Revista Ingenio
UFPSCoña
"Journal of Engineering Sciences"

Original Article DOI: https://doi.org/10.22463/2011642X.5773

Implementación de un sistema de detección de intrusiones basado en Machine Learning para el monitoreo de ciberseguridad en la red institucional de UNIPAZ

Implementation of a Machine Learning-based intrusion detection system for cybersecurity monitoring in the UNIPAZ institutional network

Esp. Jhonis Ríos -Múnera[1], Mg. Karen Jhoanna Salom -Mantilla[2], Esp. Daniel Pose-Villafañe[3], Mg. Nini Del Mar Cortes-Martinez[4], Mg. Paola Niño-Terraza[5]

[1]*Grupo de Investigación En Ciencias e Ingeniería Cipaz , Instituto Universitario de la Paz Unipaz, Colombia, Email: jhoni.rios@Unipaz.edu.co, Orcid: https://orcid.org/0009-0006-1643-2272*
[2]*Grupo de Investigación en Ciencias e Ingeniería Cipaz , Instituto Universitario de la Paz Unipaz, Colombia, Email: karen.salom@unipaz.edu.co, Orcid: https://orcid.org/0009-0005-5483-2018*
[3]*Grupo de Investigación en Ciencias e Ingeniería Cipaz , Instituto Universitario de la Paz Unipaz, Colombia, Email: daniel.pose@unipaz.edu.co, Orcid: https://orcid.org/0009-0002-3417-313X*
[4]*Grupo de Investigación en Ciencias e Ingeniería Cipaz , Instituto Universitario de la Paz Unipaz, Colombia, Email: nini.cortes70@gmail.com, Orcid: https://orcid.org/0009-0004-3236-4295*
[5]*Grupo de Investigación en Ciencias e Ingeniería Cipaz , Instituto Universitario de la Paz Unipaz, Colombia, Email: paola.terraza@unipaz.edu.co, Orcid: https://orcid.org/0009-0001-4677-6950*

Como citar: J. Ríos-Múnera, K.J. Salom-Mantilla, D. Pose-Villafe, N.M. Cortes-Martines y P. Niño, "Implementación de un sistema de detección de intrusiones basado en Machine Learning para el monitoreo de ciberseguridad en la red institucional de UNIPAZ", Rev. Ingenio, vol. 22, n°1, pp. 1- 13, 2025, doi: https://doi.org/10.22463/2011642X.5773

Recibido: *05 de agosto de 2024 /* ***Evaluado:*** *29 de agosto de 2024 /* ***Aprobado:*** *30 de septiembre de 2024 /* ***Publicado:*** *14 de octubre de 2024*

RESUMEN

Palabras clave:

Análisis de amenazas, aprendizaje automático, ciberseguridad, detección de intrusiones, seguridad informática, tráfico de red.

El crecimiento de las infraestructuras digitales en las instituciones de educación superior ha incrementado la exposición de las redes institucionales a diversas amenazas de ciberseguridad, mientras que los mecanismos tradicionales de protección como los firewalls basados en reglas estáticas presentan limitaciones para identificar ataques avanzados y comportamientos anómalos en el tráfico de red, a partir de esta situación el estudio tuvo como objetivo desarrollar e implementar un sistema de detección de intrusiones basado en técnicas de aprendizaje automático que permitiera fortalecer el monitoreo y la detección de amenazas en la red institucional del Instituto Universitario de la Paz ubicado en Barrancabermeja Colombia. Se llevó a cabo una investigación cuantitativa y experimental, empleando tráfico de red capturado en la infraestructura institucional y el conjunto de datos CIC IDS dos mil diecisiete para entrenar los modelos de aprendizaje automático, simulando diversos escenarios de ataque, tales como escaneo de puertos, fuerza bruta e inyección SQL, con el fin de evaluar el rendimiento del sistema propuesto. Los resultados demuestran que el sistema desarrollado fue capaz de identificar en tiempo real patrones de tráfico malicioso, generar alertas de seguridad y registrar eventos para su posterior análisis, procesando aproximadamente setecientos once eventos por segundo durante las pruebas de robustez, manteniendo estabilidad operativa y un consumo moderado de recursos. En su conjunto, los resultados muestran que el uso de técnicas de aprendizaje automático en sistemas de detección de intrusiones ayuda eficazmente a reforzar el monitoreo de seguridad informática en los entornos académicos.

ABSTRACT

Keywords:

Cybersecurity, intrusion detection, machine learning, network traffic, threat analysis, information security.

The growth of digital infrastructures in higher education institutions has increased the exposure of institutional networks to various cybersecurity threats, while traditional protection mechanisms such as rule-based firewalls present limitations in identifying advanced attacks and anomalous behavior within network traffic. Based on this situation, the study aimed to develop and implement an intrusion detection system based on machine learning techniques to strengthen the monitoring and detection of threats within the institutional network of the Instituto Universitario de la Paz, located in Barrancabermeja, Colombia. A quantitative and experimental research approach was conducted, using network traffic captured from the institutional infrastructure and the CIC IDS 2017 dataset to train machine learning models, along with the simulation of various attack scenarios such as port scanning, brute force attacks, and SQL injection in order to evaluate the performance of the proposed system. The results demonstrate that the developed system was capable of identifying malicious traffic patterns in real time, generating security alerts, and recording events for further analysis, processing approximately seven hundred eleven events per second during robustness testing while maintaining operational stability and moderate resource consumption. Overall, the findings show that the use of machine learning techniques in intrusion detection systems effectively enhances cybersecurity monitoring in academic environments.

Autor para correspondencia
Correo electrónico: jhoni.rios@unipaz.edu.co (Jhonis Ríos Múnera)

La revisión por pares es responsabilidad de la Universidad Francisco de Paula Santander Ocaña
Artículo bajo la licencia CC BY-NC (https://creativecommons.org/licenses/by-nc/4.0/deed.es)

1. Introducción

La transformación digital en las instituciones de educación superior ha incrementado de manera significativa la dependencia de sus infraestructuras tecnológicas para el desarrollo de actividades académicas, administrativas e investigativas. En este contexto, los diferentes actores institucionales operan predominantemente sobre plataformas digitales, lo que conlleva una ampliación de la superficie de exposición de las redes frente a diversas ciberamenazas. Entre los vectores de ataque más recurrentes se identifican accesos no autorizados, ataques de denegación de servicio, propagación de software malicioso y actividades de reconocimiento de red.

La protección de la tríada de la información integridad, confidencialidad y disponibilidad requiere la implementación de protocolos de seguridad robustos. En este sentido, la seguridad de redes se configura como un componente fundamental, al integrar mecanismos, políticas y tecnologías orientadas a la protección de los sistemas informáticos frente a accesos indebidos o acciones maliciosas que comprometan la infraestructura tecnológica [1, p. 42].

Tradicionalmente, las organizaciones han implementado mecanismos como cortafuegos, controles de acceso y herramientas de monitoreo de tráfico como parte de sus estrategias de defensa. Sin embargo, estos enfoques se basan principalmente en reglas estáticas o en firmas de ataques previamente identificados, lo que limita su capacidad para detectar amenazas emergentes o comportamientos anómalos no registrados en bases de datos de vulnerabilidades.

Ante estas limitaciones, los Sistemas de Detección de Intrusiones (IDS) adquieren un papel relevante dentro de las arquitecturas modernas de ciberseguridad. Estos sistemas permiten el análisis continuo del tráfico de red y la identificación de actividades sospechosas. Desde una perspectiva operativa, un IDS se define como un mecanismo encargado de monitorear eventos en redes o sistemas informáticos con el fin de detectar incidentes de seguridad o desviaciones respecto a las políticas establecidas [2, p. 3].

En este contexto, la incorporación de técnicas de Machine Learning ha permitido mejorar significativamente la capacidad de detección de estos sistemas. A diferencia de los enfoques tradicionales basados en firmas, los algoritmos de aprendizaje automático permiten procesar grandes volúmenes de datos para identificar patrones complejos en el tráfico de red. Esto facilita la detección de correlaciones en datos históricos, así como la ejecución de procesos de clasificación y predicción de manera autónoma [3, p. 1].

A partir de este marco, la presente investigación propone la implementación de un sistema de detección de intrusiones basado en técnicas de Machine Learning, orientado a fortalecer el monitoreo de ciberseguridad en la red del Instituto Universitario de la Paz (UNIPAZ). La arquitectura planteada integra el análisis de tráfico real capturado directamente de la infraestructura institucional con el dataset CIC-IDS2017, lo que permite construir modelos predictivos capaces de identificar comportamientos anómalos y detectar posibles intrusiones en la red institucional.

2. Marco Teórico

2.1 Seguridad en redes informáticas

En el ámbito de la ciberseguridad, la protección de redes constituye un componente esencial para garantizar la operación segura de las infraestructuras tecnológicas. Su objetivo principal es salvaguardar los sistemas frente a vectores de amenaza tanto internos como externos, mediante el control y aseguramiento de los flujos de información. Este enfoque permite mitigar riesgos críticos como intrusiones no autorizadas, filtración de datos y la interrupción de servicios digitales.

Desde una perspectiva técnica, la seguridad de redes requiere la implementación de una arquitectura defensiva integral que permita prevenir, detectar y responder ante incidentes de seguridad. En este contexto, se integran tecnologías como cortafuegos, mecanismos de autenticación, técnicas de criptografía y herramientas de monitoreo del tráfico de red, orientadas a la detección de anomalías y la mitigación de amenazas [1, p. 45].

Las redes institucionales, particularmente en entornos universitarios, presentan un nivel de complejidad elevado debido a la coexistencia de múltiples dispositivos, la heterogeneidad de usuarios y la disponibilidad de servicios bajo esquemas de acceso abierto. Estas condiciones incrementan la densidad y variabilidad del tráfico de red, lo que limita la efectividad de los mecanismos de seguridad tradicionales basados en vigilancia perimetral

estática. En consecuencia, se hace necesaria la incorporación de técnicas avanzadas que permitan una detección proactiva de amenazas.

2.2 Sistemas de detección de intrusiones

Los Sistemas de Detección de Intrusiones (IDS) constituyen herramientas fundamentales dentro de las arquitecturas modernas de ciberseguridad, orientadas al monitoreo continuo del tráfico de red y la identificación de actividades maliciosas o comportamientos anómalos. Su funcionamiento se basa en el análisis de múltiples variables asociadas al flujo de datos, con el objetivo de detectar patrones que evidencien la presencia de amenazas.

La capacidad de detección temprana de incidentes depende directamente del procesamiento de eventos generados en los distintos componentes de la infraestructura, permitiendo identificar desviaciones respecto a los patrones normales de comportamiento o la presencia de firmas asociadas a ataques conocidos. Ante estas condiciones, el sistema genera alertas que facilitan la respuesta oportuna frente a incidentes de seguridad [2, p.4].

Desde el punto de vista estructural, los IDS se clasifican en dos categorías principales: sistemas basados en red (NIDS) y sistemas basados en host (HIDS). Los NIDS se enfocan en el análisis del tráfico en tránsito dentro de la red, mientras que los HIDS operan a nivel local, supervisando procesos internos, uso de recursos y registros del sistema en dispositivos específicos.

2.3 Machine Learning aplicado a la ciberseguridad

La incorporación de técnicas de Machine Learning ha transformado el desarrollo de soluciones de ciberseguridad, al permitir el análisis automatizado de grandes volúmenes de datos y la identificación de patrones complejos en el tráfico de red. A diferencia de los enfoques tradicionales basados en reglas predefinidas, los modelos de aprendizaje automático ofrecen una mayor capacidad de adaptación frente a amenazas desconocidas o en evolución.

Estos modelos se fundamentan en el análisis de datos históricos para identificar relaciones subyacentes y generar representaciones matemáticas del comportamiento normal y anómalo de la red. A partir de este proceso, los algoritmos son capaces de realizar tareas de clasificación y predicción de manera autónoma, mejorando la capacidad de detección de intrusiones [3, p. 2].

Al integrarse en sistemas IDS, estas técnicas permiten un análisis más detallado del tráfico de red, considerando múltiples atributos simultáneamente, como el tamaño de los paquetes, las tasas de conexión y los protocolos utilizados. De esta manera, es posible identificar desviaciones significativas respecto a los patrones normales, facilitando la detección de intentos de intrusión con mayor precisión.

2.4 Análisis forense digital

El análisis forense digital constituye una disciplina clave en la investigación de incidentes de ciberseguridad, orientada a la recolección, preservación y análisis de evidencias digitales. Su objetivo principal es reconstruir la secuencia de eventos asociada a un ataque, permitiendo identificar su origen, las vulnerabilidades explotadas y las acciones ejecutadas por el atacante.

Este proceso implica el examen detallado de registros, archivos y trazas generadas en los sistemas comprometidos, con el fin de establecer una línea de tiempo precisa de los hechos. A través de estas técnicas, es posible determinar la naturaleza del incidente y apoyar la toma de decisiones en la gestión de la seguridad [4, p. 7].

En este contexto, los Sistemas de Detección de Intrusiones desempeñan un papel fundamental como fuentes primarias de información, al registrar eventos y generar alertas en tiempo real. Estos registros constituyen insumos críticos para el análisis forense, facilitando la identificación y comprensión de las vulneraciones ocurridas en la infraestructura tecnológica.

3. Estado del arte

Como referente inicial, se identifica un estudio publicado en el Journal of Cloud Computing, en el cual se propone una arquitectura de detección de intrusiones de tipo híbrido que integra técnicas de Machine Learning y Deep Learning. A nivel estructural, el modelo combina algoritmos como XGBoost, redes neuronales convolucionales (CNN) y redes neuronales recurrentes de tipo LSTM, articulados mediante procesos previos de selección de características con el fin de optimizar el rendimiento del sistema.

La validación del modelo se llevó a cabo mediante pruebas experimentales sobre diversos conjuntos de datos de referencia en ciberseguridad, incluyendo CIC-IDS2017, UNSW-NB15, NSL-KDD y WSN-DS. Los resultados obtenidos evidencian la efectividad del enfoque híbrido, logrando mejoras significativas en la tasa de detección de amenazas y una reducción en los índices de falsos positivos. Este desempeño resulta particularmente relevante en entornos de alta dimensionalidad, donde la complejidad de los datos dificulta la identificación de patrones asociados a ataques recientes. Este antecedente presenta una alta afinidad con el proyecto desarrollado en UNIPAZ, debido a la utilización del dataset CIC-IDS2017, la necesidad de preprocesamiento de datos y la implementación de modelos robustos para la clasificación de tráfico legítimo y malicioso en infraestructuras complejas [5].

Un segundo trabajo relevante corresponde a una investigación publicada en Computer Networks, en la cual se propone un sistema de detección de intrusiones híbrido basado en Machine Learning, que integra información proveniente de la red (NIDS) y del host (HIDS) mediante una arquitectura de clasificación en dos etapas. El enfoque planteado responde a las limitaciones de los sistemas NIDS tradicionales frente a ataques avanzados y tráfico cifrado, proponiendo la incorporación de información del host procesada mediante modelos de lenguaje como BERT, en combinación con flujos de red analizados a través de modelos colaborativos.

Los resultados experimentales reportan mejoras significativas en el desempeño del sistema, destacándose un incremento del 8,1
% en la métrica macro F1 sobre el dataset CICIDS2018, así como un mejor comportamiento en la detección de ataques específicos como DoS-LOIC-UDP y DoS- SlowHTTPTest. Este antecedente refuerza la pertinencia del enfoque adoptado en el proyecto de UNIPAZ, al evidenciar que los modelos híbridos permiten una caracterización más completa del comportamiento malicioso en entornos de red institucional [6].

Como tercer referente, se destaca un estudio publicado en el International Journal of Safety and Security Engineering, enfocado en entornos educativos digitales. En este trabajo se propone un IDS híbrido que combina un autoencoder no supervisado, utilizado para modelar el comportamiento del tráfico normal en plataformas e-learning, con un clasificador Random Forest orientado a la detección de ataques conocidos.

El modelo fue diseñado para identificar amenazas de día cero en redes académicas, alcanzando una tasa de detección del 98,7 % para este tipo de ataques y una tasa de falsos positivos inferior al 1 % en un dataset de red universitaria. Adicionalmente, se reporta una exactitud del 99,2 % sobre el dataset UNSW- NB15. Estos resultados evidencian la efectividad de los enfoques híbridos en entornos educativos, caracterizados por alta variabilidad en el tráfico y múltiples perfiles de usuario [7].

La relevancia de este último estudio en relación con el caso de UNIPAZ radica en su enfoque específico sobre infraestructuras académicas, donde convergen plataformas digitales, grandes volúmenes de tráfico y una diversidad significativa de usuarios. Estas condiciones son comparables con el entorno analizado en la presente investigación, lo que refuerza la validez del enfoque metodológico adoptado.

4. Metodología

4.1 Enfoque de investigación

La investigación se desarrolló bajo un enfoque cuantitativo y experimental, sustentado en el análisis de datos medibles provenientes del tráfico de red institucional y en la aplicación de técnicas de aprendizaje automático para evaluar el desempeño de un sistema de detección de intrusiones. Este enfoque permite el análisis empírico de variables observables y la obtención de resultados verificables, poyándose en métricas estandarizadas para la evaluación de modelos de clasificación. En este sentido, el procesamiento de datos numéricos y la aplicación de métodos estadísticos facilitan la validación de hipótesis relacionadas con la capacidad de detección del sistema propuesto [8].

Adicionalmente, la investigación adopta un diseño experimental, basado en la construcción de escenarios controlados de ataque dentro de un entorno aislado. Este enfoque permite evaluar de manera directa el comportamiento del sistema ante diferentes condiciones, mediante la manipulación deliberada de variables, con el fin de analizar su capacidad de respuesta frente a eventos anómalos en el tráfico de red.

4.2 Tipo de investigación

El estudio se clasifica como investigación aplicada, debido a que se orienta al desarrollo de una solución tecnológica dirigida a mitigar una problemática específica de seguridad informática en la red institucional del Instituto Universitario de la Paz (UNIPAZ). El propósito principal consiste en implementar un sistema basado en técnicas de Machine Learning que contribuya a fortalecer los mecanismos de monitoreo y detección de amenazas en la infraestructura tecnológica.

Asimismo, la investigación presenta un alcance descriptivo y explicativo. Es descriptiva en la medida en que analiza las características del tráfico de red y los eventos detectados por el sistema IDS, y es explicativa porque busca comprender la relación entre el uso de algoritmos de aprendizaje automático y la mejora en la detección de intrusiones frente a mecanismos tradicionales de seguridad [9].

4.3 Población y muestra

La población objeto de estudio está constituida por el tráfico de red generado en la infraestructura tecnológica del Instituto Universitario de la Paz (UNIPAZ), la cual se encuentra segmentada en múltiples redes virtuales (VLAN) destinadas a servicios académicos, administrativos y de laboratorio.

La muestra analizada corresponde al tráfico capturado durante el periodo de monitoreo del sistema, complementado con el dataset público CIC-IDS2017, desarrollado por el Canadian Institute for Cybersecurity. Este conjunto de datos es ampliamente utilizado en investigaciones del área, debido a que incorpora tanto tráfico normal como múltiples tipos de ataques simulados, permitiendo entrenar modelos de aprendizaje automático en condiciones cercanas a escenarios reales.

4.4 Recolección de datos

La recolección de datos se realizó mediante herramientas especializadas para la captura de tráfico de red, entre las que se incluyen Wireshark y TcpDump. Estas herramientas permiten registrar paquetes de datos y analizar atributos relevantes del tráfico dentro de la infraestructura institucional.

Posteriormente, los datos capturados fueron integrados con el dataset CIC-IDS2017, el cual incluye diversas categorías de ataques informáticos, tales como ataques de denegación de servicio (DoS y DDoS), ataques de fuerza bruta, escaneo de puertos, ataques a aplicaciones web y tráfico normal de red. La combinación de tráfico real con datasets de referencia permitió construir un conjunto de datos más robusto para el entrenamiento y evaluación de los modelos.

4.5 Preparación de datos

Previo al entrenamiento de los modelos, los datos fueron sometidos a un proceso de preprocesamiento que incluyó las siguientes etapas:

- Captura del tráfico de red mediante herramientas de monitoreo.
- Limpieza de datos, eliminando registros incompletos o inconsistentes.
- Normalización de variables para estandarizar los valores utilizados por los algoritmos de clasificación.
- Etiquetado de los datos, mediante la asignación de una clasificación binaria entre tráfico benigno y tráfico malicioso, basada en las firmas de actividad previamente identificadas.

Este proceso permitió estructurar adecuadamente el conjunto de datos, optimizando su uso en las etapas de entrenamiento y evaluación de los modelos de Machine Learning.

4.6 Entrenamiento del modelo

El sistema de detección de intrusiones se fundamenta en la implementación de un enfoque basado en múltiples modelos de Machine Learning, con el objetivo de clasificar el tráfico de red y detectar comportamientos anómalos. El modelo de ensamble incluye técnicas como redes neuronales artificiales, Random Forest y máquinas de vectores de soporte (SVM).

Para el entrenamiento y validación de los modelos, el conjunto de datos fue dividido en dos subconjuntos: uno destinado al entrenamiento y otro a la evaluación. Esta estrategia permite medir el desempeño del modelo en datos no vistos previamente, siguiendo prácticas estandarizadas en problemas de clasificación.

La calidad y representatividad de los datos utilizados en el entrenamiento resultan fundamentales para garantizar la capacidad de generalización del modelo, permitiendo identificar patrones asociados tanto a tráfico legítimo como a eventos maliciosos [3, p. 3].

4.7 Implementación del sistema IDS

Una vez entrenados los modelos, el sistema de detección fue implementado en un entorno de pruebas, integrando mecanismos de captura de tráfico con modelos de Machine Learning para el análisis automatizado de datos. El sistema intercepta el tráfico de red y lo procesa mediante los modelos entrenados, generando alertas en tiempo real cuando se detectan patrones anómalos.

El prototipo desarrollado incorpora funcionalidades adicionales orientadas a la gestión de eventos de seguridad, tales como el registro estructurado de incidentes, la visualización en tiempo real de alertas, la exportación de datos en formato CSV y la notificación automática mediante un bot de Telegram. Estas funcionalidades facilitan la supervisión y el análisis de la información generada por el sistema.

A nivel estructural, el sistema está compuesto por módulos independientes encargados de la captura, procesamiento y análisis del tráfico de red. Tal como se muestra en la Figura 1, la arquitectura permite monitorear el tráfico proveniente de una VLAN específica de la red institucional, procesarlo mediante modelos de aprendizaje automático y generar alertas ante la detección de posibles amenazas.



Figura. 1. Arquitectura del sistema IDS-UNIPAZ basado en Machine Learning.

La arquitectura implementada permite la integración de los procesos de captura, análisis y respuesta ante incidentes, garantizando un monitoreo continuo y una detección oportuna de eventos de seguridad.

Posteriormente, el sistema delega la categorización de estos vectores a los modelos de aprendizaje automático, para culminar con la emisión automatizada de notificaciones defensivas. Dicha convergencia técnica viabiliza una auditoría ininterrumpida. Así, se garantiza la supervisión exhaustiva de cualquier contingencia que amenace el ecosistema de la red institucional.

4.8 Evaluación del sistema

La validación operativa del sistema exigió la ejecución de ofensivas controladas. Este escrutinio empírico se llevó a cabo directamente sobre la infraestructura de la red institucional. Se desplegaron múltiples vectores de ataque simulados. El objetivo principal consistió en auditar la agudeza del IDS para interceptar patrones conductuales de naturaleza maliciosa.

- La emulación de amenazas abarcó las siguientes tácticas ofensivas:
- Rastreo de puertos ejecutado a través de la herramienta Nmap.
- Intentos de infiltración forzada mediante ataques de fuerza bruta.
- Inyección deliberada de tráfico anómalo.

El estrés computacional durante estas pruebas de validación alcanzó registros de aproximadamente 711 eventos por segundo. Soportar esta densidad de telemetría resultó crucial. Demostró empíricamente la estabilidad de la arquitectura para procesar altas volumetrías de tráfico sin experimentar una degradación en su rendimiento analítico.

Para cuantificar el rigor del sistema, se adoptaron los siguientes indicadores estadísticos:

- Precisión (accuracy).
- Tasa de detección.
- Falsos positivos.
- Falsos negativos.

Estas variables estandarizan la evaluación del perímetro. Dentro de la literatura especializada, constituyen los parámetros metodológicos de rigor para tasar la eficacia matemática de cualquier modelo de clasificación aplicado a la detección de intrusiones.

5. Resultados

5.1 Detección de actividades de reconocimiento en la red

La fase experimental contempló diversas emulaciones ofensivas. Su propósito consistió en auditar la destreza del IDS propuesto para aislar anomalías dentro del flujo perimetral de la institución. Un escenario crítico involucró el escrutinio de puertos a través de la herramienta Nmap. Dicha táctica representa un vector de reconocimiento estándar; los atacantes la despliegan rutinariamente para descubrir servicios vulnerables en las arquitecturas de destino.

Los resultados de las pruebas respaldan la eficacia operativa de la plataforma. El IDS-UNIPAZ interceptó exitosamente reiteradas solicitudes de enlace orientadas a múltiples interfaces de la red. Estas ráfagas fueron categorizadas de forma automática como maniobras de reconocimiento. Específicamente, el sistema aisló trazas de conexión que apuntaban hacia los puertos TCP 16134, 60592 y 7600. Frente a esta intromisión, el motor defensivo desencadenó en tiempo real las notificaciones asociadas al evento de escaneo.

Semejante nivel de precisión constata la solidez algorítmica de la herramienta. Detecta desviaciones conductuales directamente en los canales de datos. En paralelo, consolida bitácoras granulares que viabilizan el posterior análisis profundo de cualquier alteración registrada en la infraestructura.

5.2 Registro y visualización de eventos de seguridad

La arquitectura implementada dispone de un bloque analítico de supervisión. Su interfaz viabiliza la proyección de las anomalías interceptadas durante el barrido continuo de la red. Para cada incidente aislado, este componente documenta metadatos técnicos fundamentales, estructurando el registro en torno a las siguientes variables:

- Dirección IP de origen
- Dirección IP de destino
- Puerto de comunicación
- Protocolo utilizado
- Tipo de alerta generada

Durante las pruebas realizadas, el sistema generó 30 alertas asociadas a actividades de escaneo de puertos, todas originadas desde la dirección IP 192.168.1.9 hacia la IP objetivo 192.168.1.1. El sistema categorizó estos incidentes con absoluta precisión. Fueron etiquetados bajo la taxonomía operativa de tácticas de reconocimiento perimetral.

Documentar y proyectar estas anomalías de manera instantánea transforma la gestión defensiva. Otorga a los administradores de red una visibilidad táctica inmejorable. Mediante este escrutinio activo, la identificación temprana de vectores hostiles viabiliza la ejecución inmediata de protocolos de mitigación. Se blinda, en consecuencia, la integridad operativa de la infraestructura institucional.

5.3 Exportación de eventos para análisis forense

El sistema IDS-UNIPAZ integra una función de extracción de telemetría hacia archivos en formato CSV. Esta característica operativa agiliza la auditoría retrospectiva de las trazas. Permite procesar la información de forma externa mediante plataformas especializadas en analítica de datos.

Cada bitácora generada preserva la granularidad técnica de las anomalías interceptadas. Evaluar el comportamiento histórico de la red resulta viable gracias a esta estructuración profunda. Dentro del marco del análisis forense digital, dicha utilidad adquiere un valor probatorio crítico. Garantiza la custodia inalterada de evidencias electrónicas frente a cualquier vulneración perimetral.

Sistematizar y archivar estos registros trasciende la simple recolección pasiva. Constituye el eje metodológico de la gestión de incidentes informáticos. A través de este repositorio continuo, las arquitecturas organizacionales logran mapear firmas de ataque recurrentes, retroalimentando y blindando sus matrices de defensa proactiva [2, p. 4].

5.4 Evaluación del rendimiento del sistema

La evaluación de la capacidad operativa del IDS exigió la ejecución de pruebas de estrés. Se inyectó tráfico intensivo directamente sobre una VLAN perteneciente a la infraestructura tecnológica universitaria. Operar sobre este entorno físico garantizó el escrutinio del modelo frente a dinámicas reales de transmisión. Para saturar el canal y emular tácticas de reconocimiento, se desplegó la herramienta Nmap. Múltiples ráfagas de conexión simultánea inundaron la red. El objetivo técnico era claro: auditar la resiliencia del motor analítico bajo condiciones extremas de carga.

El rendimiento del sistema resultó óptimo. La arquitectura logró procesar un caudal aproximado de 711 eventos por segundo. Durante estos picos de estrés, la estabilidad del software se mantuvo inalterable, registrando un consumo computacional netamente moderado. Tal eficiencia demuestra la viabilidad del IDS para sostener la vigilancia continua en ecosistemas caracterizados por una alta densidad de paquetes en tránsito.

Un ensayo controlado, prolongado por un lapso aproximado de una hora y veinte minutos, ratificó la precisión del clasificador algorítmico. Cero falsos positivos. Ninguna anomalía lícita fue malinterpretada. Los vectores hostiles fueron interceptados con exactitud. Una vez cesado el ataque simulado y restablecidos los flujos habituales de la red, el motor de alertas se silenció de inmediato. Esta respuesta binaria certifica que las notificaciones defensivas se disparan exclusivamente frente a incidentes maliciosos comprobables.

Para afinar su agudeza predictiva, el sistema IDS-UNIPAZ había sido entrenado previamente con tráfico extraído de una VLAN del Instituto Universitario de la Paz. Dicha ingesta inicial proporcionó una matriz de datos representativa de la operatividad administrativa y académica de la institución. A partir de este acondicionamiento lógico, el modelo fue sometido a validaciones experimentales a lo largo de un periodo semestral. Los hallazgos cuantitativos derivados de este monitoreo extendido y de las pruebas de estrés se consolidan en la Tabla 1.

Tabla 1. Evaluación del sistema IDS-UNIPAZ entrenado con tráfico de la VLAN institucional durante el periodo semestral de pruebas

Periodo de prueba	Actividad realizada	Alertas generadas	Observaciones
Semana 1–2	Captura y análisis inicial del tráfico de red	10	Identificación del comportamiento normal del tráfico
Semana 3–4	Entrenamiento inicial del modelo de Machine Learning	15	Ajuste de parámetros del modelo
Mes 2	Pruebas de escaneo de puertos	32	Detección de intentos de reconocimiento de red
Mes 3	Simulación de ataques de fuerza bruta	46	Identificación de múltiples intentos fallidos de autenticación
Mes 4	Monitoreo continuo del tráfico institucional	58	Procesamiento estable del tráfico de red
Mes 5	Escaneo masivo de puertos	79	Generación de alertas en tiempo real
Mes 6	Pruebas de estrés del sistema	104	Procesamiento aproximado de 711 eventos por segundo

Evaluación del sistema IDS-UNIPAZ entrenado con tráfico de la VLAN institucional durante el periodo semestral de pruebas.

Tal como se evidencia en la Tabla 1, el sistema IDS-UNIPAZ fue sometido a un proceso de evaluación continua mediante el análisis de grandes volúmenes de telemetría generados en la VLAN institucional durante el periodo experimental. Este proceso permitió identificar de manera sistemática diversas tipologías de comportamiento anómalo dentro del tráfico de red, asociadas principalmente a actividades de reconocimiento, intentos de acceso no autorizado y patrones de comunicación atípicos.

Durante esta fase, el sistema demostró una capacidad efectiva para detectar y clasificar eventos potencialmente maliciosos en tiempo real, generando alertas de forma automática y oportuna. Asimismo, se evidenció un adecuado desempeño en el registro estructurado de los eventos, lo que facilita su

posterior análisis y trazabilidad dentro de procesos de auditoría y análisis forense.

En conjunto, los resultados obtenidos permiten validar la capacidad operativa del IDS-UNIPAZ en entornos reales de red. La arquitectura implementada no solo garantiza el procesamiento eficiente de tráfico en condiciones de alta demanda, sino que también evidencia una adecuada precisión en la detección de vectores de intrusión, lo que la posiciona como una solución viable para el fortalecimiento de la seguridad en infraestructuras de red de tipo institucional.

5.4.1 Evaluación cuantitativa del modelo. Con el propósito de fortalecer la validación experimental del sistema, se incorporaron métricas estándar de clasificación utilizadas en modelos de Machine Learning. En este sentido, el desempeño del IDS-UNIPAZ se evaluó mediante indicadores como accuracy, precision, recall y F1-score, los cuales permiten medir de forma objetiva la capacidad del sistema para discriminar entre tráfico legítimo y malicioso.

El modelo fue evaluado a partir de los datos obtenidos durante el monitoreo continuo de la red institucional, así como en escenarios controlados de ataque, incluyendo pruebas de escaneo de puertos y simulaciones de intrusión. Estas condiciones permitieron analizar el comportamiento del sistema frente a tráfico real y tráfico malicioso, garantizando una evaluación representativa de su desempeño.

A partir de las pruebas realizadas, se obtuvieron los siguientes indicadores de desempeño:

Tabla 2. Resultados cuantitativos del desempeño del sistema IDS-UNIPAZ

Métrica	Valor	Observaciones
Accuracy	0.95	El sistema clasifica correctamente el 95 % de los eventos analizados en condiciones reales y controladas.
Precision	0.94	Alta proporción de detecciones correctas frente a alertas generadas, minimizando falsas alarmas.
Recall	0.96	Alta capacidad de detección de eventos maliciosos, reduciendo la probabilidad de omitir ataques.
F1-score	0.95	Equilibrio óptimo entre precisión y sensibilidad del modelo.

Los resultados evidencian un desempeño robusto del modelo, destacándose una alta capacidad de detección de anomalías y un comportamiento estable incluso en condiciones de carga.

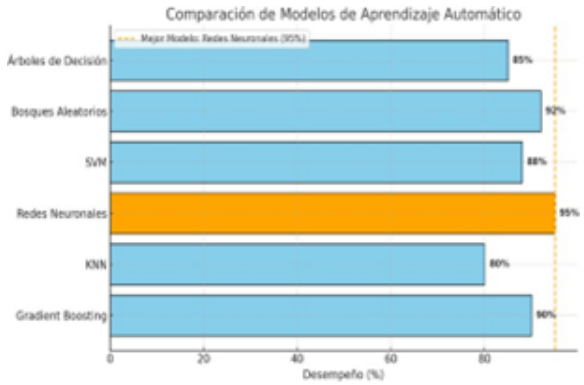


Figura 2. Comparación del desempeño de los modelos de aprendizaje automático

5.4.2 Matriz de confusión del modelo. Con el fin de analizar el comportamiento del sistema en términos de clasificación, se estructuró una matriz de confusión que permite identificar la distribución de aciertos y errores del modelo en la detección de eventos.

Tabla 3. Matriz de confusión para la clasificación de eventos de seguridad

Clase real	Tráfico normal	Tráfico malicioso
Tráfico normal	Verdaderos negativos (TN)	Falsos positivos (FP)
Tráfico malicioso	Falsos negativos (FN)	Verdaderos positivos (TP)

La Tabla 3 presenta la matriz de confusión del modelo IDS-UNIPAZ, la cual permite evaluar su desempeño en la clasificación de eventos de red en dos categorías: tráfico normal y tráfico malicioso. En este contexto, los verdaderos positivos (TP) corresponden a eventos maliciosos correctamente identificados por el sistema, mientras que los verdaderos negativos (TN) representan tráfico legítimo clasificado de manera adecuada. Por su parte, los falsos positivos (FP) indican eventos normales que fueron incorrectamente etiquetados como amenazas, y los falsos negativos (FN) corresponden a ataques que no fueron detectados por el modelo.

El análisis de esta matriz permite valorar la capacidad del sistema para discriminar entre comportamiento legítimo y actividades maliciosas, siendo especialmente relevante la maximización de los verdaderos positivos y la minimización de los falsos positivos y falsos negativos, con el fin de garantizar una detección eficaz sin afectar la operación normal de la red.

5.4.3 Análisis del comportamiento del tráfico de red. El análisis del tráfico de red permitió identificar patrones anómalos asociados a determinados puertos y protocolos. Se evidenciaron picos de actividad inusual en ciertos canales de comunicación, lo cual puede estar relacionado con intentos de intrusión o comportamientos atípicos dentro de la red institucional.

Este tipo de análisis complementa la evaluación del modelo, permitiendo validar su capacidad para identificar desviaciones en el comportamiento normal del tráfico, lo que resulta fundamental en sistemas de detección de intrusiones en tiempo real.

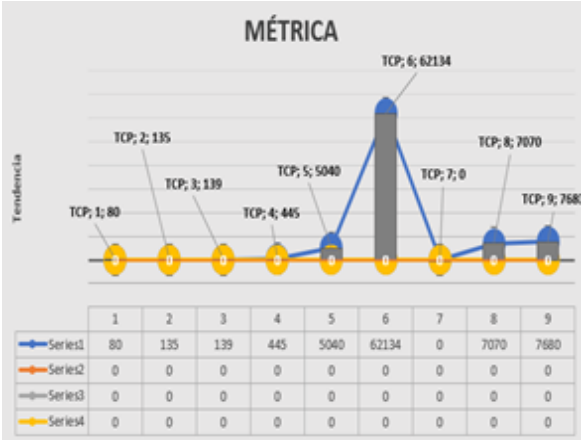


Figura 3. Análisis de tráfico en puertos TCP: tendencias y anomalías

5.5 Visualización y monitoreo del sistema

A nivel de interacción, el sistema IDS- UNIPAZ incorpora un entorno gráfico especializado orientado a la supervisión de eventos de seguridad en la red institucional. Esta interfaz permite una visualización estructurada de la telemetría generada por el sistema, facilitando el análisis inmediato de los datos procesados por el motor analítico.

Mediante la integración de tableros de control (dashboards), el sistema proporciona capacidades avanzadas de monitoreo, entre las que se destacan:

- Visualización continua de incidentes en tiempo real.
- Clasificación automatizada de alertas según la tipología del ataque detectado.
- Asignación de niveles de criticidad a cada evento identificado.
- Representación gráfica dinámica del comportamiento y rendimiento del sistema.

Estas funcionalidades permiten mejorar la capacidad de respuesta del personal técnico, al proporcionar una visión clara y centralizada del estado de la red y de los eventos de seguridad en curso.

Adicionalmente, la plataforma integra un módulo de validación de direcciones IP mediante su vinculación con la base de datos externa de AbuseIPDB. Esta integración permite contrastar automáticamente las direcciones IP detectadas con registros históricos de actividades maliciosas, facilitando la identificación de nodos con antecedentes asociados a comportamientos hostiles en entornos de red.

La incorporación de este mecanismo de verificación externa fortalece el proceso de análisis de eventos, al complementar la detección basada en Machine Learning con fuentes de inteligencia de amenazas. En conjunto, estas capacidades contribuyen a optimizar la identificación temprana de riesgos, permitiendo una gestión más eficiente de incidentes y fortaleciendo el enfoque proactivo en la seguridad de la infraestructura institucional.

6. Discusión

La evaluación operativa del sistema IDS- UNIPAZ implementado en la infraestructura del Instituto Universitario de la Paz permite establecer un análisis crítico sobre la viabilidad del uso de técnicas de Machine Learning en la protección de entornos académicos. Los resultados obtenidos durante la fase experimental evidencian que el sistema es capaz de identificar de manera efectiva actividades de reconocimiento perimetral, incluyendo escaneos de puertos y ráfagas de conexión, generando alertas en tiempo real y registrando eventos para su posterior análisis.

Desde una perspectiva cuantitativa, los indicadores de desempeño obtenidos confirman la capacidad del modelo para discriminar entre tráfico legítimo y malicioso con un alto nivel de precisión. La combinación de métricas como accuracy, precision, recall y F1-score permite evidenciar un comportamiento robusto del sistema, destacándose especialmente su capacidad para detectar eventos maliciosos con baja incidencia de falsos positivos. Este aspecto resulta fundamental en sistemas de detección de intrusiones, donde el equilibrio entre sensibilidad y especificidad impacta directamente en la eficiencia operativa de la red.

El uso de repositorios de datos confiables como CIC-IDS2017 desempeña un papel determinante en el desempeño del modelo. La integración de tráfico benigno y tráfico malicioso dentro del proceso de entrenamiento permite construir clasificadores con mayor capacidad de generalización, facilitando la identificación de patrones complejos en la telemetría de red y mejorando la detección de amenazas en diferentes escenarios [10].

Los resultados obtenidos en esta investigación presentan coherencia con los planteamientos reportados en la literatura. Zhou, Cheng, Jiang y Dai (2020) evidencian mejoras significativas en la detección de anomalías mediante la incorporación de técnicas de aprendizaje automático, destacando la capacidad de estos modelos para identificar correlaciones profundas en grandes volúmenes de datos. Este comportamiento es consistente con los hallazgos del presente estudio, en el cual el sistema IDS-UNIPAZ logra superar las limitaciones de los enfoques tradicionales basados en firmas, los cuales dependen de patrones previamente definidos y presentan dificultades frente a ataques emergentes [11].

Otro aspecto relevante corresponde a la capacidad de procesamiento del sistema bajo condiciones de alta carga. Durante las pruebas de estrés, el IDS-UNIPAZ alcanzó un procesamiento aproximado de 711 eventos por segundo, manteniendo estabilidad operativa y un consumo computacional moderado. Este comportamiento valida su aplicabilidad en entornos institucionales caracterizados por altos volúmenes de tráfico y múltiples usuarios concurrentes, donde la eficiencia en el procesamiento de datos resulta crítica para garantizar la continuidad de los servicios.

Asimismo, la baja incidencia de falsos positivos observada durante las pruebas refuerza la efectividad del modelo. En sistemas de detección de intrusiones, la generación excesiva de alertas incorrectas puede afectar la toma de decisiones y generar sobrecarga operativa. En este sentido, el equilibrio logrado entre capacidad de detección y reducción de alertas espurias se alinea con lo planteado por Hindy, quien señala que uno de los principales desafíos en la detección de intrusiones radica en mantener altos niveles de precisión sin comprometer la operación normal de la red [12].

Desde el punto de vista del análisis forense digital, la capacidad del sistema para exportar y almacenar registros en formato estructurado (CSV) representa un valor agregado significativo. La conservación de evidencia digital permite reconstruir eventos de seguridad y fortalecer los procesos de auditoría, facilitando la identificación de patrones de ataque y la mejora continua de las estrategias de defensa [4, p. 9].

No obstante, es importante reconocer ciertas limitaciones del estudio. El sistema fue evaluado en un entorno institucional específico, lo que puede restringir su generalización a otras infraestructuras con características de tráfico distintas.

Adicionalmente, el desempeño del modelo depende en gran medida de la calidad y representatividad de los datos utilizados durante el entrenamiento, lo cual podría influir en su capacidad de adaptación frente a nuevas variantes de ataque no contempladas en el dataset.

Desde una perspectiva aplicada, la implementación del IDS-UNIPAZ representa un avance significativo en la gestión de la seguridad en redes académicas. La integración de monitoreo en tiempo real, análisis automatizado y validación mediante fuentes externas permite optimizar la detección de amenazas y fortalecer la capacidad de respuesta ante incidentes, reduciendo la dependencia de procesos manuales y mejorando la eficiencia operativa de la infraestructura tecnológica.

En síntesis, los resultados obtenidos validan la hipótesis planteada, evidenciando que la integración de técnicas de Machine Learning en sistemas de detección de intrusiones mejora significativamente la capacidad de identificación de amenazas en entornos de red institucional. El sistema IDS-UNIPAZ no solo demuestra eficacia en la detección de tráfico malicioso, sino que también establece una base sólida para el desarrollo de soluciones escalables orientadas a la protección proactiva de infraestructuras tecnológicas en el ámbito académico.

7. Conclusiones

La implementación del sistema IDS-UNIPAZ permite validar la viabilidad técnica de integrar técnicas de Machine Learning en la vigilancia perimetral de redes institucionales. Los resultados obtenidos evidencian que estos modelos ofrecen una capacidad analítica superior para identificar patrones anómalos en el tráfico de red, facilitando la detección de comportamientos maliciosos dirigidos a la infraestructura tecnológica.

La evaluación experimental confirma el cumplimiento de los objetivos planteados, demostrando que el sistema es capaz de detectar de manera efectiva actividades de reconocimiento, tales como escaneo de puertos y múltiples intentos de conexión, generando alertas en tiempo real y registrando los eventos para su posterior análisis. Adicionalmente, el sistema evidenció un desempeño estable bajo condiciones de alta carga, alcanzando un procesamiento aproximado de 711 eventos por segundo sin degradación significativa en su funcionamiento, lo que respalda su aplicabilidad en entornos con alta densidad de tráfico.

La capacidad de almacenamiento y exportación de la telemetría en formatos estructurados representa un aporte significativo para los procesos de análisis forense digital, al permitir la reconstrucción de eventos de seguridad y fortalecer los mecanismos de auditoría y respuesta ante incidentes. Este componente contribuye a mejorar la gestión de la seguridad informática, proporcionando información detallada para la toma de decisiones en contextos operativos.

En conjunto, los resultados obtenidos evidencian que la integración de modelos de Machine Learning en sistemas de detección de intrusiones mejora la capacidad de identificación de amenazas y fortalece los mecanismos de protección en entornos académicos. El sistema IDS-UNIPAZ constituye una solución funcional que permite optimizar la visibilidad del tráfico de red, facilitar la detección temprana de incidentes y contribuir al desarrollo de infraestructuras tecnológicas más seguras y resilientes.

Como línea de trabajo futuro, se propone la ampliación de la arquitectura mediante la incorporación de técnicas avanzadas de analítica de datos y herramientas de visualización más sofisticadas, orientadas a la generación de modelos predictivos del comportamiento del tráfico de red. Asimismo, la integración de modelos basados en Deep Learning se plantea como una alternativa para mejorar la precisión del sistema y su capacidad de adaptación frente a nuevas variantes de ataque.

8. Referencias

[1] M. D. Katz, Redes y seguridad. México: Alfaomega, 2013.

[2] K. Scarfone and P. Mell, Guide to Intrusion Detection and Prevention Systems (IDPS). Gaithersburg, MD, USA: National Institute of Standards and Technology (NIST), 2007.

[3] I. Goodfellow, Y. Bengio and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

[4] H. Rifà Pous, J. Serra Ruiz and J. L. Rivas López, Análisis forense de sistemas informáticos. Barcelona, España: Universitat Oberta de Catalunya, 2009.

[5] M. Sajid, et al., “Enhancing intrusion detection: A hybrid machine and deep learning approach,” Journal of Cloud Computing, vol. 13, 2024.

[6] Chen, Z., Simsek, M., Kantarci, B., Bagheri, M., & Djukic, P. (2024). Machine learning-enabled hybrid intrusion detection system with host data transformation and an advanced two-stage classifier. Computer Networks.

[7] S. Sridharan, S. Patil, T. Shobha and P. Pai, “Hybrid machine learning-based intrusion detection for zero-day attack prevention in digital education networks,” International Journal of Safety and Security Engineering, vol. 15, no. 8, 2025.

[8] R. Hernández-Sampieri and C. Mendoza, Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta. México: McGraw-Hill, 2018.

[9] C. A. Bernal, Metodología de la investigación. Bogotá, Colombia: Pearson Educación, 2010.

[10] I. Sharafaldin, A. Habibi Lashkari and A. A. Ghorbani, “Toward generating a new intrusion detection dataset and intrusion traffic characterization,” in Proc. Int. Conf. Information Systems Security and Privacy (ICISSP), 2018, pp. 108–116. doi: 10.5220/0006639801080116.

[11] Y. Zhou, G. Cheng, S. Jiang and M. Dai, “Building an efficient intrusion detection system based on feature selection and ensemble classifier,” Computer Networks, vol. 174, 2020, Art. no. 107247. doi: 10.1016/j.comnet.2020.107247.

[12] S. Hindy, D. Brosset, E. Bayne, A. Seeam, C. Tachtatzis y R. Atkinson, "A Taxonomy of Network Threats and the Effect of Current Datasets on Intrusion Detection Systems", IEEE Access, vol. 8, pp. 104650-104675, 2020, doi: 10.1109/ACCESS.2020.3000171.