en protocolos sin conexión y de rápida transmisión como DNS y streaming, indica que gran parte de la actividad irregular previamente detectada en el Puerto 53 (DNS), basado en UDP, es la principal causa de este volumen de tráfico. Además, este comportamiento refuerza la posibilidad de que exista un uso inusual o automatizado de consultas DNS que incrementa significativamente la carga sobre la red. Este patrón podría

estar asociado tanto a fallos de configuración como a actividades maliciosas que buscan aprovechar la naturaleza ligera del protocolo UDP.

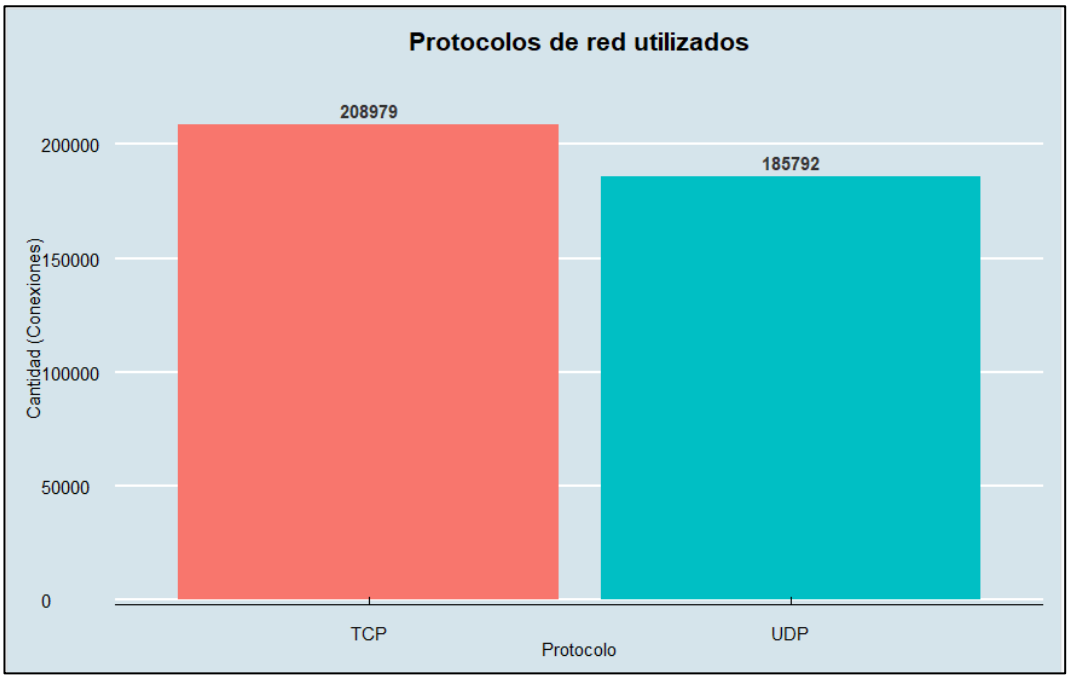


Ilustración 15: Protocolo de red utilizados: Elaboración propia

2.5.3 FASE DE PREPARACIÓN DE LOS DATOS

Tras el análisis exploratorio, se procederá a aplicar ingeniería de características, proceso que incluye la limpieza y preparación de los datos, la extracción y creación de nuevas variables predictoras y la transformación de los datos para el modelado posterior.

```
# Verificar las columnas
print(f"Información de columnas:")
print(df.dtypes.value_counts())
print(f"\nColumnas object: {list(df.select_dtypes(include=['object']).columns)}")

# 🔧 Target Encoding mejorado para variables categóricas
def target_encode_con_verificacion(df, categorical_col, target_col='Anomaly', smooth=10):
    """
    Target Encoding con verificación exhaustiva
    """
    print(f"\n TARGET ENCODING PARA: {categorical_col}")
    print(f"   {'-' * 50}")

    # Verificar datos de entrada
    print(f"Verificación de datos:")
    print(f"      - Columna: {categorical_col}")
    print(f"      - Tipo: {df[categorical_col].dtype}")
    print(f"      - Valores únicos: {df[categorical_col].nunique()}")
    print(f"      - No nulos: {df[categorical_col].count()}/{len(df)}")

    #Calcular estadísticas
    print(f"Calculando cantidad de cada categoria...")
```

```python
# Agrupar por categoría
stats = df.groupby(categorical_col)[target_col].agg(['count', 'mean']).reset_index()

print(f"      - Categorías encontradas: {len(stats)}")

#Aplicar Target Encoding
print(f"Aplicando encoding...")
global_mean = df[target_col].mean()
print(f"      - Media global: {global_mean:.4f}")

stats['encoded'] = (stats['count'] * stats['mean'] + smooth * global_mean) / (stats['count'] + smooth)

# Crear mapping
encoding_map = dict(zip(stats[categorical_col], stats['encoded']))

# Aplicar encoding
new_col_name = f'{categorical_col}_encoded'
df[new_col_name] = df[categorical_col].map(encoding_map)
df[new_col_name] = df[new_col_name].fillna(global_mean)

# Verificar resultados
print(f"Resultados:")
print(f"      - Nueva columna: {new_col_name}")
```

Ilustración 16:Target Encoding en columnas categóricas

En la Ilustración 16 presenta el código responsable de aplicar la técnica de Target Encoding a variables categóricas en un conjunto de datos. Primero, se cargan y verifican las columnas del dataset, identificando aquellas de tipo "object" (que suelen ser categóricas). Luego, se emplea un método de Target Encoding que usa la media de la variable "Anomaly" para asignar valores numéricos a cada categoría, incluyendo un factor de suavizado para evitar que categorías con pocos registros tengan un impacto excesivo en el modelo. Tras realizar el Target Encoding, se eliminan las columnas originales, dejando solo las columnas transformadas para análisis posterior.

```python
# 1. CONVERSIÓN DE FECHAS Y TIEMPOS
df['Date'] = pd.to_datetime(df['Date'])

# Convertir Time directamente a datetime para extraer componentes
df['Time_dt'] = pd.to_datetime(df['Time'], format='%H:%M:%S', errors='coerce')

# Si falla la conversión, intentar formato alternativo
if df['Time_dt'].isna().any():
    print("   Usando formato alternativo para Time...")
    df['Time_dt'] = pd.to_datetime(df['Time'], errors='coerce')

# Crear datetime combinado
df['datetime'] = pd.to_datetime(
    df['Date'].dt.strftime('%Y-%m-%d') + ' ' + df['Time_dt'].dt.strftime('%H:%M:%S')
)

# 2. CARACTERÍSTICAS TEMPORALES (sin duplicaciones)
df['hour'] = df['Time_dt'].dt.hour
df['minute'] = df['Time_dt'].dt.minute
```

Ilustración 17:Conversión de la característica Time

En la Ilustración 17 se realizó la separación de los componentes horarios desde la columna "Time", extrayendo por separado la hora, minuto y segundo en columnas individuales.

```
# Ventana móvil para el ratio de bytes enviados/recibidos
df['Upload_Download_Ratio'] = df['Sentbyte'] / (df['Rcvdbyte'] + 1)  # Evitar división por cero

# Sumar bytes enviados y recibidos para obtener el volumen total de tráfico
df['ByteTotal'] = df['Sentbyte'] + df['Rcvdbyte']
 # Calcular bytes por segundo (total bytes / duración)

# Evitar división por cero en duración
df['bytes_per_second'] = np.where(
    df['Duration'] > 0,
    df['ByteTotal'] / df['Duration'],
    0  # Si duración es 0, bytes_per_second = 0
)

# Calcular ratio de paquetes (enviados/recibidos)
df['packet_ratio'] = np.where(
    df['Rcvdpkt'] > 0,
    df['Sentpkt'] / df['Rcvdpkt'],
    df['Sentpkt']  # Si no hay paquetes recibidos, usar solo enviados
```

Ilustración 18:Creación de nuevas características

Realizando la ingeniería de características se crearon cuatro nuevas variables numéricas que se muestran en la Ilustración 18, que permitirán crear otras características que ayudan identificar capturan patrones tácticos documentados en campañas APT reales y permitirán a los modelos de machine learning detectar comportamientos que escapan a las reglas de seguridad tradicionales como son:

Detección de Beaconing (Comando y Control)

El beaconing es importante para detectar las APT, ya que permite comunicaciones periódicas con servidores C2. Esta métrica identifica conexiones regulares entre 3.5 minutos y 1 hora, patrón típico usado por grupos avanzados para mantener persistencia sin ser detectados.

```
# 1. Comportamiento de beaconing (periodicidad en comunicaciones)
# Usar 'datetime' como timestamp y 'Sessionid' como IP origen
df_sorted = df.sort_values(['Sessionid', 'Eventtime'])
df_sorted['beaconing_score'] = (
df_sorted.groupby('Sessionid')['Eventtime']
.diff()
.div(1e9)  # Convertir nanosegundos a segundos
.apply(lambda x: 1 if (x > 210 and x < 3600) else 0 if pd.notna(x) else 0)
```

Ilustración 19: Comportamiento de beaconing (periodicidad en comunicaciones: Elaboración Propia

Detección de Exfiltración por Volumen

Las APTs exfiltran datos sensibles mediante transferencias anómalas grandes. El percentil 95 identifica valores extremos en el volumen de datos, detectando cuando se transfieren archivos completos o bases de datos a través de canales encubiertos.

```
# 2. Patrones de exfiltración (tráfico inusualmente grande)
# Usar 'ByteTotal' como tamaño de paquete
if 'ByteTotal' in df_sorted.columns:
    q95 = df_sorted['ByteTotal'].quantile(0.95)
    df_sorted['exfiltration_risk'] = (df_sorted['ByteTotal'] > q95).astype(int)
    print(f"   - Umbral exfiltración (Q95): {q95:.2f} bytes")
```

Ilustración 20:Patrones de exfiltración: Elaboración Propia

Detección de Escaneo de Puertos

El reconocimiento de red es la fase inicial de toda APT. Esta métrica detecta escaneos verticales donde un mismo host consulta múltiples puertos destino, indicando mapeo de servicios y vulnerabilidades en la infraestructura.

```
# 3. Scanning behavior (múltiples puertos destino desde misma IP)
# Usar 'Dstport' como puerto destino
if 'Dstport' in df_sorted.columns:
    df_sorted['scanning_score'] = df_sorted.groupby('Srcip')['Dstport'].transform('nunique')
    df_sorted['scanning_score'] = (df_sorted['scanning_score'] > 5).astype(int)
```

Ilustración 21: Scanning behavior: Elaboración Propia

Actividad en Horario No Laboral

Las APTs operan estratégicamente durante horarios de baja supervisión. Esta detección identifica actividad sospechosa en horarios nocturnos o fines de semana, cuando los equipos de seguridad tienen menor capacidad de respuesta.

```
# 4. Características temporales avanzadas
if 'hour' in df_sorted.columns:
    df_sorted['es_horario_no_laboral'] = ((df_sorted['hour'] < 8) | (df_sorted['hour'] > 18)).astype(int)
```

Ilustración 22: Características temporales avanzadas: Elaboración Propia

Detección de Exfiltración por Asimetría

El tráfico normal tiene ratios balanceados. Cuando el upload excede significativamente al download, indica posible exfiltración de datos. Ratio >10:1 sugiere transferencia masiva de información hacia el exterior.

```
# 5. Comportamiento anómalo adicional
# Ratio alto de upload/download puede indicar exfiltración
if 'Upload_Download_Ratio' in df_sorted.columns:
    df_sorted['high_upload_ratio'] = (df_sorted['Upload_Download_Ratio'] > 10).astype(int)
```

Ilustración 23:Detección de Exfiltración por Asimetría: Elaboración Propia

Sesiones de Larga Duración (C2 Persistente)

Las sesiones C2 persistentes mantienen conexiones activas por largos períodos para evitar reconstrucción de patrones de beaconing. Sesiones >1 hora son anómalas en tráfico corporativo normal y sugieren canal de comando activo.

```
# 6. Sesiones de larga duración (posible C2)
if 'Duration' in df_sorted.columns:
    df_sorted['long_session'] = (df_sorted['Duration'] > 3600).astype(int)  # Más de 1 hora
```

Ilustración 24:Sesiones de larga duración: Elaboración Propia

Se procederá a realizar una depuración selectiva de columnas, eliminando aquellas que presenten un bajo desempeño en la matriz de correlación con la variable objetivo, así como las que contengan valores mayoritariamente vacíos o muestren información redundante con otras características del dataset, la característica de horario laborar se elimina ya que el dataset no aplica y así quedaría la dimensionalidad del conjunto de datos para el posterior modelado.

Adicionalmente, este proceso de reducción de atributos permite optimizar el rendimiento de los algoritmos de aprendizaje automático, disminuyendo el ruido y evitando el sobreajuste. Al conservar únicamente las variables más relevantes y de mayor impacto predictivo, se facilita la interpretación del modelo y se mejora la eficiencia en términos de tiempo de entrenamiento y consumo de recursos computacionales. Esta depuración también contribuye a una mayor estabilidad en los resultados, al eliminar datos que podrían distorsionar la capacidad de generalización del sistema.

```
# Primero: Añadir las características temporales básicas
df_with_temporal = add_temporal_features(df)

# Segundo: Añadir características específicas para APT
df_with_apt = extraer_caracteristicas_apt(df_with_temporal)

# Tercero: Eliminar columnas específicas (DESPUÉS de crear todas las características)
columns_to_drop = [
     # 'Eventtime',
    'Tz', 'Logid', 'Type', 'Srcname', 'Srcintf', 'Dstintf',
    'Dstintfrole', 'Poluuid', 'Policyname', 'Transip',
    'Srchwvendor', 'Osname', 'Mastersrcmac', 'Srcmac', 'Appid', 'Applist',
    'Utmref', 'Authserver', 'Countwaf', 'hour','minute','Date' , 'Srcip', 'Dstip', 'datetime'
    ,'Time', 'Level', 'Transport', 'Lanout','Appcat', 'es_horario_no_laboral'
]
```

Ilustración 25:Limpieza de datos: Elaboración Propia

La matriz de correlación en la Ilustración 26 muestra las relaciones lineales entre las variables numéricas del dataset ya limpio. Las correlaciones detectadas son útiles para reconocer relaciones fuertes entre variables, lo que puede facilitar la reducción de la multicolinealidad en modelos predictivos mediante la eliminación o la combinación de variables altamente correlacionadas.

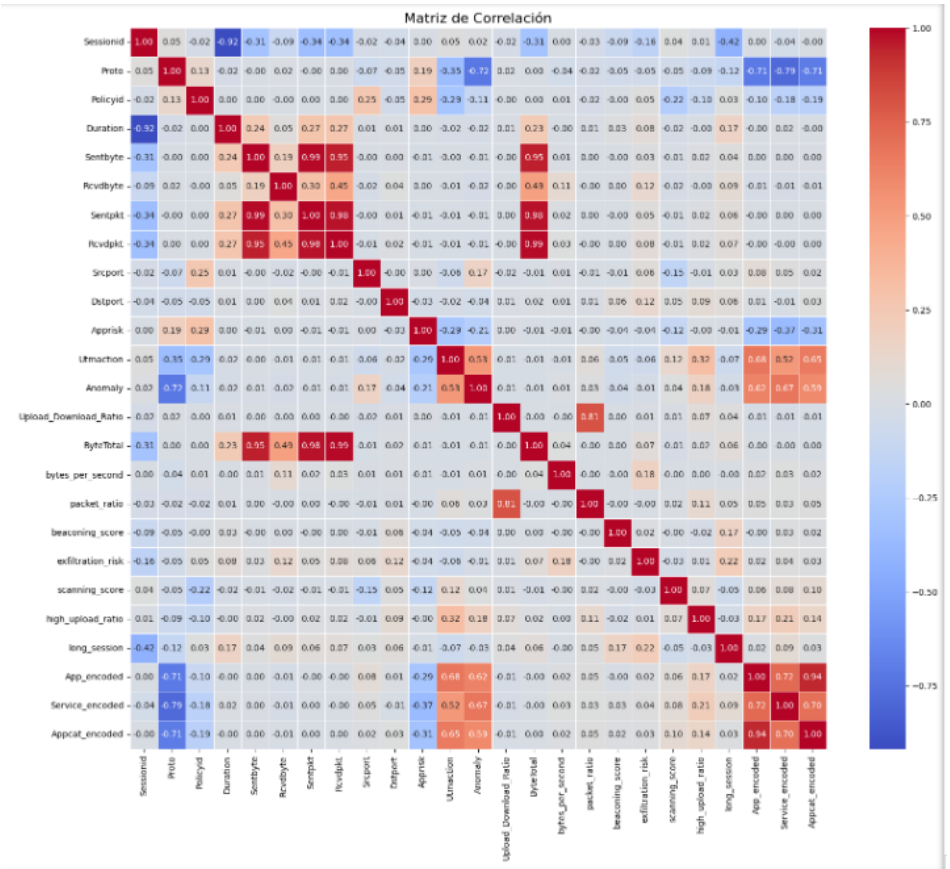


Ilustración 26: Matriz de correlación de dataset limpio: Elaboración Propia

En este análisis de selección de características, se ha identificado un conjunto optimizado de variables predictoras para la clase objetivo "Anomaly", basado en su alta correlación. Las características con mayor poder predictivo incluyen Utmaction, Apprisk, Service_encoded, Duration, Appcat_encoded, y long_session, todas con correlaciones moderadas a altas con "Anomaly". Se conservan variables relacionadas con el tráfico, como Sentbyte y Sentpkt, además de otras útiles como Proto, Dstport y scanning_score. Por otro lado, se eliminan variables redundantes o que muestran baja correlación, como Rcvdbyte, Rcvdpkt, ByteTotal, y también las que no aportan un valor predictivo significativo, como Eventtime, Sessionid, Policyid, Srcport, y Dstinetsvc_encoded. Este proceso optimiza las características para modelos predictivos y reduce la multicolinealidad.

Después del análisis en la Ilustración 27 se muestra la matriz de confusión de las características seleccionadas para comenzar con los algoritmos elegidos.

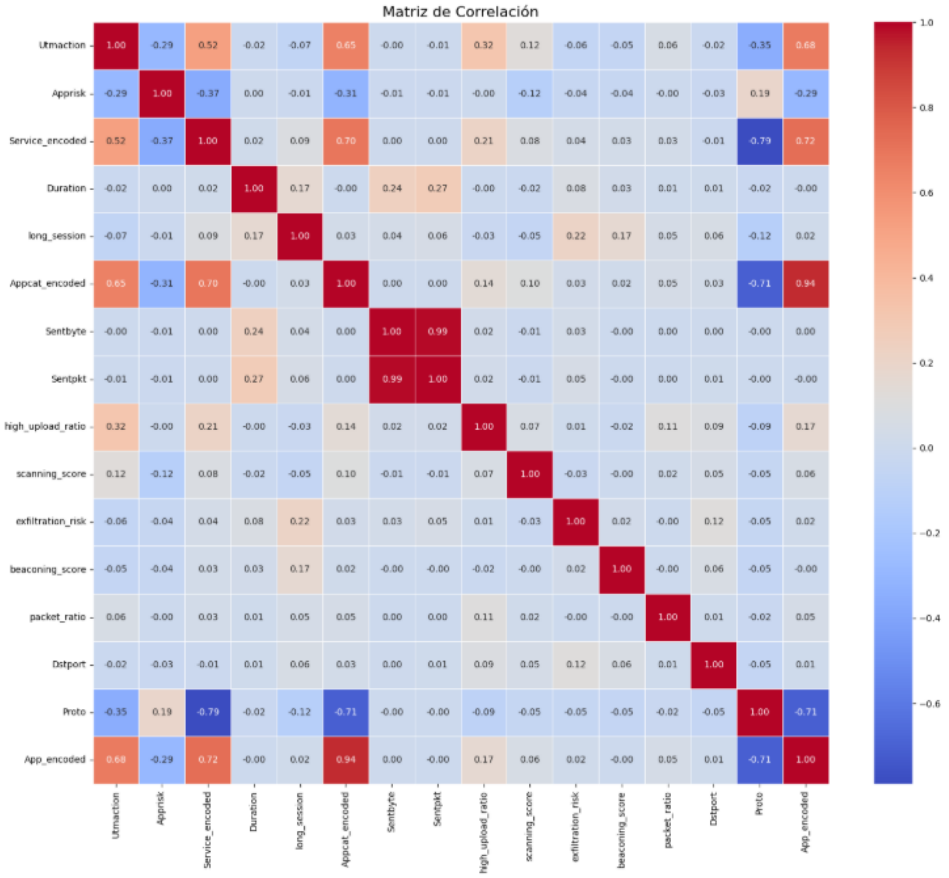


Ilustración 27:Matriz de correlación de las características seleccionadas: Elaboración Propia

2.5.4 FASE DE MODELADO

En la fase de modelado se probarán diferentes algoritmos de aprendizaje automático utilizando Jupyter Notebook como entorno de trabajo para identificar cuál ofrece el mejor desempeño al trabajar con las columnas seleccionadas durante la fase de selección de características. Los modelos que se probarán son K-Means, Isolation Forest, Random Forest y SVM.

Se entrenarán estos modelos con las columnas seleccionadas, y se evaluarán utilizando métricas como precisión, Recall, F1-Score y AUC-ROC, para identificar cuál de estos modelos ofrece el mejor desempeño en la predicción de "Anomaly". Además, se realizarán ajustes en los parámetros de cada modelo para optimizar sus resultados.

Random Forest

Este modelo es un conjunto de árboles de decisión que predicen la clase como en el trabajo realizado que son "normal" o "anómalo" mediante la votación de múltiples árboles. Se utiliza principalmente para tareas de clasificación y es especialmente útil cuando hay un gran número de características y cuando las relaciones entre las características no son lineales.

El código carga un conjunto de datos desde un archivo CSV y seleccionó las características más relevantes para la detección de anomalías utilizando Random Forest. Primero, se seleccionaron las columnas que representan las características relevantes y la variable objetivo Anomaly. Luego, se aplicó StandardScaler para normalizar las características, asegurando que todas tuvieran la misma escala. El conjunto de datos se dividió en entrenamiento y prueba en proporciones del 80 % y 20%, respectivamente, y se entrenó el modelo con los datos de entrenamiento. Después, se evaluó el modelo usando varias métricas, como precisión, recall, F1-score y accuracy. Las métricas se evaluaron en conjuntos de entrenamiento y

prueba. Se mostró el rendimiento del modelo con una matriz de confusión y también se calculó el AUC de la curva ROC.

```python
import pandas as pd
import numpy as np
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import classification_report, accuracy_score, roc_auc_score,
precision_recall_curve, roc_curve, confusion_matrix
import matplotlib.pyplot as plt
import seaborn as sns

# Cargar los datos
df = pd.read_csv('output_gru_processed.csv')

# Selecciona las características relevantes para la detección de anomalías

X = df[[
    'Utmaction', 'Apprisk', 'Service_encoded','Duration','long_session','Appcat_encoded',
    'Sentbyte','Sentpkt',
    'high_upload_ratio',
    'scanning_score',
    'exfiltration_risk',
    'beaconing_score',
    'packet_ratio',
    'Dstport',
    'Proto',
    'App_encoded'
]]
# Separar características y etiquetas
y = df['Anomaly']
# Información inicial del dataset
print(f"Dimensiones del dataset: {df.shape}")
print(f"Columnas: {df.columns.tolist()}")
print("\nDistribución de etiquetas originales:")
print(df['Anomaly'].value_counts())
print(f"Proporción anomalías: {df['Anomaly'].mean():.3f}")

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

# Dividir los datos en conjunto de entrenamiento y prueba (80% entrenamiento, 20% prueba)
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.2, random_state=42,stratify=y)
print(f"\nCaracterísticas utilizadas: {X.columns.tolist()}")
print(f"Número de características: {X.shape[1]}")
```

```python
# Crear y entrenar el modelo de Random Forest
model_rf = RandomForestClassifier(class_weight='balanced', n_estimators=150, random_state=42)
```

Ilustración 28:Código de modelo Random Forest

Parámetro	*Valor*	*Descripción*
class_weight	'balanced'	Ajusta los pesos de las clases para manejar el desbalance entre las clases (en este caso, las anomalías).
n_estimators	150	Número de árboles en el bosque. Un valor más alto puede mejorar la precisión, pero también incrementa el tiempo de computación.
random_state	42	Semilla aleatoria para asegurar que los resultados sean reproducibles.

Tabla 15:Parámetros utilizados en el modelo Random Forest

Isolation Forest

Este modelo es un algoritmo de detección de anomalías basado en un conjunto de árboles de decisión aleatorios que no votan por una clase, sino que miden qué tan fácil es aislar cada registro: las observaciones que se aíslan con pocas particiones del árbol se consideran “anómalas”, mientras que las que requieren más particiones se consideran “normales”.

```
# Importar las librerías necesarias
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.ensemble import IsolationForest
from sklearn.metrics import roc_curve, auc, classification_report, confusion_matrix
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler

# Cargar los datos
df = pd.read_csv('output_gru_processed.csv')
# Verificar si hay valores nulos o faltantes en el DataFrame
print(df.isnull().sum())
print("SELECCIÓN DE CARACTERÍSTICAS ESPECÍFICAS")
print("=" * 50)

# Definir las características que quieres usar
X = df[[ 'Utmaction','Apprisk','Service_encoded','Duration','long_session','Appcat_encoded',
         'Sentbyte','Sentpkt', 'high_upload_ratio', 'scanning_score','exfiltration_risk',
         'beaconing_score', 'packet_ratio','Dstport','Proto', 'App_encoded']]
y = df['Anomaly']
```

```
print(f"\n Estadísticas del dataset con características seleccionadas:")
print(f"   • Forma de X: {X.shape}")
print(f"   • Anomalías: {y.sum()} ({y.mean()*100:.2f}%)")

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

# Dividir los datos en conjunto de entrenamiento y prueba (80% entrenamiento, 20% prueba)
X_train, X_test, y_train, y_test = train_test_split(X_scaled, y, test_size=0.2, random_state=42)

# Inicializar el modelo Isolation Forest con PARÁMETROS ÓPTIMOS
iso_forest = IsolationForest(
    n_estimators=150,
    max_samples='auto',
    contamination=0.41,
    max_features=0.5,
    random_state=42,
    bootstrap=True,
    n_jobs=-1,
    verbose=0
)

# Ajustar el modelo con los datos de entrenamiento
print("Entrenando modelo con parámetros óptimos...")
iso_forest.fit(X_train)
```

Tabla 16:Código de modelo Isolation Forest

El código cargó un conjunto de datos desde un archivo CSV y seleccionó las características más relevantes para la detección de anomalías utilizando Isolation Forest. Primero, se seleccionaron las columnas que representan las características clave y la variable objetivo Anomaly. A continuación, se aplicó StandardScaler para normalizar las características, asegurando que todas tuvieran la misma escala. El conjunto de datos se dividió en entrenamiento y prueba en proporciones del 80 % y 20%, respectivamente, y se entrenó el modelo con los datos de entrenamiento. Posteriormente, el modelo fue ajustado con parámetros óptimos y evaluado utilizando varias métricas, como la precisión, recall, F1-score y la matriz de confusión. Además, se calculó el AUC de la curva ROC para visualizar el rendimiento del modelo.

Asimismo, se verificó la estabilidad del modelo analizando la distribución de valores predichos y evaluando si existían sesgos hacia ciertas clases, lo cual es común en problemas de detección de anomalías con datos desbalanceados. También se generaron gráficas complementarias, como curvas de densidad y proyecciones en dos dimensiones, para observar la separación entre instancias normales y anómalas después de la estandarización. Finalmente, los resultados obtenidos permitieron validar la efectividad del enfoque implementado y establecer una base sólida para futuras mejoras en los métodos de detección de intrusiones.

Parámetro	*Valor*	*Descripción*
n_estimators	150	Número de árboles en el bosque. Un valor más alto puede mejorar la precisión, pero también incrementa el tiempo de computación.
contamination	41	Establece la proporción de anomalías calculadas anteriormente.
bootstrap	True	Indica si se deben usar muestras con reemplazo al construir los árboles.
n_jobs	-1	Utiliza todos los núcleos del procesador disponibles para paralelizar la computación.
verbose	0	Controla el nivel de información impresa durante el entrenamiento (0 para silencio).

Tabla 17:Parámetros utilizados en el modelo Isolation Forest

SVM (Máquina de Vectores de Soporte)

SVM (Support Vector Machine) es un modelo de clasificación supervisada que identifica el mejor margen de separación entre dos clases, en este caso, "normales" y "anómalas". Su objetivo principal es encontrar un hiperplano que discrimine eficazmente las clases, maximizando la distancia del margen, lo que lo hace especialmente útil en problemas de dos clases. Aunque funciona mejor con datos con una separación clara, también puede manejar casos no lineales mediante kernels, lo que le permite trazar fronteras de decisión complejas en espacios de alta dimensión.

SVM destaca por su robustez en entornos donde el ruido es moderado y los datos se encuentran desbalanceados, ya que el margen maximizado ayuda a evitar decisiones influenciadas por valores atípicos. Su capacidad para generalizar bien incluso con pocas muestras lo convierte en una herramienta valiosa en escenarios donde la cantidad de datos etiquetados es limitada. Asimismo, su flexibilidad al incorporar distintos tipos de *kernels* —como RBF, polinomial o sigmoidal—

permite adaptar el modelo a diferentes formas de distribución de los datos, aumentando la precisión en la detección de anomalías en sistemas complejos.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.svm import SVC
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import (confusion_matrix, classification_report, roc_curve, auc, accuracy_score,
precision_score, recall_score, f1_score, roc_auc_score)
from imblearn.over_sampling import SMOTE
import warnings
warnings.filterwarnings('ignore')

# Configurar para usar menos memoria
plt.rcParams['figure.max_open_warning'] = 0

# Cargar datos
df = pd.read_csv('output_gru_processed.csv')

print("Distribución de Anomaly:")
print(df['Anomaly'].value_counts())

# Separar características y variable objetivo
X = df[[  'Utmaction','Apprisk','Service_encoded','Duration','long_session','Appcat_encoded',
          'Sentbyte','Sentpkt','high_upload_ratio', 'scanning_score','exfiltration_risk','beaconing_score',
          'packet_ratio','Dstport', 'Proto','App_encoded']]
y = df['Anomaly']

# Liberar memoria del dataframe original
del df

# Dividir los datos
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42, stratify=y)

print(f"\nDatos de entrenamiento: {X_train.shape}")

# Aplicar SMOTE
print("Aplicando SMOTE...")
smote = SMOTE(random_state=42)
X_train_balanced, y_train_balanced = smote.fit_resample(X_train, y_train)
print(f"Después de SMOTE: {X_train_balanced.shape}")
```

```
# Aplicar StandardScaler
print("Aplicando StandardScaler...")
scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train_balanced)
X_test_scaled = scaler.transform(X_test)

# Liberar memoria intermedia
del X_train_balanced

# Entrenar modelo SVM con probability=True solo para AUC
print("Entrenando modelo SVM...")
model = SVC(
    kernel='rbf',
    probability=True,  # True para calcular AUC
    random_state=42,
    cache_size=500
)
model.fit(X_train_scaled, y_train_balanced)
```

Ilustración 29:Código de modelo SVM

El código cargó un conjunto de datos desde un archivo CSV y seleccionó las características más relevantes para la clasificación utilizando SVM. Primero, se seleccionaron las columnas clave que representan las características y la variable objetivo Anomaly. Luego, se aplicó StandardScaler para estandarizar las características, asegurando que todas tuvieran la misma escala. Posteriormente, el conjunto de datos se dividió en entrenamiento y prueba en proporciones del 80 % y 20%, respectivamente. El modelo SVM fue entrenado utilizando los datos de entrenamiento, ajustando los parámetros para optimizar el rendimiento. Después del entrenamiento, se evaluó el modelo utilizando diversas métricas como precisión, recall, F1-score y la matriz de confusión. También se calculó el AUC de la curva ROC para visualizar la capacidad de discriminación del modelo.

Parámetro	***Valor***	***Descripción***
C	1.0	Parámetro de penalización que controla el margen de error permitido.
kernel	'rbf'	Función de núcleo (kernel) utilizada para transformar los datos en un espacio de mayor dimensión. Se usa la función kernel radial (RBF), que es útil para manejar datos no lineales.
random_state	42	Semilla aleatoria que se utiliza para garantizar la reproducibilidad de los resultados.
probability	True	Indica que se deben calcular las probabilidades de pertenecer a cada clase. Esto es útil para calcular métricas como AUC.
degree	3	Parámetro utilizado solo para el kernel polinómico, especifica el grado del polinomio (no utilizado en el kernel RBF).

Tabla 18:Parámetros utilizados en el modelo SVM

K-means

K-Means es un algoritmo de clustering no supervisado utilizado para categorizar datos en distintos clústeres basados en su similitud. Su objetivo principal es dividir un conjunto de datos en un número específico de grupos definido de clústeres (K), en donde cada dato se asigna al clúster cuyo centroide (el centro de masa) está más cercano. Mediante un proceso repetitivo, el algoritmo asigna los puntos a los clústeres y actualiza los centroides hasta que las asignaciones de puntos y centroides dejan de cambiar significativamente. Aunque K-Means es eficiente en la agrupación de datos con estructuras claras y separadas, también puede ser sensible a la elección de K y a los puntos de inicio de los centroides. A pesar de su simplicidad, K-Means es muy utilizado en problemas de segmentación y clasificación cuando las relaciones entre los datos no son necesariamente lineales.

```python
import pandas as pd
import numpy as np
from sklearn.preprocessing import StandardScaler, LabelEncoder
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt
import seaborn as sns

df = pd.read_csv('output_gru_processed.csv')
def preparar_datos_kmeans(df):
    # Seleccionar y limpiar características numéricas relevantes para APT
    features = df[['Utmaction','Apprisk','Service_encoded','Duration','long_session','Sentbyte','Sentpkt',
                   'high_upload_ratio','scanning_score','exfiltration_risk','beaconing_score','packet_ratio',
                   'Dstport','Proto','App_encoded' ,'Appcat_encoded'
    ]].copy()

    # Limpiar valores NaN
   # features = features.fillna(0)
    return features

# Preparar datos
features = preparar_datos_kmeans(df)
print("Características para K-Means:")
print("Características para K-Means:")
print(features.columns.tolist())

# Aplicar el filtro IQR (Interquartile Range) para eliminar outliers
Q1 = features.quantile(0.25)
Q3 = features.quantile(0.75)
IQR = Q3 - Q1

# Filtrar los datos fuera del rango intercuartílico
features_filtered = features[~((features < (Q1 - 1.5 * IQR)) | (features > (Q3 + 1.5 * IQR))).any(axis=1)]

# Mostrar el número de filas antes y después de eliminar outliers
print(f"Número de filas antes de eliminar outliers: {features.shape[0]}")
print(f"Número de filas después de eliminar outliers: {features_filtered.shape[0]}")

# Si se necesita estandarizar los datos
scaler = StandardScaler()
features_scaled = scaler.fit_transform(features_filtered)

# Aplicar K-Means con un número de clusters determinado
kmeans = KMeans(n_clusters=2, random_state=42)
kmeans.fit(features_scaled)

# Predicción de los clusters
labels = kmeans.predict(features_scaled)
```

Ilustración 30:Código de modelo K-means

El código proporcionado carga un conjunto de datos desde un archivo CSV y selecciona las características relevantes para el análisis utilizando el algoritmo K-Means. Se definen las características clave del conjunto de datos, luego, se aplica un filtro de valores nulos para asegurarse de que los datos estén completos, y se eliminan los outliers utilizando el Rango Intercuartílico (IQR) para crear la visualización, se utilizó PCA para disminuir la cantidad de dimensiones de los datos a dos componentes principales y facilitar la representación gráfica. Tras limpiar los datos, se escalan utilizando StandardScaler para asegurar que todas las características tengan la misma influencia en el algoritmo de K-Means. Se procedió a calcular la inercia (suma de distancias cuadradas dentro de los clústeres) para diferentes valores de K (número de clústeres) mediante el Método del Codo, tras evaluar los resultados de la gráfica plasmada en la Ilustración 31, se observó que el valor óptimo para K es 2, ya que el codo de la gráfica de inercia se encuentra en este punto, lo que indica una mejora marginal al aumentar el número de clústeres más allá de ese valor según, y se realiza la predicción de los clústeres.

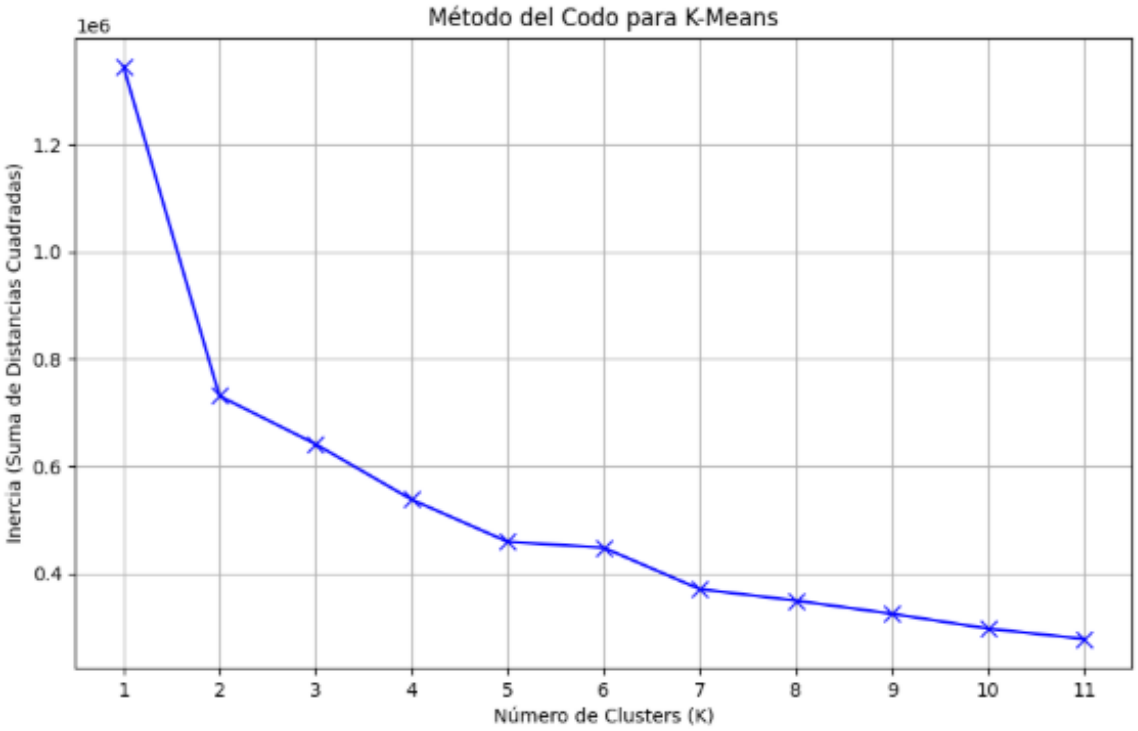


Ilustración 31: Método del Codo para calcular K: Elaboración Propia

2.5.5 FASE DE EVALUACIÓN

Evaluación de resultados de los modelos

Para evaluar los resultados de los modelos presentados en este trabajo, se utilizaron varias métricas clave que nos permiten medir el rendimiento de los modelos de manera sencilla y efectiva. Entre las métricas utilizadas se incluyen:

- Recall (Sensibilidad o Exhaustividad)
  Mide qué proporción de los casos positivos reales fueron detectados correctamente.
- F1-score
  Es la media armónica entre precision y recall; equilibra ambas métricas.
- Accuracy (Exactitud)
  Mide qué proporción de todas las predicciones (positivas y negativas) fueron correctas.
- Precision (Precisión)
  Mide qué proporción de las predicciones positivas realmente eran correctas.

La Ilustración 32 muestra que desde la matriz de confusión se asocian las métricas que evalúan los algoritmos.

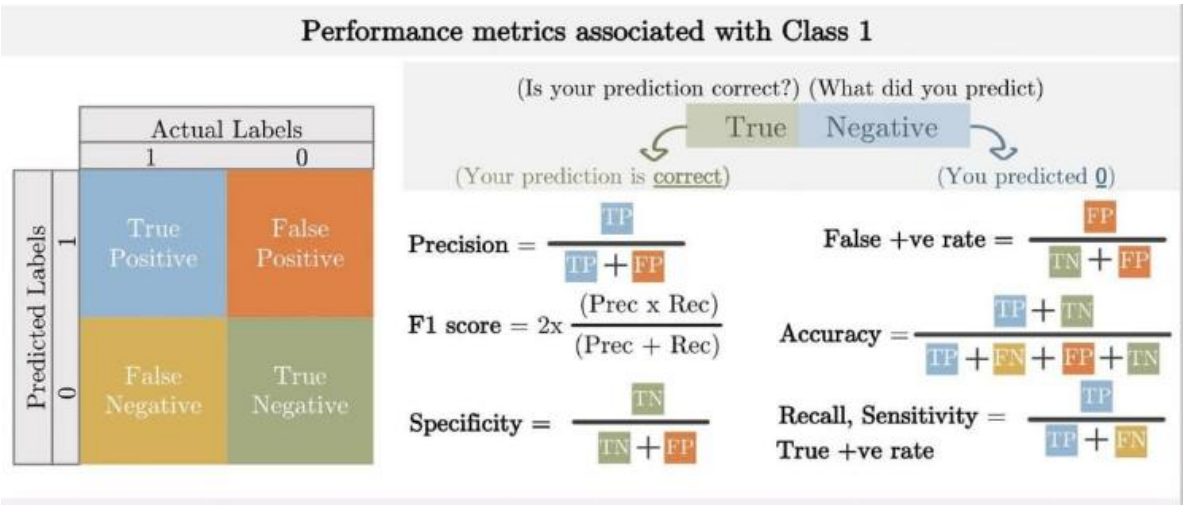


Ilustración 32:Métricas de rendimiento:Cheatsheets [82]

Además, se calculó el AUC (Área bajo la Curva) de la curva ROC, lo que nos permitió evaluar la su habilidad para distinguir entre las clases anomala y normal a través de diferentes umbrales de decisión. Estas métricas son esenciales para comprender tanto el desempeño general del modelo como su efectividad en la detección de anomalías.

SVM	
Métricas	Resultados
Recall	0.97
F1-score	0.85
Accuracy	0.87
Precision	0.76

Tabla 19: Métricas de presión de SVM

Random Forest	
Métricas	Resultados
Recall	0.90
F1-score	0.87
Accuracy	0.90
Precision	0.85

Tabla 20: Métricas de presión de Random Forest

Isolation Forest	
Métricas	Resultados
Recall	0.27
F1-score	0.38
Accuracy	0.64
Precision	0.62

Tabla 22:Métricas de presión de Isolation Forest

K-means	
Métricas	Resultados
Recall	0.91
F1-score	0.81
Accuracy	0.83
Precision	0.73

Tabla 21: Métricas de presión de K-means

Las métricas de evaluación de algoritmos son fundamentales para medir el rendimiento y la capacidad de generalización de un modelo, especialmente en tareas de clasificación, regresión y detección de anomalías. Estas métricas permiten comparar diferentes enfoques, identificar debilidades y validar si el comportamiento del modelo es adecuado para el problema planteado.

La Ilustración 33 muestra dos gráficas de evaluación de K-means. A la izquierda, la Curva ROC con un AUC de 0.85 indica que el modelo tiene una buena capacidad para distinguir entre anomalías y tráfico normal, ya que un valor cercano a 1 (como 0.85) refleja un alto desempeño en la clasificación. A la derecha, la Matriz de Confusión dada por todos los datos revela que el modelo identificó correctamente 145,617 anomalías (verdaderos positivos) y 126,811 casos normales (verdaderos negativos), pero también cometió errores: clasificó 53,205 anomalías como normales (falsos negativos) y 183,268 casos normales como anomalías (falsos positivos), lo que sugiere una tasa de falsas alarmas elevada pero una detección razonable de anomalías reales.

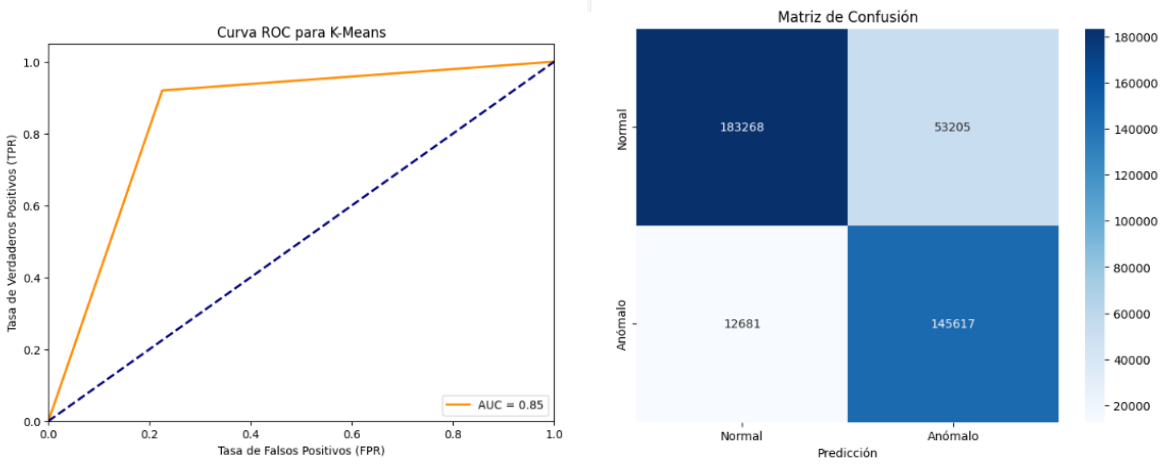


Ilustración 33: Curva ROC y matriz de confusión de K-means

La Ilustración 34 presenta la evaluación del modelo Isolation Forest, mostrando una Curva ROC con AUC de 0.746 que indica un desempeño moderado en la detección de anomalías, la matriz de confusión que es dada por los datos de entrenamiento revela que el modelo identificó correctamente 5,119 anomalías (verdaderos positivos) y 42,171 casos normales (verdaderos negativos), pero cometió 8,659 falsos negativos (anomalías no detectadas) y 23,006 falsos positivos (tráfico normal clasificado como anómalo), demostrando una precisión muy baja con alta tasa de falsas alarmas y baja sensibilidad. Este comportamiento sugiere que el modelo no logra separar adecuadamente los patrones anómalos de los normales, posiblemente debido al desbalance de clases o a la complejidad del comportamiento del tráfico analizado.

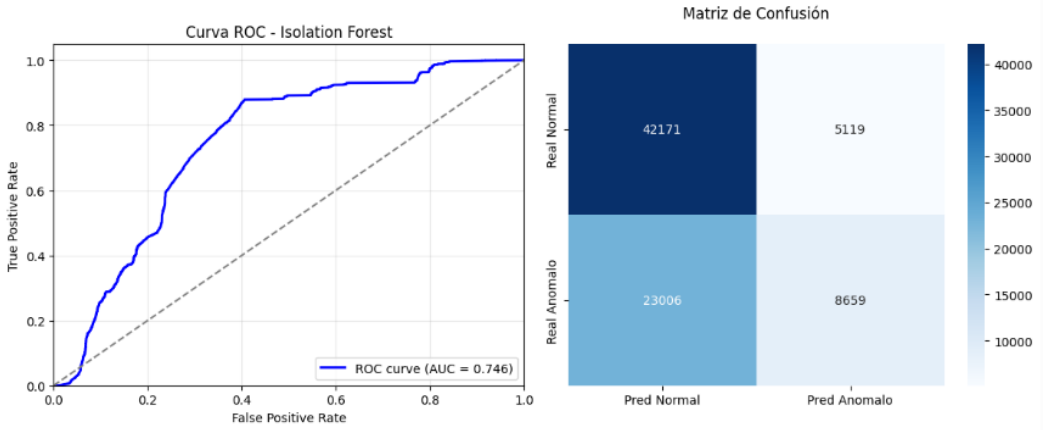


Ilustración 34: Curva ROC y matriz de confusión de Isolation Forest

El modelo Random Forest demostró ser muy eficaz para detectar anomalías, alcanzando un Área Bajo la Curva (AUC) de 0.97 en la Curva ROC, lo que refleja una capacidad casi perfecta para distinguir un caso anómalo de uno normal. Además, su matriz de confusión, basada en los datos de entrenamiento, mostró un rendimiento muy sólido, reduciendo tanto los falsos negativos como los falsos positivos, convirtiéndolo en un clasificador altamente confiable.

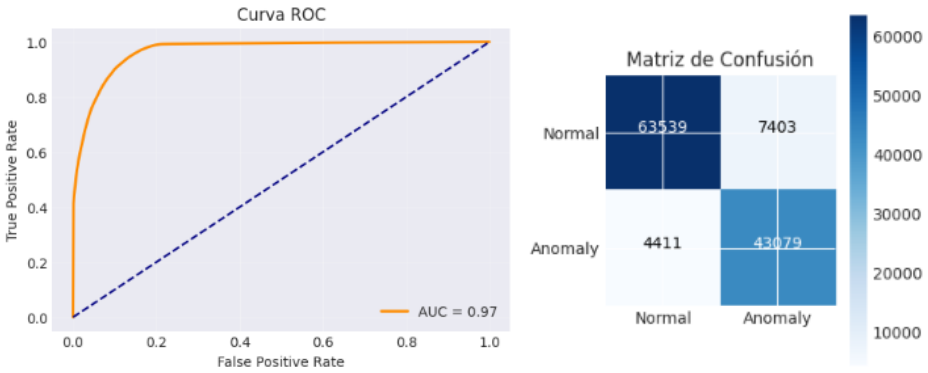


Ilustración 35: Curva de ROC y matriz de confusión de Random Forest

El modelo de Máquina de Vectores de Soporte (SVM) ofrece un rendimiento muy bueno para la detección de anomalías, como lo demuestra su Área Bajo la Curva (AUC) de 0.944 en la Curva ROC, confirmando su alta capacidad discriminatoria. Su Matriz de Confusión destaca por tener la menor cantidad de Falsos Negativos entre los modelos analizados, lo que significa que es extremadamente efectivo para no omitir anomalías reales. Sin embargo, esta alta sensibilidad viene con el costo

de generar 14,346, resultando en más falsas alarmas, lo que puede afectar su precisión general.

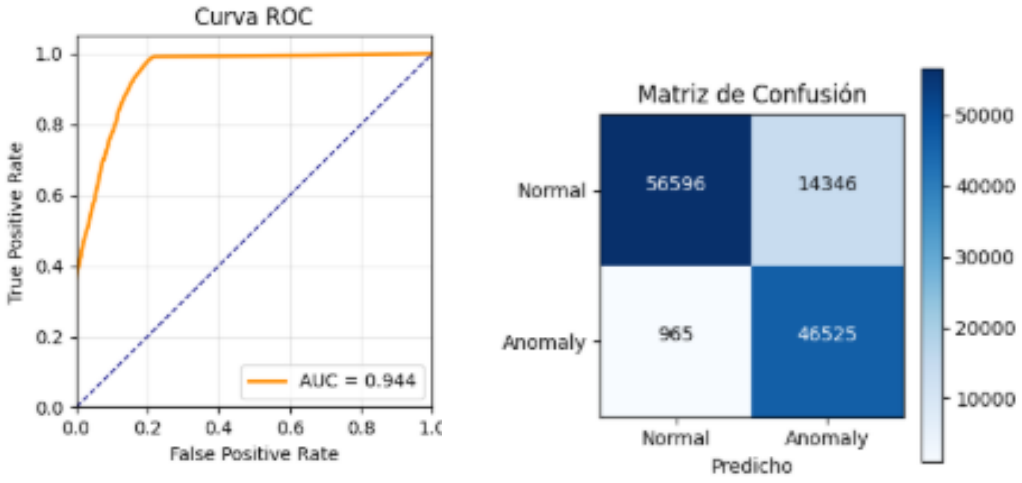


Ilustración 36: Curva de ROC y matriz de confusión de SVM

2.5.6 FASE DE EXPLOTACIÓN

Para la fase de explotación, el dashboard presentado constituye la interfaz operativa que materializa los hallazgos del modelo en un entorno práctico de monitoreo de seguridad. Permite a los usuarios cargar nuevos archivos de red en formato CSV, utilizando Random Forest y visualizar de manera centralizada los resultados del análisis, incluyendo la clasificación del tráfico, las alertas generadas y las métricas de rendimiento, las curvas de aprendizaje y una gráfica que muestra la distribución del número de conexiones por protocolo/puerto destino, etiquetando cada conexión como normal (rosa) o anomalía (marrón). Esta herramienta transforma la complejidad de los datos y las predicciones del modelo en información accionable, facilitando la toma de decisiones y la respuesta ante posibles incidentes, y constituye el puente esencial entre la investigación y la aplicación operativa continua del sistema, permitiendo la explotación sostenible del conocimiento generado y cerrando efectivamente el ciclo de la metodología CRISP-DM con una herramienta adaptada a las necesidades operativas reales.

Además, proporciona visualizaciones claras y reportes automatizados que permiten a los analistas identificar rápidamente patrones de riesgo y áreas críticas de mejora. Su integración con sistemas de monitoreo y alerta asegura que las decisiones

basadas en los modelos se implementen de manera oportuna, fortaleciendo la resiliencia y seguridad de la infraestructura.

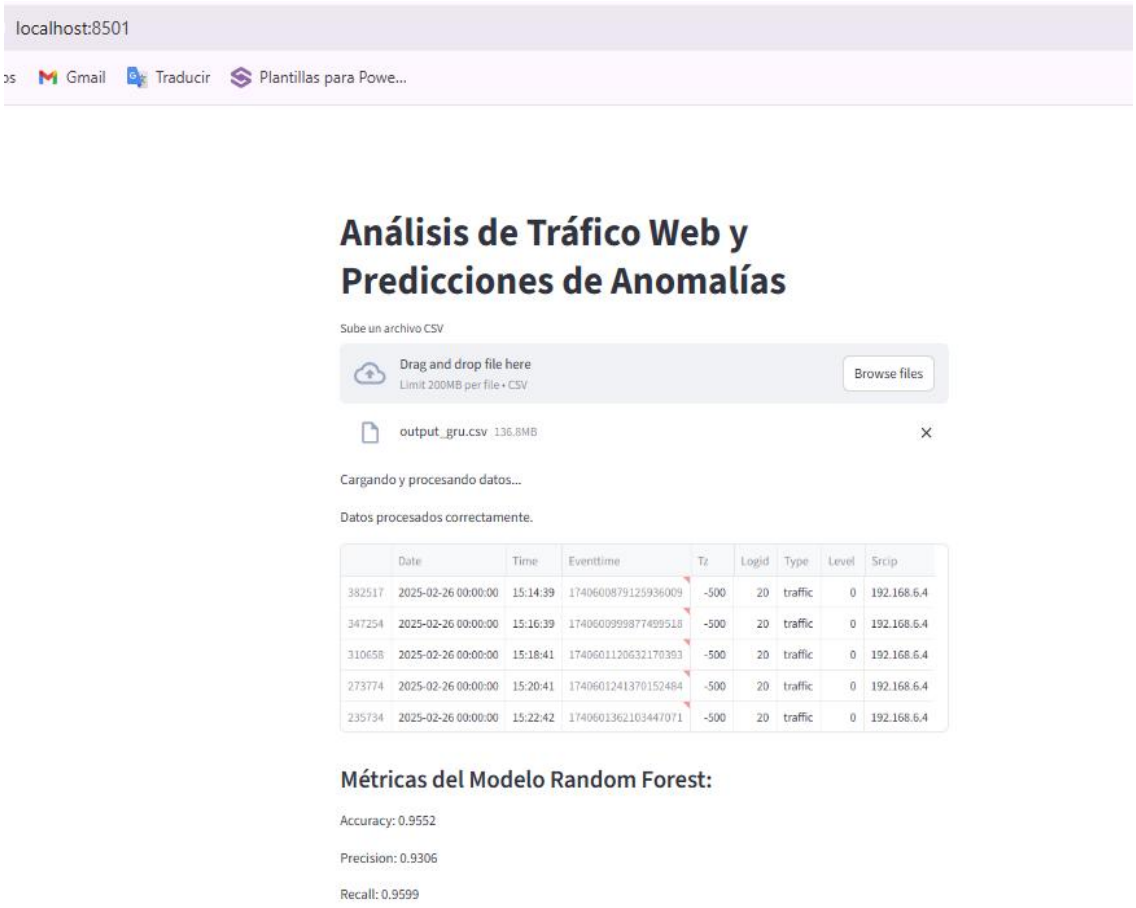


Ilustración 37: Dashboard

El proceso de detección de anomalías en el tráfico de red se realiza en varias fases: carga de datos, preprocesamiento, entrenamiento del modelo, evaluación y visualización de los resultados. Todo el proceso se desarrolló utilizando una aplicación interactiva con Streamlit.

El proceso inicia con la carga de un archivo CSV a través de una interfaz de usuario en Streamlit. Tras esto, se llevan a cabo varias transformaciones.

- El archivo subido debe ser de tipo .csv; de lo contrario, el sistema solicitará que se cargue un archivo con el formato adecuado.
- Conversión de fechas y horas a formato adecuado.
- Generación de características adicionales, como la relación de bytes enviados y recibidos y la velocidad de transmisión.

- Aplicación de Target Encoding en columnas categóricas como "Servicio" y "Aplicación".
- Detección de patrones sospechosos asociados a APT mediante la extracción de características adicionales, como beaconing, exfiltración y escaneo de puertos.

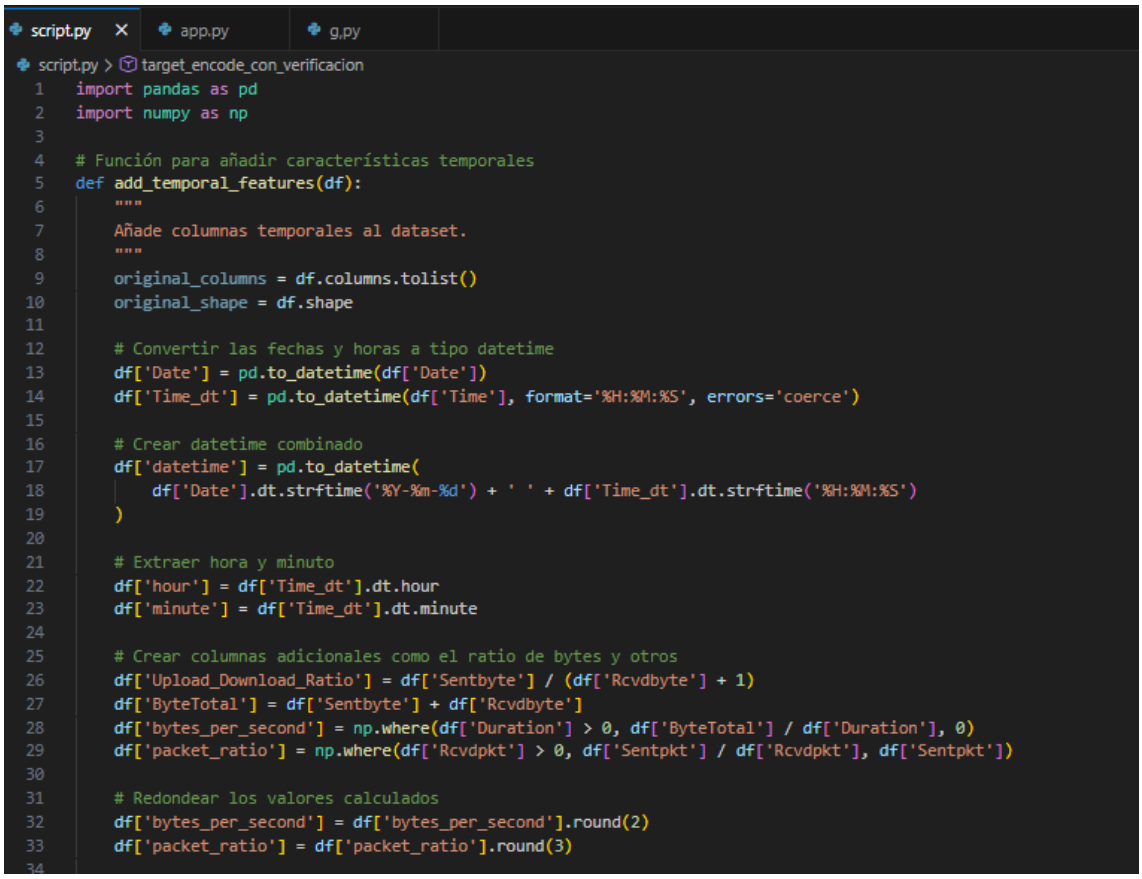


Ilustración 38: Código de script.py

El modelo Random Forest se entrena con los datos preprocesados en un conjunto de entrenamiento y se guarda como un archivo .pkl para futuras predicciones.

Se evalúa el rendimiento del modelo mediante métricas como accuracy, precision, recall, F1-score y AUC-ROC. Además, se generan curvas de aprendizaje para analizar cómo varía el rendimiento del modelo con diferentes volúmenes de datos.

Este análisis permite identificar posibles problemas de sobreajuste o subajuste y determinar si el modelo se beneficia de más datos de entrenamiento. Asimismo,

facilita la selección de hiperparámetros y estrategias de optimización para mejorar la capacidad de generalización del modelo en nuevos datos.

```python
# Predicción
y_pred = model.predict(X_scaled)
y_proba = model.predict_proba(X_scaled)[:, 1]

# Métricas
acc = accuracy_score(y, y_pred)
prec = precision_score(y, y_pred)
rec = recall_score(y, y_pred)
f1 = f1_score(y, y_pred)
auc_score = roc_auc_score(y, y_proba)

st.write("### Métricas del Modelo Random Forest:")
st.write(f"Accuracy: {acc:.4f}")
st.write(f"Precision: {prec:.4f}")
st.write(f"Recall: {rec:.4f}")
st.write(f"F1-score: {f1:.4f}")
st.write(f"AUC-ROC: {auc_score:.4f}")

# Matriz de confusión
cm = confusion_matrix(y, y_pred)
fig, ax = plt.subplots(figsize=(8,6))
im = ax.imshow(cm, interpolation='nearest', cmap=plt.cm.Blues)
ax.set_title('Matriz de Confusión')
plt.colorbar(im, ax=ax)
ax.set_xticks([0,1])
ax.set_yticks([0,1])
ax.set_xticklabels(['Normal', 'Anomaly'])
ax.set_yticklabels(['Normal', 'Anomaly'])

for i in range(2):
    for j in range(2):
        ax.text(j, i, str(cm[i, j]),
```

Ilustración 39: Evaluación de algoritmo

En la Ilustración 40 e Ilustración 41 muestra las gráficas generadas por el algoritmo Random Forest aplicado al dashboard.

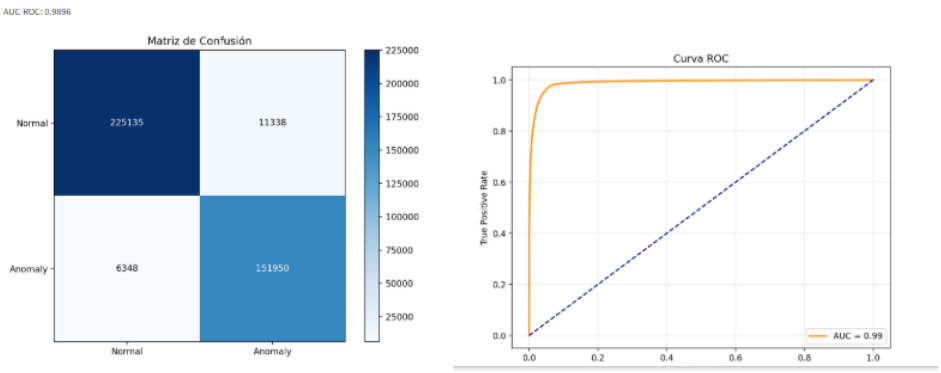


Ilustración 40:Gráfica 1 final dashboard

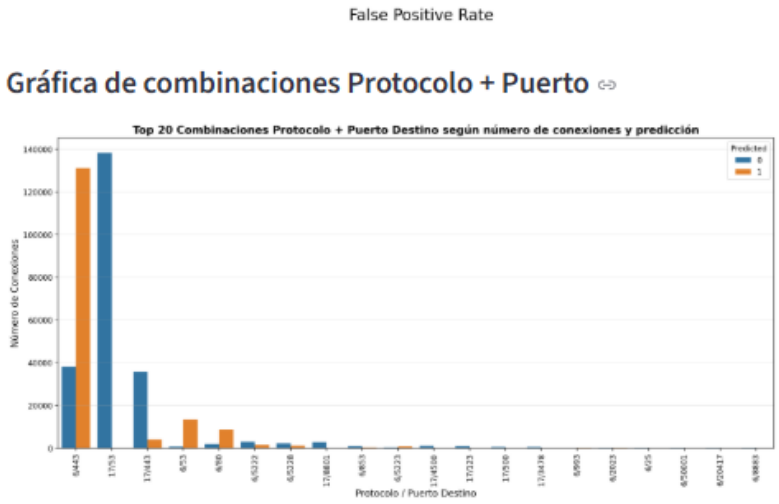


Ilustración 41:Gráfica 2 final dashboard

RESULTADOS

El análisis comparativo de los cuatro algoritmos implementados reveló diferencias significativas en su capacidad para detectar Amenazas Persistentes Avanzadas en el tráfico web. En los resultados de los modelos evaluados, Random Forest destaca como el más efectivo en general, con un AUC de 0.97, indicando una excelente capacidad para distinguir entre clases positivas y negativas. Además, presenta una precisión de 0.85 y un recall de 0.90, aunque genera más falsos positivos (7,403) que otros modelos, clasifica correctamente tanto las observaciones normales como las anomalías. Su F1-Score de 0.87 refleja un buen equilibrio entre precisión y recall, lo que garantiza fiabilidad en términos generales. Por otra parte, SVM obtiene un AUC de 0.944, mostrando también buen rendimiento, pero con una precisión de 0.76 y un recall de 0.97, lo que significa que, aunque identifica bien las anomalías, tiene una mayor tasa de falsos positivos (14,346), representando un reto en aplicaciones que minimizan las falsas alarmas.

K-means presenta un AUC de 0.85 dando a paso una buena distribución, la matriz de confusión indica que clasifica correctamente los casos normales (103,149), pero tiene dificultades para reconocer las anomalías, evidenciado por una alta tasa de falsos negativos (34,811), y un F1-Score de 0.81. Por otro lado, Isolation Forest obtiene el AUC más bajo, de 0.74, lo que sugiere que es el menos efectivo para distinguir entre clases. Su precisión es de 0.62, con un recall bajo de 0.27, lo que

indica problemas para detectar correctamente las anomalías y un F1-Score de 0.38, reflejando un rendimiento deficiente en la clasificación.

Random Forest es el modelo más robusto y equilibrado en términos de AUC, precisión, recall y F1-Score, aunque con el inconveniente de generar muchas falsas alarmas. SVM mantiene un rendimiento destacado, principalmente en recall, aunque con menor precisión. K-means y Isolation Forest son menos efectivos por lo que se recomienda principalmente Random Forest para tareas que necesitan una clasificación precisa y confiable.

Modelo	AUC (ROC)	Accuracy	Precision	Recall	F1-Score
SVM	**0.944**	**0.87**	**0.76**	**0.97**	**0.85**
Random Forest	**0.97**	**0.90**	**0.85**	**0.90**	**0.87**
K-means	**0.85**	**0.83**	**0.73**	**0.91**	**0.81**
Isolation Forest	**0.746**	**0.64**	**0.62**	**0.27**	**0.38**

Tabla 23:Resultados de Modelos

Resultado Final Gráficas Random Forest

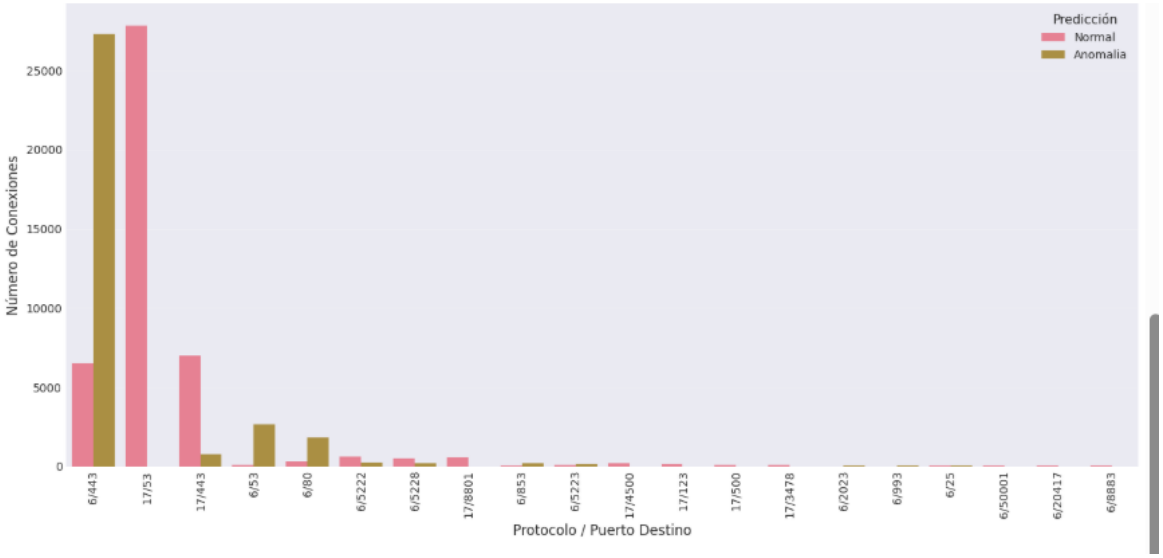


Ilustración 42:Gráfica barras de predicción modelo Random Forest

El gráfico de barras presenta la clasificación de un modelo Random Forest aplicado a datos de prueba del dataset, mostrando la distribución del Número de Conexiones por Protocolo - Puerto Destino y etiquetando cada conexión como Normal (rosa) o Anomalía (marrón). Se analiza que el tráfico más frecuente en puertos web comunes, como 6/443 (HTTPS) y 6/80 (HTTP), fue clasificado principalmente como anomalía, lo que sugiere una alta sensibilidad del modelo o la presencia de actividad anómala significativa en estos servicios durante la prueba. Por otro lado, el tráfico en el puerto 17/53 (DNS), que ocupa el segundo lugar en cantidad, fue mayormente clasificado como Normal, indicando que este flujo se ajusta mejor a los patrones típicos reconocidos por el modelo de detección.

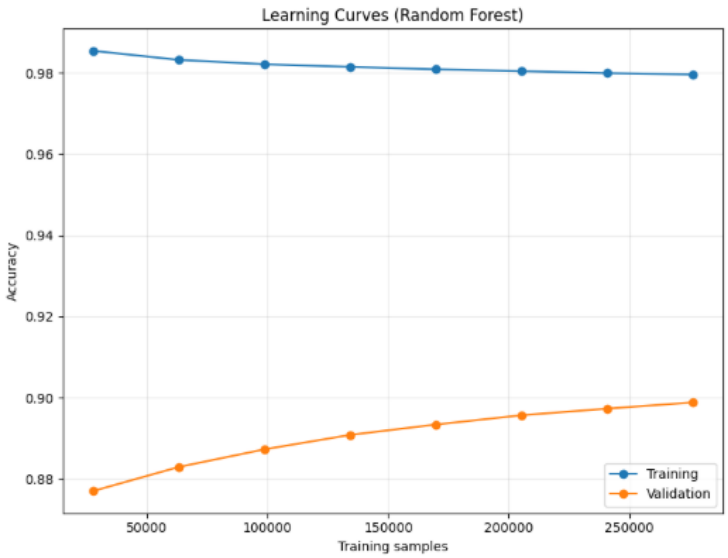


Ilustración 43: Curva de aprendizaje Random Forest

La Ilustración 43 ilustra cómo cambia la precisión del modelo según aumenta el número de muestras de entrenamiento. La línea azul refleja la precisión en el entrenamiento, que se estabiliza al incrementar las muestras, señalando que el modelo alcanza su máxima capacidad de ajuste a los datos. La línea naranja representa la precisión en la validación, que va mejorando progresivamente con más muestras. Esto indica que el modelo mejora su precisión a medida que recibe más datos, aunque ya está acercándose a una convergencia en el proceso de entrenamiento.

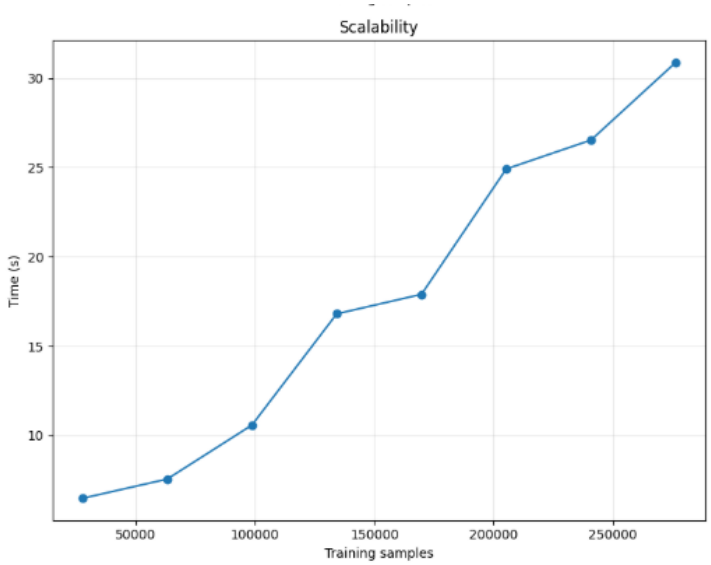


Ilustración 44: Escalabilidad de Random Forest

La Ilustración 44 muestra cómo el tiempo de entrenamiento del modelo Random Forest se incrementa a medida que aumentan las muestras de entrenamiento. Se puede notar que el tiempo crece de forma proporcional al tamaño del conjunto de datos, lo cual es esperado para modelos de Random Forest, ya que requieren la construcción de varios árboles de decisión. A medida que se añaden más datos, el tiempo de cómputo requerido también aumenta, lo que refleja una mayor complejidad computacional y de memoria.

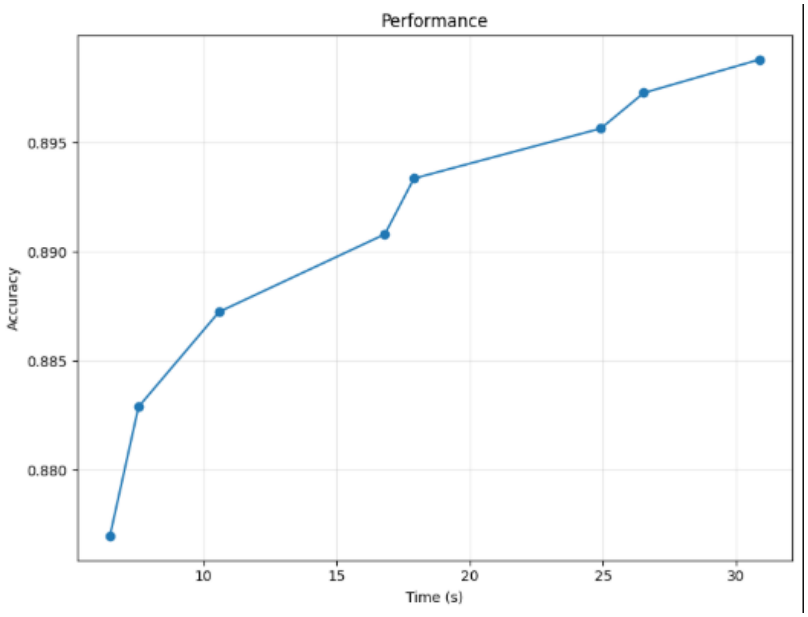


Ilustración 45: Comportamiento de Random Forest

En la Ilustración 45, se observa la relación entre el tiempo de entrenamiento y la precisión en la validación. A medida que aumenta el tiempo de entrenamiento (es decir, el número de muestras de entrenamiento), la precisión también aumenta de manera progresiva. Esto sugiere que el modelo mejora su capacidad de generalización a medida que se expone a más datos, alcanzando una precisión de aproximadamente 0.895 al final del gráfico.

Resultado Final Gráficas K-means

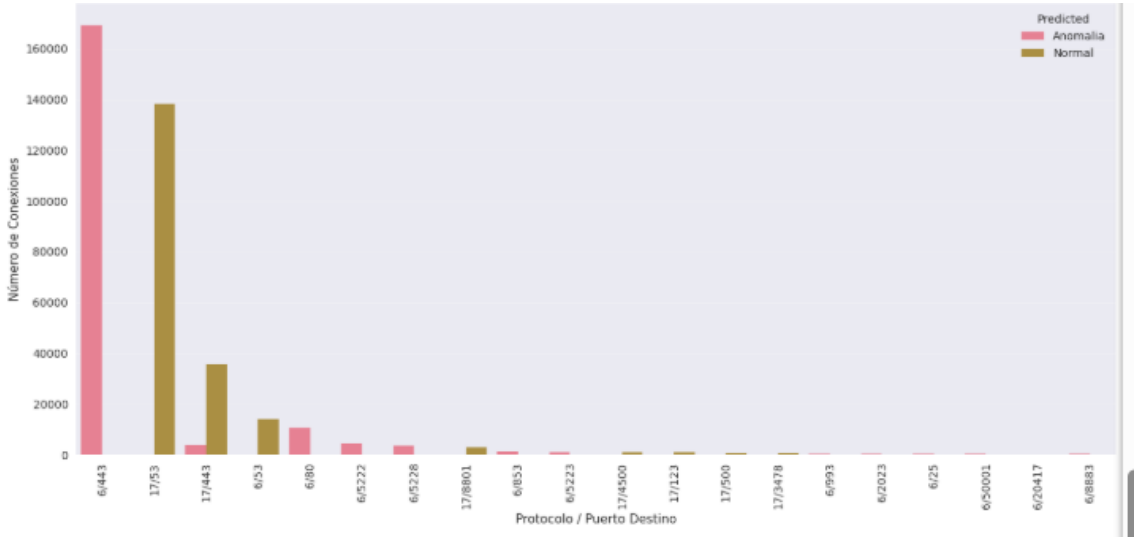


Ilustración 46:Gráfica barras de predicción modelo K-means

El gráfico de K-Means, un modelo no supervisado, muestra una clasificación de clusters más polarizada respecto al Random Forest, donde la mayoría de las conexiones en el puerto 6/443 (HTTPS) se agruparon en el cluster etiquetado como Anomalía (rosa), sugiriendo que este tráfico difiere estructuralmente del resto en cuanto a las características de clustering. En cambio, casi todas las conexiones 17/53 (DNS), junto con el tráfico de 17/443 y 6/80 (que el Random Forest había clasificado como anómalas), quedaron en el cluster etiquetado como Normal (marrón).

Esto indica que K-Means detecta patrones de agrupamiento basados en similitudes estructurales de los datos más que en etiquetas de clase previamente definidas, lo que puede llevar a discrepancias con modelos supervisados como Random Forest.

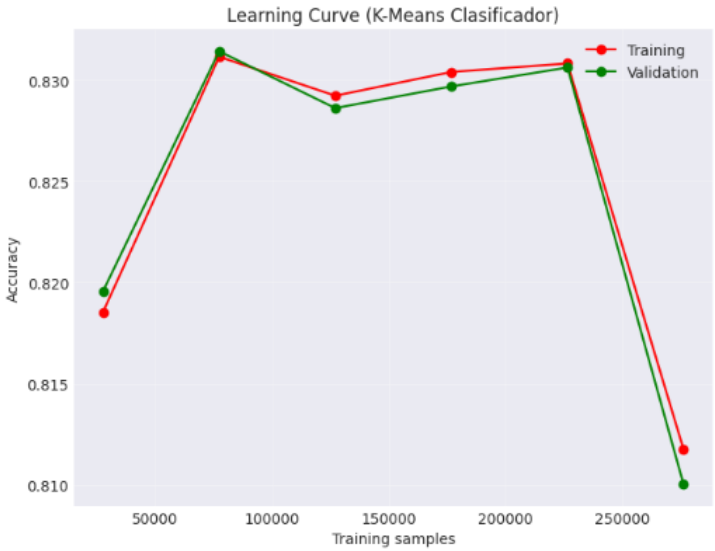


Ilustración 47: Curva de aprendizaje K-means

La Ilustración 47 muestra cómo la precisión del modelo K-Means varía con el incremento de las muestras de entrenamiento. La línea roja muestra la precisión en el conjunto de entrenamiento, mientras que la verde indica la precisión en el conjunto de validación. Con el aumento en el número de muestras las dos precisiones alcanzan una estabilidad cercana a 0.82, aunque presentan algunas fluctuaciones. Estas variaciones en la precisión de validación sugieren que el modelo podría estar sobreajustando en ciertos tamaños de datos. Aunque la precisión se estabiliza, el rendimiento no mejora notablemente con un aumento en el número de muestras, lo que sugiere que K-Means posiblemente no está explotando al máximo el potencial de los datos grandes.

En esta **¡Error! No se encuentra el origen de la referencia.**, el tiempo de e ntrenamiento se presenta según la cantidad de muestras utilizadas. La línea azul muestra que el ajuste tarda más con un conjunto reducido de muestras (50,000), pero luego se estabiliza en torno a 1.8-2 segundos al aumentar las muestras. A medida que se añaden más datos, el tiempo de ajuste no crece de forma lineal, lo que podría señalar que el modelo se ajusta rápidamente tras cierto número de muestras.

Esto sugiere que el modelo alcanza un punto de saturación en el que la incorporación de datos adicionales aporta un beneficio marginal al entrenamiento, optimizando así el uso de recursos computacionales.

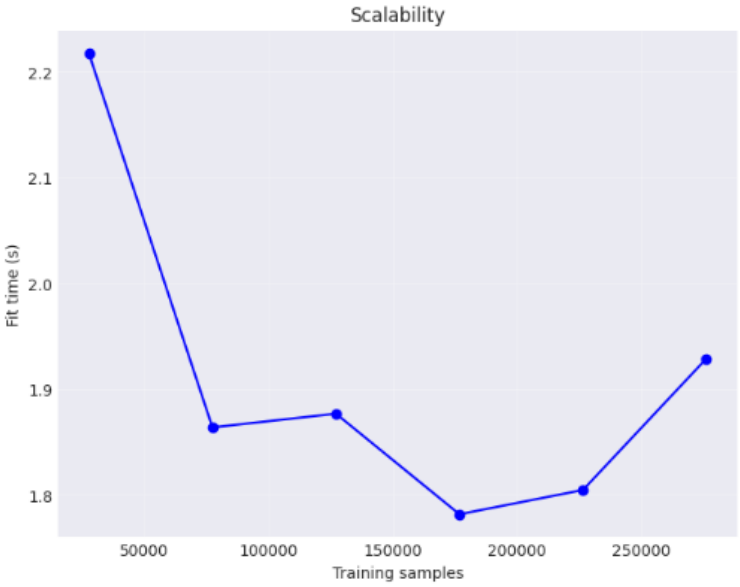


Ilustración 48: Escalabilidad de K-means

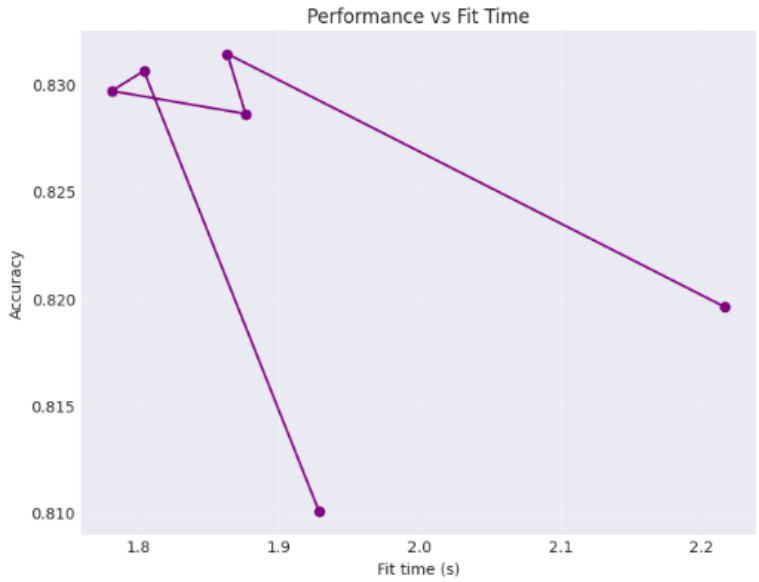


Ilustración 49: Comportamiento de K-means

En la Ilustración 49, se puede ver que la relación entre el tiempo de ajuste (Fit time) y la precisión (Accuracy) muestra una variabilidad significativa. Aunque en algunos puntos la precisión alcanza valores cercanos a 0.830, también hay fluctuaciones a medida que aumenta el tiempo de ajuste. Esto indica que el tiempo de ajuste no sigue una tendencia lineal en cuanto a la mejora de la precisión del modelo. Las fluctuaciones podrían deberse a la convergencia de los centroides en el algoritmo K-means, lo que sugiere que un mayor tiempo de ajuste no garantiza

necesariamente mayor precisión, sino que puede estar influido por otros factores, como la inicialización de los centroides o el número de iteraciones. Estos hallazgos resaltan la importancia de optimizar tanto el tiempo de ajuste como los parámetros del algoritmo para mejorar su rendimiento.

Resultado Final Gráficas SVM

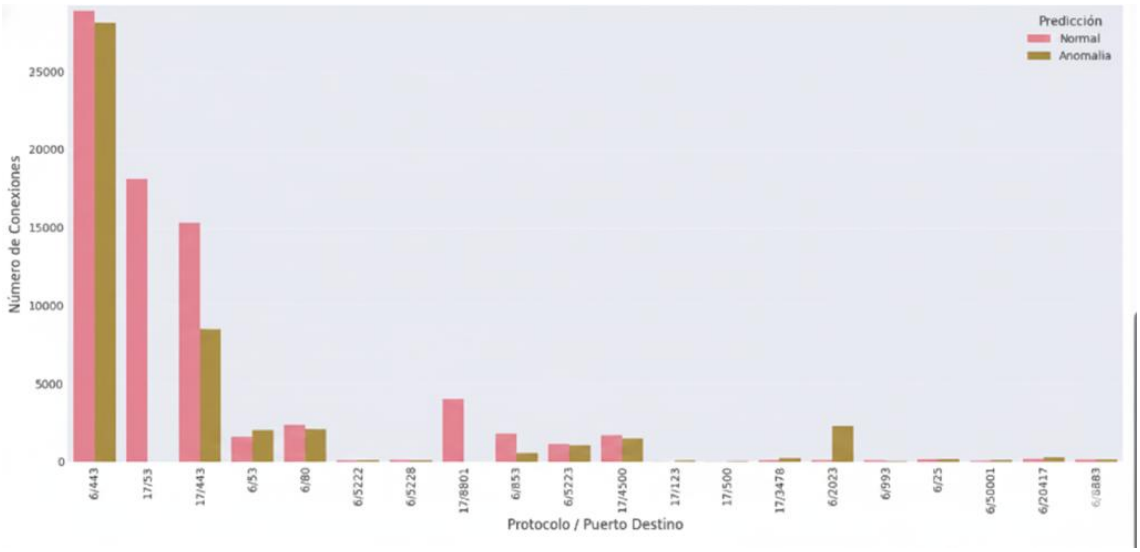


Ilustración 50:Gráfica barras de predicción modelo SVM

El modelo de Máquinas de Vectores de Soporte (SVM), como clasificador supervisado, introdujo una perspectiva más equilibrada en la detección de anomalías en comparación con Random Forest y K-Means, aunque los tres coincidieron en clasificar el tráfico 17/53 (DNS) como predominantemente Normal. En el tráfico 6/443 (HTTPS), el SVM mostró una distribución casi equitativa entre Normal y Anomalía, en marcado contraste con la fuerte tendencia anómala vista en RF y KM, indicando una menor propensión a marcar el tráfico web de alto volumen como atípico. No obstante, el SVM mantuvo la tendencia del RF de clasificar gran parte del tráfico 17/443 como Anomalía, mientras que se acercó al KM al clasificar el puerto 6/80 como ligeramente más Normal. Finalmente, el SVM destacó por etiquetar el tráfico del puerto de bajo volumen 6/2023 casi en su totalidad como Anomalía, sugiriendo una sensibilidad única a patrones específicos en puertos menos comunes.

El modelo SVM logró una precisión de 0.94 demostrando su eficacia para identificar patrones y realizar predicciones precisas, este alto nivel de rendimiento implica un elevado costo computacional, particularmente al trabajar con grandes cantidades de datos. Debido a la complejidad y al tiempo requerido para el procesamiento, se decidió usar una GPU para acelerar el entrenamiento y disminuir la carga computacional.

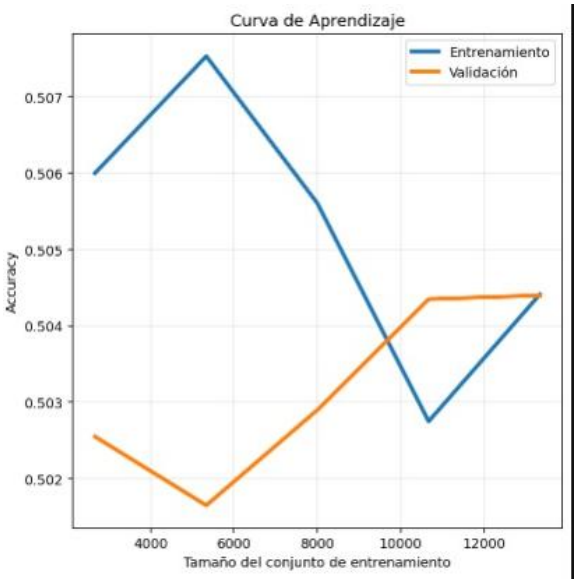


Ilustración 51:Curva de Aprendizaje SVM

La primera Ilustración 51, conocida como la Curva de Aprendizaje, ilustra cómo la precisión en el conjunto de validación incrementa con el aumento del tamaño del conjunto de entrenamiento. Al principio, el modelo se beneficia de contar con más datos, pero luego la mejora se estabiliza, lo que indica que ha aprendido la mayoría de los patrones relevantes en los datos. Esta estabilización sugiere que, aunque los resultados son buenos, agregar más muestras no mejora significativamente la precisión y que el tiempo de entrenamiento puede ser una restricción.

Por lo tanto, es importante balancear la cantidad de datos utilizados con los recursos computacionales disponibles para evitar incrementos innecesarios en el tiempo de procesamiento. Además, esta observación indica que el modelo ha capturado la mayoría de los patrones relevantes presentes en el conjunto de datos actual.

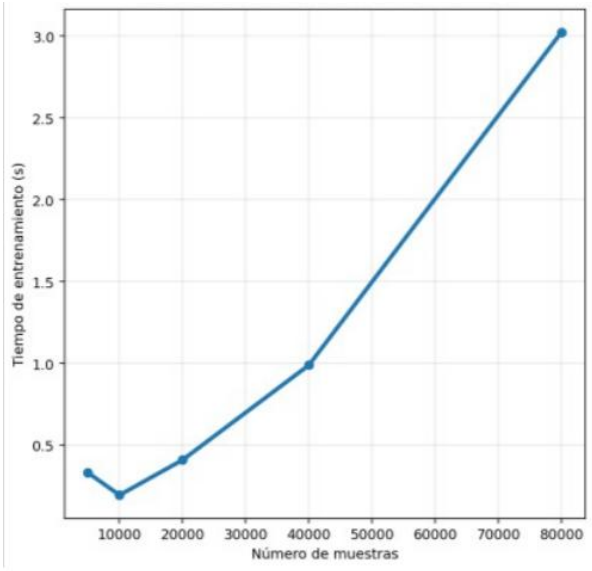


Ilustración 52:Escalabilidad SVM

En la Ilustración 52, se aprecia cómo el tiempo de entrenamiento crece en proporción al incremento del número de muestras

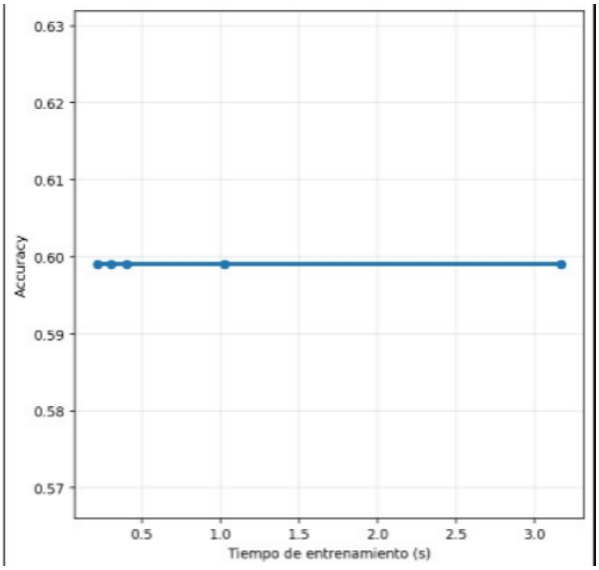


Ilustración 53:Comportamiento SVM

Finalmente, en la Ilustración 53 presenta cómo varía el tiempo de entrenamiento en relación con la precisión lograda aunque el tiempo de entrenamiento aumenta al agregar más muestras, la precisión se mantiene estable tras llegar a un número suficiente de datos. Esto señala que introducir datos adicionales no garantiza una mejora en el rendimiento, aunque sí aumenta el tiempo de procesamiento.

Resultado Final Gráficas Isolation Forest

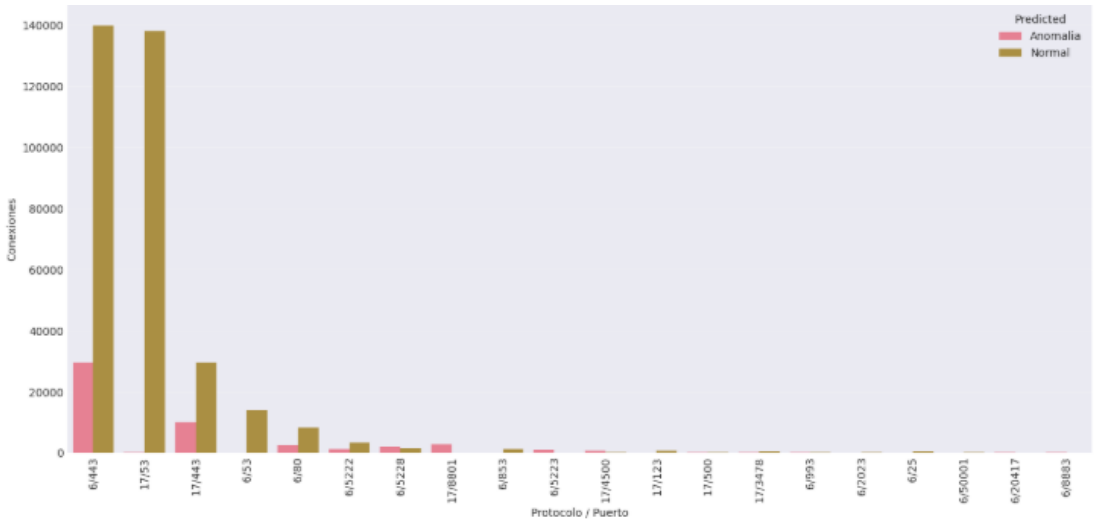


Ilustración 54:Gráfica barras de predicción modelo Isolation Forest

El modelo Isolation Forest, el último de los algoritmos probados, mostró el mayor contraste con los modelos supervisados (RF y SVM) al clasificar la inmensa mayoría del tráfico de alto volumen en el puerto 6/443 (HTTPS) como Normal, invirtiendo la tendencia anómala fuerte o total detectada por los otros clasificadores. Coincidió con los modelos anteriores en que el tráfico 17/53 (DNS) es abrumadoramente Normal, lo que establece una consistencia en la caracterización de este tráfico. Para los demás puertos web (17/443 y 6/80), Isolation Forest también favoreció la etiqueta Normal, lo que sugiere que, desde una perspectiva de aislamiento (es decir, qué tan raros son los datos), la mayoría de los flujos de conexión en el conjunto de prueba no son lo suficientemente aislados o atípicos como para ser marcados como anomalía.

Esto refuerza la idea de que el modelo es conservador al identificar anomalías, priorizando la reducción de falsos positivos. Sin embargo, también indica que podría pasar por alto patrones sutiles de comportamiento anómalo si estos no se destacan lo suficiente frente al resto de los datos. Por lo tanto, la combinación de Isolation Forest con otros métodos supervisados podría mejorar la detección de casos más complejos. Además, resalta la importancia de un preprocesamiento

cuidadoso y la selección de características relevantes que acentúen las diferencias entre tráfico normal y anómalo.

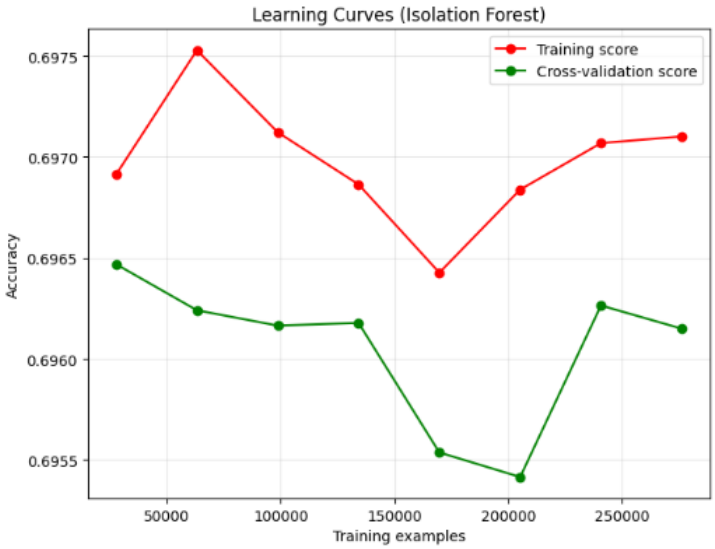


Ilustración 55: Curva de aprendizaje de Isolation Forest

La Ilustración 55 nos dice cómo la precisión del modelo Isolation Forest varía según el número de muestras de entrenamiento. A medida que aumenta el número de muestras, las dos precisiones tienden a estabilizarse, aunque la precisión en entrenamiento es ligeramente superior a la de validación. Esto indica un posible ligero sobreajuste del modelo a los datos de entrenamiento y una capacidad limitada para generalizar a datos nuevos.

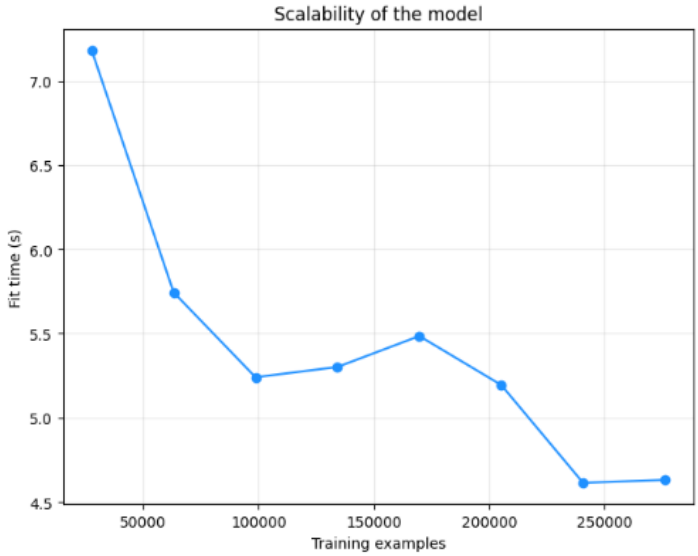


Ilustración 56: Escalabilidad de Isolation Forest

Esta Ilustración 56 muestra el tiempo que tarda en entrenarse el modelo Isolation Forest con diferentes tamaños de muestra. La línea azul muestra que el tiempo de ajuste disminuye significativamente al incrementar el número de muestras, estabilizándose después de 150,000 muestras. Esto muestra que Isolation Forest es altamente escalable, ya que el tiempo de entrenamiento no aumenta de manera lineal con más datos, sino que se vuelve más eficiente a medida que crecen datos.

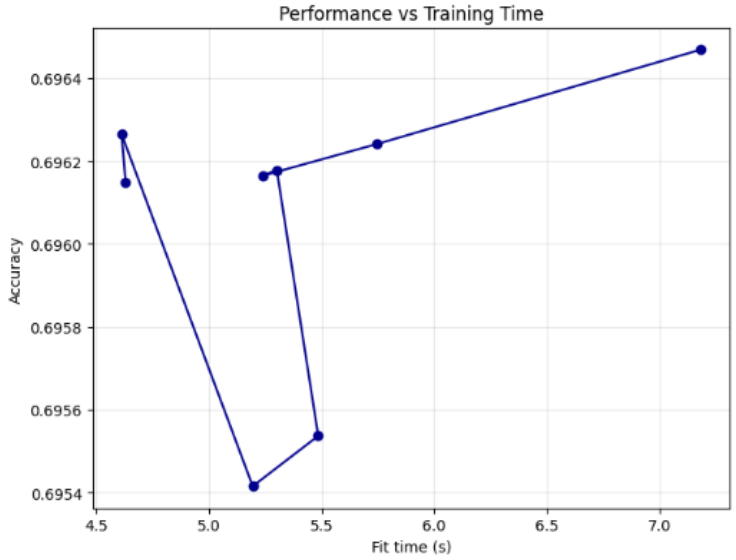


Ilustración 57: Comportamiento de Isolation Forest

La Ilustración 57 se compara la precisión en relación con el tiempo de entrenamiento. La precisión muestra una ligera mejora a medida que aumenta el tiempo de ajuste, aunque se estabiliza cerca de 0.696 después de un cierto punto. Esto sugiere que, aunque el modelo continúa mejorando ligeramente con más tiempo de entrenamiento, la precisión no aumenta de manera significativa con el incremento del tiempo, lo que indica que Isolation Forest tiene una capacidad limitada para mejorar con mayor tiempo de entrenamiento.

Por lo tanto, prolongar el entrenamiento más allá de cierto punto puede no justificar el costo computacional adicional. Esto resalta la necesidad de enfocarse en la optimización de parámetros, la selección de características relevantes y la calidad de los datos, más que en la mera extensión del tiempo de entrenamiento. Además, subraya que la eficiencia y la efectividad del modelo dependen en gran medida de la naturaleza intrínseca de los datos y del patrón de anomalías presentes.

CONCLUSIONES

Se aplicaron y evaluaron exitosamente cuatro modelos predictivos (SVM, Random Forest, K-means e Isolation Forest) capaces de identificar comportamientos anómalos en el tráfico web. La ingeniería de características, que incluyó métricas específicas para patrones APT como beaconing y exfiltración, fue fundamental para la calidad predictiva. Los modelos supervisados demostraron una capacidad superior para la clasificación precisa en comparación con los no supervisados en este contexto de datos de red.

La implementación de técnicas de análisis exploratorio de datos (EDA) y minería de datos facilitó la identificación y cuantificación de patrones anómalos clave en el conjunto de datos, confirmando comportamientos sospechosos como tráfico cifrado masivo en el puerto 443, actividades de beaconing y sesiones prolongadas. El análisis correlacional y la generación de características derivadas fueron esenciales para convertir datos brutos en indicadores de APTs útiles para los modelos.

Se diseñó e implementó un dashboard que presenta de manera clara y consolidada los resultados del análisis, incluyendo métricas de rendimiento de los modelos, distribución de tráfico normal vs. anómalo y gráfica por puertos y protocolos. Este dashboard demostró la ventaja de la técnica desarrollada, traduciendo la salida compleja de los modelos en información accionable.

RECOMENDACIONES

Se recomienda adoptar e implementar en producción el modelo Random Forest por su equilibrio óptimo entre precisión (90%), recall (90%) y robustez , lo que minimiza tanto falsos negativos como falsos positivos. Además, se sugiere institucionalizar un pipeline de preprocesamiento y ingeniería de características que actualice y refine continuamente las métricas basadas en nuevas tácticas APT documentadas en frameworks como MITRE ATT&CK.

Se recomienda automatizar y programar la ejecución periódica del análisis exploratorio (EDA) sobre nuevos volúmenes de logs de red. Esto permitirá monitorear tendencias emergentes, validar la vigencia de las características creadas y descubrir nuevos patrones anómalos no contemplados inicialmente, manteniendo el sistema de detección adaptativo ante la evolución de las amenazas.

Se recomienda mejorar le dashboard para poder integrarlo con los sistemas de monitoreo de red existentes en la UPSE para que las alertas generadas por los modelos se reflejen en tiempo real. Para lograrlo, se propone desarrollar una API REST que exponga las predicciones del modelo y permita la ingestión automatizada de alertas, se sugiere incorporar en el dashboard un módulo de retroalimentación que permita a los administradores marcar falsos positivos/negativos, optimizando así el modelo de forma continua y adaptativa.

BIBLOGRAFIA

[1] N. J. Vallejo Ayala, «La Ciencia Y La Tecnología En La Sociedad Contemporánea, Una Perspectiva Desde La Biótica,» *Universidad Militar Nueva Granada,* Pp. 156-158, 2019.

[2] C. A. Loor, K. Morocho Y M. Hallo, «Uso De Técnicas De Minería De Datos Para La Detección De Ataques De Inyección De Sql En Sistemas De Bases De Datos,» *Revista Politécnica,* Vol. 51, 2023.

[3] Upse, «Upse,» Upse, 26 07 2018. [En Línea]. Available: Https://Www.Upse.Edu.Ec/Index.Php?Option=Com_Content&View=Article&Id=10&Itemid=188. [Último Acceso: 08 2025].

[4] W. O. Montaña Sierra Y W. Daza Tibocha , *Diseño E Implementación De Un Sistema De Detección De Anomalías De Red Mediante El Análisis Avanzado De Logs Utilizando Software Libre En Un Ambiente De Laboratorio Que Permita La Optimizacion De Recursos Tecnologicos,* Bogotá : Universidad Piloto De Colombia Especialización En Seguridad Informática , 2018.

[5] N. Ilzam, N. Jamil Y M. L. Yunus Yusoff, «A Systematic Literature Review On Advanced Persistent Threat Behaviors And Its Detection Strategy,» *Insitute Of Informatics And Computing In Energy,* Pp. 1-18, 2023.

[6] J. E. Velasco López , *Uso De Técnicas De Web Mining: Aplicación Empírica En El Sector De La Administración Pública,* Madrid: Universidad Complutense De Madrid, 2013.

[7] J. L. Say Valdez , *Minería Web Como Herramienta De Análisis De Ficheros Log En Servidores Web,* Guatemala: Universidad De San Carlos De Guatemala , 2010.

[8] Santiago Quintero B., «Nuevas Perspectivas En El Estudio De Amenazas Persistentes Avanzadas,» Universidad De Salamanca - Departamento De Matemática Aplicada, España , 2021.

[9] . S. Quintero Bonilla, *Nuevas Perspectivas En El Estudio De Amenazas Persistentes Avanzadas,* Salamanca: Universidad De Salamanca, 2021.

[10] D. Carrillo Rico Y R. Merchan Patarroyo , *Fortalecimiento Del Esquema De Defensa En Profundidad En La Anh Para Incrementar El Nivel De Protección Frente A Las Amenazas Persistentes Avanzadas.,* Bogotá: Universidad Piloto De Colombia , 2015.

[11] El Comercio , «Ecuador Denuncia 40 Millones De Ciberataques Tras Retiro De Asilo A Assange,» 15 Abril 2019. [En Línea]. Available: Vhttps://Www.Elcomercio.Com/Actualidad/Seguridad/Ecuador-Denuncia-Millones-Ciberataques-Assange/. [Último Acceso: 10 Octubre 2025].

[12] Joseph Guamán S., «Seguridad De La Información En Las Fuerzas Armadas Del Ecuador Para La Predicción De Ciberataques En Redes Militares,» Universidad De Guadalajara, Zapopan - Jalisco , 2024.

[13] P.-N. S. M. &. K. V. Tan, «Introduction To Data Mining,» *Pearson Education,* 2006.

[14] C. Vasant, A. Banerjee Y V. Kumar, «Anomaly Detection: A Survey,» *Acm Computing Surveys,* Vol. 41, Nº 3, Pp. 1-58, 2009.

[15] L. M. Guillermo, «Plan-De-Creación-De-Oportunidades-2021-2025,» Secretaria Nacional De Planificacion, 2021. [En Línea]. Available: Https://Www.Planificacion.Gob.Ec/Wp-Content/Uploads/2021/09/Plan-De-Creacio%Cc%81n-De-Oportunidades-2021-2025-Aprobado.Pdf. [Último Acceso: 2025].

[16] E. Rus Arias, «Investigación Exploratoria,» Economipedia, 1 11 2020. [En Línea]. Available: Https://Economipedia.Com/Definiciones/Investigacion-Exploratoria.Html. [Último Acceso: 19 08 2025].

[17] X. A Meléndez De Cooper , *Proceso De Investigación-Diagnostica Y Elaboración De Informes Sociales,* San Salvador: Ciudad Universitaria "Dr Fabio Castillo Figueroa", 2022.

[18] J. Mejía, «Crisp-Dm Frente A La Revolución De La Ia: ¿Evolución O Extinción?,» Linkedin, 07 2025. [En Línea]. Available: Https://Www.Linkedin.Com/Pulse/Impacto-De-Crispdm-En-Proyectos-Ia-Y-Su-Vigencia-Frente-Jes%C3%Bas-Mej%C3%Ada-70wif/. [Último Acceso: 2025].

[19] Upse, «Reseña Histórica De La Creación De La Universidad,» Upse, 2018. [En Línea]. Available: Https://Www.Upse.Edu.Ec/Index.Php?Option=Com_Content&View=Article&Id=10 &Itemid=166 .

[20] Upse, «Misión-Visión,» Upse, 2024. [En Línea]. Available: Https://Www.Upse.Edu.Ec/Archivos/Index.Php?Option=Com_Sppagebuilder&View=Page&Id=10&Itemid=183. [Último Acceso: 2025].

[21] Kaspersky, «¿Qué Es Una Amenaza Avanzada Persistente (Apt)?,» 2023. [En Línea]. Available: Https://Latam.Kaspersky.Com/Resource-Center/Definitions/Advanced-Persistent-Threats. [Último Acceso: 09 Septiembre 2025].

[22] Ibm, «¿Qué Es La Minería De Datos?,» 28 Junio 2024. [En Línea]. Available: Https://Www.Ibm.Com/Es-Es/Topics/Data-Mining. [Último Acceso: 07 Septiembre 2025].

[23] R. D. Ecuador, «Constitución De La República Del Ecuador,» Lexis, 2024. [En Línea]. [Último Acceso: 2025].

[24] L. Guillermo, Ley Orgánica De Protección De Datos Personales, Ecuador, 2023.

[25] Coip, «Código Orgánico Integral Penal (Coip),» Informática Jurídica, 2014. [En Línea]. Available: Https://Www.Informatica-Juridica.Com/Codigo/Codigo-Organico-Integral-Penalcoip-10-Agosto-Del-2014/ . [Último Acceso: 2025].

[26] Dennis Juilland & Miguel Osorio, «Desarrollo De Modelos De Inteligencia Artificial Para La Detección De Phishing En Encabezados De Correos Electrónicos, Contribuidos Por Grupos De Amenazas Persistentes Avanzadas (Atp),» Universidad Eia, Envigado, 2024.

[27] Rosas Paredes K., «Análisis Exploratorio De Ataques Informáticos Aplicando Herramientas De Minería De Datos Para La Gestión De La Seguridad De Redes Inalámbricas En Universidades De Arequipa,» Escuela Universitaria De Posgrado Federico Villareal, Lima - Perú, 2020.

[28] Gregg Lindemulder & Amber Forrest , «¿Qué Son Las Amenazas Persistentes Avanzadas?,» Ibm, [En Línea]. Available: Https://Www.Ibm.Com/Es-Es/Think/Topics/Advanced-Persistent-Threats. [Último Acceso: 30 Noviembre 2025].

[29] Microsoft, «What Is An Advanced Persistent Threat (Apt)?,» Microsoft, [En Línea]. Available: Https://Www.Microsoft.Com/En-Us/Security/Business/Security-101/What-Is-Advanced-Persistent-Threat-Apt?Utm_Source=Chatgpt.Com. [Último Acceso: 30 Noviembre 2025].

[30] Aws, «¿Por Qué Es Importante La Detección De Anomalías?,» Aws, [En Línea]. Available: Https://Aws.Amazon.Com/Es/What-Is/Anomaly-Detection/. [Último Acceso: 30 Noviembre 2025].

[31] Ibm, «Detección Basada En Firmas,» Ibm, [En Línea]. Available: Https://Www.Ibm.Com/Mx-Es/Think/Topics/Intrusion-Prevention-System. [Último Acceso: 30 Noviembre 2025].

[32] Welivesecurity, «Heurística Antivirus Y La Detección Proactiva De Amenazas,» Welivesecurity, 18 Marzo 2013. [En Línea]. Available: Https://Www.Welivesecurity.Com/La-Es/2013/03/18/Heuristica-Antivirus-Deteccion-Proactiva-Amenazas/. [Último Acceso: 30 Noviembre 2025].

[33] Jim Holdsworth & Mark Scapicchio, «¿Qué Es El Deep Learning?,» Ibm, [En Línea]. Available: Https://Www.Ibm.Com/Es-Es/Think/Topics/Deep-Learning. [Último Acceso: 30 Noviembre 2025].

[34] Josh Fruhlinger, «Whitelisting Explained: How It Works And Where It Fits In A Security Program,» 07 Junio 2024. [En Línea]. Available: Https://Www.Csoonline.Com/Article/569493/Whitelisting-Explained-How-It-Works-And-Where-It-Fits-In-A-Security-Program.Html. [Último Acceso: 30 Noviembre 2025].

[35] Mozilla Developer Network (Mdn), «What Is A Web Page?,» Mozilla Developer Network (Mdn), [En Línea]. Available: Https://Developer.Mozilla.Org/En-Us/Docs/Learn/Getting_Started_With_The_Web/What_Is_A_Web_Page.

[36] Mdn Web Docs, «¿Qué Es La Seguridad De Sitios Web?,» 27 Marzo 2025. [En Línea]. Available: Https://Developer.Mozilla.Org/Es/Docs/Learn_Web_Development/Extensio

ns/Server-Side/First_Steps/Website_Security. [Último Acceso: 08 Septiembre 2025].

[37] Cisco, «What Is Network Traffic Analysis?,» [En Línea]. Available: Https://Www.Cisco.Com/Site/Us/En/Learn/Topics/Security/What-Is-Network-Traffic-Analysis.Html. [Último Acceso: 30 Noviembre 2025].

[38] Ibm, «What Is Network Traffic Analysis?,» [En Línea]. Available: What Is Network Traffic Analysis?. [Último Acceso: 30 Noviembre 2025].

[39] Cisco, «What Are Netflow And Sflow Protocols?,» [En Línea]. Available: Https://Www.Cisco.Com/C/En/Us/Td/Docs/Iosxr/Cisco8000/Netflow/Configuration/B-Netflow-Configuration-Ios-Xr-8000/Netflow-Sflow-Key-Concepts.Pdf. [Último Acceso: 30 Noviembre 2025].

[40] Linkedin, «¿Cómo Se Pueden Utilizar Los Analizadores De Paquetes Para Solucionar Problemas De Red?,» [En Línea]. Available: Https://Es.Linkedin.Com/Advice/1/How-Can-You-Use-Packet-Analyzers-Troubleshoot?Lang=Es&Lang=Es. [Último Acceso: 30 Noviembre 2025].

[41] Ankush Thakur, «12 Analizadores De Paquetes De Red Para Administradores De Sistemas Y Analistas De Seguridad,» Geekflare, 12 Septiembre 2024. [En Línea]. Available: Https://Geekflare.Com/Es/Network-Packet-Analyzers/. [Último Acceso: 30 Noviembre 2025].

[42] Ahmed, M., Mahmood, A. N., & Hu, J, «A Survey Of Network Anomaly Detection Techniques.,» Journal Of Network And Computer Applications, 2018.

[43] Karspersky, «¿Qué Es Una Ip?,» [En Línea]. Available: Https://Latam.Kaspersky.Com/Resource-Center/Definitions/What-Is-An-Ip-Address. [Último Acceso: 30 Noviembre 2025].

[44] Amazon Web Services, «¿Qué Es La Minería De Datos?,» 2023. [En Línea]. Available: Https://Aws.Amazon.Com/Es/What-Is/Data-Mining/. [Último Acceso: 08 Septiembre 2025].

[45] Sap, «¿Por Qué Utilizar La Minería De Datos?,» 2023. [En Línea]. Available: Https://Www.Sap.Com/Spain/Products/Data-Cloud/Hana/What-Is-Data-Mining.Html. [Último Acceso: 08 Septiembre 2025].

[46] Guaña Moya, J., Álvarez-Carpio, G., Briones-Montalvo, C., Ortiz-Terán, I., & Moya Carrera, B. H., «Minería De Datos Aplicada Al Tráfico De Red De Aplicaciones Android Para Detectar Actividades Maliciosas,» *Ingeniería E Innovación Del Futuro,* Vol. 4, Nº 1, Pp. 160-176, 2025.

[47] «¿Qué Es Machine Learning?,» Ibm, [En Línea]. Available: Https://Www.Ibm.Com/Mx-Es/Think/Topics/Machine-Learning. [Último Acceso: 2025].

[48] Ivan Belcic & Cole Stryker, «¿Qué Es El Aprendizaje Supervisado?,» [En Línea]. Available: Https://Www.Ibm.Com/Es-Es/Think/Topics/Supervised-Learning. [Último Acceso: 30 Noviembre 2025].

[49] Google Cloud, «¿Qué Es El Aprendizaje No Supervisado?,» [En Línea]. Available: Https://Cloud.Google.Com/Discover/What-Is-Unsupervised-Learning?Hl=Es-419. [Último Acceso: 30 Noviembre 2025].

[50] Sap, «¿Qué Es La Minería De Datos?,» 2024. [En Línea]. Available: Https://Www.Sap.Com/Spain/Products/Data-Cloud/Hana/What-Is-Data-Mining.Html. [Último Acceso: 15 Septiembre 2025].

[51] Ia Blog, «Machine Learning Vs. Deep Learning: Diferencias Clave Y Aplicaciones,» Ia Blog, [En Línea]. Available: Https://Iartifi Cial.Blog/Aprendizaje/Machine-Learning-Vs-Deep-Learning/. [Último Acceso: 2025].

[52] Yingian Z. & Marten D., «Ccsw '21: Proceedings Of The 2021 On Cloud Computing Security Workshop,» Association For Computing Machinery, New York - United States, 2021.

[53] Han, J., Kamber, M., & Pei, J., «Data Mining: Concepts And Techniques (3rd Ed.). Morgan Kaufmann.,» 2011.

[54] Probyto Ia, «Explorando El Crisp-Dm: Importancia, Fases Y Consideraciones Para La Toma De Decisiones Basada En Datos,» Linkedin, 18 Enero 2024. [En Línea]. Available: Https://Www.Linkedin.Com/Pulse/Exploring-Crisp-Dm-Significance-Phases-Considerations-Data-Driven-Qyioc. [Último Acceso: 30 Noviembre 2025].

[55] Oracle , «Oracle - Oci,» ¿Qué Es Una Base De Datos?, 24 Noviembre 2020. [En Línea]. Available: Https://Www.Oracle.Com/Latam/Database/What-Is-Database/. [Último Acceso: 01 10 2025].

[56] Aws, «Https://Aws.Amazon.Com/Es/What-Is/Python/,» [En Línea]. Available: Https://Aws.Amazon.Com/Es/What-Is/Python/. [Último Acceso: 25 11 2025].

[57] «Tutorial De Jupyter Notebook,» [En Línea]. Available: Https://Www.Programaenpython.Com/Miscelanea/Tutorial-De-Jupyter-Notebook/. [Último Acceso: 25 Noviembre 2025].

[58] Bryan Clark, «¿Qué Es Scikit-Learn (Sklearn)?,» [En Línea]. Available: Https://Www.Ibm.Com/Think/Topics/Scikit-Learn. [Último Acceso: 25 Noviembre 2025].

[59] IBM, «Visualize Data With The Matplotlib Library,» [En Línea]. Available: Https://Dataplatform.Cloud.Ibm.Com/Exchange/Public/Entry/View/55f7f605960d9a4d578aeeaa2b62bdea. [Último Acceso: 25 Noviembre 2025].

[60] Ibm, «Rstudio,» 23 Octubre 2025. [En Línea]. Available: Https://Dataplatform.Cloud.Ibm.Com/Docs/Content/Wsj/Analyze-Data/Rstudio-Overview.Html?Context=Wx&Locale=Es. [Último Acceso: 30 Noviembre 2025].

[61] Microsfot , «¿Qué Es Excel?,» [En Línea]. Available: Https://Support.Microsoft.Com/Es-Es/Office/-Qu%C3%A9-Es-Excel-94b00f50-5896-479c-B0c5-Ff74603b35a3. [Último Acceso: 30 Noviembre 2025].

[62] Microsoft , «Crear O Editar Archivos .Csv Para Importarlos A Outlook,» [En Línea]. Available: Https://Support.Microsoft.Com/Es-Es/Office/Crear-O-Editar-Archivos-Csv-Para-Importarlos-A-Outlook-4518d70d-8fe9-46ad-94fa-1494247193c7#:~:Text=Un%20archivo%20CSV%20(Valores%20separados,De%20un%20programa%20a%20otro.. [Último Acceso: 30 Noviembre 2025].

[63] S. Few, «Information Dashboard Design: The Effective Visual Communication Of Data,» *O'reilly Media,* 2006.

[64] D. S. Bolivia, Artist, *Importancia De Un Dashboard.* [Art]. Data School Bolivia, 2025.

[65] M. Att&Ck, «Application Layer Protocol: Web Protocols,» Mitre Att&Ck, [En Línea]. Available: Https://Attack.Mitre.Org/Techniques/T1071/001/. [Último Acceso: 09 2025].

[66] M. Att&Ck, «Exfiltration Over C2 Channel,» Mitre Att&Ck, [En Línea]. Available: Https://Attack.Mitre.Org/Techniques/T1041/. [Último Acceso: 2025].

[67] M. Att&Ck, «Network Service Discovery,» Mitre Att&Ck, [En Línea]. Available: Https://Attack.Mitre.Org/Techniques/T1046/. [Último Acceso: 2025].

[68] M. Att&Ck, «Remote Services: Remote Desktop Protocol,» Mitre Att&Ck, [En Línea]. Available: Https://Attack.Mitre.Org/Techniques/T1021/001/. [Último Acceso: 2025].

[69] M. Att&Ck, «Exfiltration Over Alternative Protocol: Exfiltration Over Unencrypted Non-C2 Protocol, Sub-Technique T1048.003 - Enterprise,» Mitre Att&Ck, [En Línea]. Available: Https://Attack.Mitre.Org/Techniques/T1048/003/. [Último Acceso: 27 09 2025].

[70] M. Att&Ck, «Application Layer Protocol,» Mitre Att&Ck, [En Línea]. Available: Https://Attack.Mitre.Org/Techniques/T1071/. [Último Acceso: 2025].

[71] G. Betancour, «Las Máquinas De Soporte Vectorial (Svms),» *Scientia Et Technica,* Vol. 11, Nº 27, Pp. 67-72, 2005.

[72] I. Ghafir, M. Hammoudeh Y V. Prenosil, «Detection Of Advanced Persistent Threat Using Machine-Learning Correlation Analysis.,» *Future Generation Computer Systems,* Vol. 89, Pp. 349-359, 2018.

[73] M. I. Abiodun, E.-S. E. Absalón, A. Laith Y A. Belal, «K-Means Clustering Algorithms: A Comprehensive Review, Variants Analysis, And Advances In The Era Of Big Data,» *Information Sciences,* Vol. 622, 2022.

[74] Y. Zhou Y Y. Tang, «A K-Means Clustering Based Intrusion Detection System For Detecting Apts.,» *Journal Of Computational Science,,* Vol. 26, Pp. 142-149, 2018.

[75] D. Quirumbay Yagual , D. Fernandez Iglesias Y F. Novoa, «A Hybrid Deep Learning-Based Architecture For Network Traffic Anomaly Detection Via Efms-Enhanced Kmeans Clustering And Cnn-Gru Models,» *Mdpi,* Vol. 15, 2025.

[76] J. A. Rodrigo, «Detección De Anomalías: Isolation Forest,» Ciencia De Datos, 05 2020. [En Línea]. Available: Https://Cienciadedatos.Net/Documentos/66_Deteccion_Anomalias_Isolationforest. [Último Acceso: 27 11 2025].

[77] L. Fei Tony, T. Kai Ming Y Z. Zhi-Hua, «Isolation Forest,» *Eighth Ieee International Conference On Data Mining,* Pp. 413-422, 2008.

[78] P. Aakash, K. Rakesh Y P. Vatsal, «Reseña De Random Forest: Un Clasificador De Conjuntos,» *Conferencia Internacional Sobre Tecnologías De Comunicación De Datos Inteligentes E Internet De Las Cosas,* Pp. 758-763, 2018.

[79] A. G. Manoel Fernando, W. Xidi Y P. D. L. Alair, «Credit Card Fraud Detection With Artificial Immune System,» *In International Conference On Artificial Immune Systems,* P. 119–131, 2008.

[80] M. Ester, H.-P. Kriegel, J. Sander Y X. Xu, «A Density-Based Algorithm For Discovering Clusters In Large Spatial Databases With Noise,» Pp. 226-231, 1996.

[81] V. Duma Silva, «Clasificación De Tráfico Web Orientadas A La Identificación Oportuna De Ataques Usando Técnicas De Deep Learning,» 2024.

[82] A. Anwar, «Cheat Sheets For Machine Learning And Data Science,» Cheatsheets, [En Línea]. Available:

Https://Sites.Google.Com/View/Datascience-Cheat-Sheets. [Último Acceso: 2025].

[83] M. Walsh, C. Worrell Y T. Scanlon, *Toward Use Of Artificial Intelligence For Advanced Persistent Threat Detection,* Carnegie Mellon University, 2024.

[84] L. Zhang Y Q. Yan, «Detect Malicious Websites By Building A Neural Network To Capture Global And Local Features Of Websites,» *Computers Y Security,* Vol. 137, 2024.

[85] B. Tang, J. Yang, X. Li, Y. Cao Y J. Y Wang, «Apt Detector: Detect And Identify Apt Malware Based On Deep Learning Framework,» *Proceedings Of The 2023 9th International Conference On Computing And Artificial Intelligence,* Pp. 576-583, 2023.

[86] M. M. Hasan, M. U. Islam Y J. Uddin, «Advanced Persistent Threat Identification With Boosting And Explainable Ai.,» *Sn Computer Science,* 2023.

[87] Ibm, «Spss Modeler,» 2011.

[88] Huerta, «Amenazas Persistentes Avanzadas,» *Nau Llibres,* 2016.

[89] B. K. Chandola, «Anomaly Detection: A Survey,» *Acm Computing Surveys,* Pp. 1-58, 2009.

[90] J. Villanueva Morales, J. Lugo Rodríguez, L. Landeros Vázquez, D. Ramírez Buenrostro Y N. Ramírez Pérez, *Aplicación De Algoritmos De Clasificación Para El Análisis De Tejido Mamario Y Detección De Cáncer De Mama,* Madrid: Universidad Carlos Iii De Madrid, 2015.

[91] Ibm, «¿Qué Son Las Amenazas Persistentes Avanzadas?,» 03 Abril 2024. [En Línea]. Available: Https://Www.Ibm.Com/Es-Es/Topics/Advanced-Persistent-Threats. [Último Acceso: 03 Septiembre 2025].

[92] Kaspersky, «Una Amenaza Permanente,» 2024. [En Línea]. Available: Https://Www.Kaspersky.Es/Resource-Center/Definitions/Advanced-Persistent-Threats. [Último Acceso: 08 Septiembre 2025].

[93] Kaspersky Team, «Los Ataques Financieros Crecen En América Latina Y Aumenta La Preocupación Por El Uso De La Piratería,» 17 Noviembre 2022. [En Línea]. Available: Https://Latam.Kaspersky.Com/Blog/Panorama-Amenazas-Latam-2022/25509/. [Último Acceso: 16 Septiembre 2025].

[94] Eset, «Análisis De Tráfico De Red,» [En Línea]. Available: Https://Help.Eset.Com/Eea/10.1/Es-Es/Idh_Config_Epfw_Scan_Main_Page.Html. [Último Acceso: 28 Septiembre 2025].

[95] Ibm, «¿Qué Es El Análisis Dej Tráfico De Red (Nta)?,» 06 Agosto 2025. [En Línea]. Available: Https://Www.Ibm.Com/Es-Es/Think/Topics/Network-Traffic-Analysis. [Último Acceso: 27 Septiembre 2025].

[96] Ibm, «¿Qué Es El Pentesting?,» [En Línea]. Available: Https://Www.Ibm.Com/Mx-Es/Think/Topics/Penetration-Testing. [Último Acceso: 28 Septiembre 2025].

[97] Cloudflare, «¿Cómo Se Lleva A Cabo Una Prueba De Penetración Típica?,» [En Línea]. Available: Https://Www.Cloudflare.Com/Es-Es/Learning/Security/Glossary/What-Is-Penetration-Testing/. [Último Acceso: 28 Septiembre 2025].

[98] O. Suplemento, «Código Orgánico Integral Penal, Coip,» Lexis, 2024. [En Línea]. [Último Acceso: 2025].