PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

FACULTAD DE INGENIERÍA

TRABAJO DE TITULACIÓN PREVIO A LA OBTENCIÓN DEL GRADO DE MAGISTER EN SISTEMAS DE INFORMACIÓN MENCIÓN CIENCIA DE DATOS

TEMA:

Desarrollar un modelo predictivo de detección de ataques a herramientas de seguridad perimetral, de la COAC Jardín Azuayo

AUTOR:

Romel Mauricio Cando Jara

DIRECTOR:

Eduardo José Montero Bermúdez Msc.

Quito – Ecuador

2024

Indice de contenidos

Indice de ilustraciones

Indice de tablas

RESUMEN

En el panorama actual de la ciberseguridad, las amenazas han evolucionado considerablemente siendo mucho más complejas, con un mayor incremento e impacto, afectando tanto a organizaciones como a usuarios. Este incremento en los últimos años ha planteado un desafío significativo para las instituciones, que deben estar preparadas y emplear estrategias efectivas para mitigar estas amenazas.

En este contexto, la institución financiera COAC Jardin Azuayo se enfrenta a la tarea crucial de analizar los *logs* generados por herramientas de seguridad de frontera, como *firewalls*. Estos dispositivos, producen grandes volúmenes de datos, cuya gestión y análisis se ven dificultados por la limitada capacidad de almacenamiento de los dispositivos actuales, los cuales no están diseñados para tal propósito.

Existen regulaciones internas y externas que requieren la conservación de datos históricos, con la necesidad de garantizar la disponibilidad de *logs*, conforme a las políticas institucionales. Este proyecto busca contribuir a la implementación de medidas efectivas, optimizando el tiempo de análisis de datos mediante la aplicación de técnicas de *machinelearning*.

La adecuada preparación y explotación de estos datos permiten no solo la extracción de información clave, sino también el desarrollo de modelos predictivos y técnicas de aprendizaje automático, elementos fundamentales para fomentar la innovación en campos como la ciencia y la tecnología.

Esta tesis presenta un análisis mediante el uso de algoritmo de regresión logística que permite realizar una predicción basada en la severidad de los ataques registrados en los logs de las herramientas de seguridad, se realizaron evaluaciones de los resultados para estimar la efectividad del modelo. Los resultados demostraron que el modelo clasificó correctamente los ataques en diferentes niveles de severidad, como se reflejó en las métricas de rendimiento de cada clase. Esto indicó que el agrupamiento de comportamientos según la severidad fue adecuado.

ABSTRACT

In the current cybersecurity landscape, threats have evolved considerably, becoming much more complex, with greater growth and impact, affecting both organizations and users. This increase in recent years has posed a significant challenge for institutions, which must be prepared and employ effective strategies to mitigate these threats.

In this context, the financial institution COAC Jardin Azuayo faces the crucial task of analyzing the logs generated by border security tools, such as firewalls. These devices produce large volumes of data, the management and analysis of which are made difficult by the limited storage capacity of current devices, which are not designed for this purpose.

There are internal and external regulations that require the conservation of historical data, with the need to guarantee the availability of logs, in accordance with institutional policies. This project seeks to contribute to the implementation of effective measures, optimizing data analysis time through the application of machine learning techniques.

The proper preparation and exploitation of these data allow not only the extraction of key information, but also the development of predictive models and machine learning techniques, fundamental elements to foster innovation in fields such as science and technology.

This thesis presents an analysis using a logistic regression algorithm that allows a prediction based on the severity of the attacks recorded in the logs of the security tools. Evaluations of the results were carried out to estimate the effectiveness of the model. The results showed that the model correctly classified the attacks into different levels of severity, as reflected in the performance metrics of each class. This indicated that the grouping of behaviors according to severity was adequate.

Dedicatoria

Dedico este trabajo con todo mi amor y cariño a los seres más maravillosos de este mundo mi hermosa familia, mi esposa e hijos, son el motor y ocupan mi corazón en la vida, de igual manera a mis queridos padres quienes con su esfuerzo velaron siempre por mi formación y me dieron los valores necesarios; juntos ayudaron a dar rumbo a mi vida y cumplir estos grandes objetivos. Su influencia ha sido la fuerza impulsora detrás de muchos de mis logros, y siempre estaré agradecido por su amor incondicional.

Este logro es también suyo, y cada página de esta tesis lleva impresa mi gratitud hacia cada uno de ellos por hacer posible este sueño.

Culmina una meta más, pero en el camino de seguro vendrán otras.

Agradecimiento

Primero dar gracias a Dios, seguido quiero expresar mi más sincero agradecimiento a mi amada esposa, Johanna Flores, cuya paciencia, amor y apoyo incondicional han sido mi refugio y fortaleza, a lo largo de este desafiante proceso. Su comprensión y motivación han sido esenciales para mantener mi enfoque y energía.

A mis amados hijos, Romina Abigail y Joan Matias, quienes con su hermosa presencia y alegría diaria me recordaron la importancia de perseguir mis sueños y la responsabilidad de ser un ejemplo para ellos. Su entusiasmo y curiosidad por el mundo me impulsan a ser mejor cada día.

También debo un profundo agradecimiento a mis queridos padres, Oscar Cando y Ernestina Jara, por su amor eterno y las enseñanzas de vida que han sembrado en mí desde niño. Su apoyo emocional y espiritual ha sido un pilar fundamental en mi vida y especialmente durante mi viaje académico.

A mis queridos hermanos, por darme el apoyo y respaldo en todo el tiempo que ha durado este proceso, siempre han tenido una palabra para llenarme de motivación.

A mis apreciados docentes, quienes han sido fuente de inspiración, especialmente a mi director de tesis Dr. Eduardo Montero, gracias por su apoyo, paciencia y colaboración durante el desarrollo de este trabajo.

CAPITULO I: GENERALIDADES

1.1 Introducción

Las amenazas en el ámbito de la *ciberseguridad*, han tenido un impacto considerable en las organizaciones y usuarios, debido a que son cada vez más complejas, metódicas y sofisticadas, con mucha más trascendencia en los últimos años, el desafío que se presenta es grande por lo que es importante estar preparados y aprovechar estrategias efectivas para mitigar estos riesgos.

En este contexto en la institución financiera cooperativa de ahorro y crédito Jardín Azuayo, existe entre las actividades una muy relevante que es el análisis de *logs* generados por herramientas de seguridad de frontera. Estos *logs* de seguridad registran eventos provenientes de la actividad del *firewall*, dispositivos gestionados por el departamento de Seguridad Informática de la institución. Estos dispositivos registran *data* en gran volumen, lo que dificulta su almacenamiento en el dispositivo actual, debido a que sus funciones son otras y a su limitada capacidad de almacenamiento.

Existen regulaciones internas y directrices de organismos de control externos que exigen la garantía de al menos los últimos meses de datos, conforme a una política institucional. En este contexto, se busca mejorar la generación de medidas efectivas por parte del área de Seguridad Informática, reduciendo el tiempo necesario para analizar estos datos mediante procesos convencionales. Para lograrlo, se pretende implementar técnicas de *machinelearning* que optimicen y agilicen el análisis.

Las organizaciones actuales están cada vez más convencidas de que los datos son uno de los activos más valiosos. Su adecuada preparación y uso no solo permiten la extracción de información crucial, sino que también facilitan el desarrollo de modelos predictivos y técnicas de aprendizaje automático. Es por esta razón que este proceso resulta fundamental en el desarrollo de proyectos de investigación, con el objetivo de fomentar la innovación en diversos campos, incluyendo la ciencia y la tecnología.

1.2 Justificación

En la institución, se disponen de dispositivos críticos como *firewalls* y *WAF* en la frontera, los cuales generan una gran cantidad de información sobre el tráfico y posibles ataques. Esta información se almacena en un tercer dispositivo en forma de registros (*logs*), los cuales son datos no estructurados de gran volumen. A través del proceso ETL (*Extract, Transform, Load*), se busca estructurar y almacenar estos datos en una base de datos *PostgreSQL*, asegurando su integridad y autenticidad.

El procesamiento actual de los datos es bajo demanda debido a que su análisis toma demasiado tiempo por lo que se aísla a casos puntuales; para cumplir requerimientos o realizar seguimientos sobre sospechas o alertas presentadas. Se debe mejorar este proceso para aportar en su administración, brindando una respuesta más efectiva, y permitiendo adoptar varias contramedidas frente a amenazas cambiantes y constantes en el ámbito de *ciberseguridad*.

La aplicación de técnicas de *machine learning* nos permitirá descubrir características en el conjunto de datos, mejorando la comprensión de las variables y sus relaciones. Adicionalmente, podremos validar nuestra hipótesis y a su vez, asistiremos al departamento de seguridad.

Este proyecto también tiene como objetivo apoyar la realización de otras actividades dentro de la institución financiera. A través de los datos estructurados obtenidos, se podrá complementar los seguimientos operativos necesarios para el procesamiento de transacciones financieras generadas por los clientes, que deben ser verificados por los organismos de control interno.

1.3 Hipótesis

Algunas direcciones *IP* de origen, atacan repetidamente las mismas direcciones *IP* de destino.

Variable dependiente: direccionIPdestino (dirección IP fija asignada por aplicativo)

Variables independientes: direccionIPorigen, paisOrigen (Diferentes direcciones *IP* y países)

Palabras clave: dirección *IP*, herramienta de seguridad, *fortiAnalyzer*, extracción, transformación y carga (*ETL*), base de datos, *machine learning*, *firewall* y del *WAF web application firewall*, volumen de datos.

CAPITULO II: FUNDAMENTOS TEÓRICOS

2.1 Planteamiento del problema

Objetivo de Estudio

Comportamiento de ataques – SOCIEDAD

Campo de Acción

Ataques registrados desde ubicación geográfica.

Problema Científico

Contexto general: *Logs* de ataques en herramienta de seguridad, técnicas de *machine learning*, *(Firewall, WAF)*

Contexto Particular: Tráfico / Comportamiento de ataques, aprendizaje automático, minería de datos, visualización.

Problemática: volumen de datos no estructurados, repositorio limitado de almacenamiento, minimizar tiempos de análisis de tareas manuales, prevención, actualmente no existen reportes, cumplimento normativo.

Hipótesis /QR: ¿Algunas *IP* de origen atacan repetidamente las mismas *IP* de destino?

¿La mayoría de los ataques provienen de un país específico?

Propuesta: Aplicación de técnicas de *machine learning*, y visualización de datos en el análisis de *logs* de herramienta de seguridad perimetral *firewall*, de la institución COAC Jardín Azuayo.

2.2 Contextualización del tema

Aplicación de técnicas de *machine learning* y visualización de datos, que permiten identificar patrones en *logs*, generados por la herramienta de seguridad *firewall*, para correlacionar, clasificar y

analizar eventos de posibles ataques y generar listas negras o guiar otras acciones; minimizando tiempos de análisis de tareas manuales, aportando en la detección preventiva de manera proactiva.

2.3 Objetivos

Objetivo General

Desarrollar de un modelo predictivo de detección de ataques a herramientas de seguridad perimetral de la COAC Jardín Azuayo.

Objetivos Específicos

- Cargar los *logs* de seguridad a un repositorio de almacenamiento que garantice la disponibilidad de la *data*.
- Preprocesar los *logs* de la herramienta de seguridad *firewall* y *WAF*.
- Identificar patrones en los datos en función de la ubicación geográfica, dirección *IP* de origen, entre otras variables.
- Agrupar comportamientos de acuerdo con el nivel de severidad de los ataques.

2.4 Marco teórico y conceptual

Los ataques a nivel mundial han crecido exponencialmente de hecho, en América Latina y particularmente en Ecuador se ha incrementado en los últimos años hasta en un 24% según Forbes (2021). Varios de estos ataques se han dado debido a la pandemia, que generó el incremento en el uso de la tecnología, el teletrabajo, la educación virtual, transacciones electrónicas, etc.

Las técnicas de *machine learning* están siendo muy utilizadas actualmente en diferentes ámbitos entre ellos la *ciberseguridad*, permitiendo tomar acciones en beneficio de las organizaciones.

En la institución, se desea aprovechar estos beneficios, lo cual ha generado una necesidad específica en el departamento de seguridad. El objetivo es facilitar el análisis centralizando los datos actuales, que están en formato de *logs*, son no estructurados y de gran volumen. Mediante estrategias que

incluyen el proceso *ETL*, se pretende estructurar estos datos sin alterar su integridad y almacenarlos en un repositorio de base de datos que posteriormente permita la aplicación de técnicas de machine *learning*.

2.4.1 Marco Teórico

La ciberseguridad es una disciplina en constante evolución, caracterizada por la presencia de amenazas persistentes y cambiantes. La respuesta efectiva a estas amenazas es crucial, lo que implica disponer de herramientas adecuadas y optimizar las contramedidas basadas en datos precisos y fáciles de procesar. Para los usuarios, es fundamental que estas soluciones sean tanto eficaces como accesibles en términos de costo.

Los dispositivos perimetrales realizan su proceso de control basado en reglas, las cuales son cruciales para la monitorización, administración eficiente y uso efectivo de recursos. Para lograr la optimización de estas reglas y aumentar su eficacia, es necesario realizar análisis de datos tanto en tiempo real como históricos.

El análisis exploratorio de datos (EDA) permitirá descubrir características en el conjunto de datos, entendiendo mejor las dimensiones, variables, sus relaciones ocultas permitiendo visualizarlas y diseñar un modelo predictivo que nos permita hacer pruebas de nuestra hipótesis; asistiendo al departamento de seguridad a realizar acciones y mitigar riesgos comprobados con datos más concretos.

2.4.2 Marco Conceptual

Incremento de ataques a ciertos aplicativos de la institución financiera

El incremento en el uso de tecnología trae consigo un crecimiento de ataques a nivel mundial, todas las organizaciones deben invertir en *hardware* y *software* para mitigar estos riesgos. La institución financiera no es ajena a ello y recibe muchos ataques a diferentes aplicativos, para lo cual se toman

medidas de seguridad, pero debido al volumen de datos que se generan, toma mucho tiempo manual realizar exploración de la *data*para poder tomar acciones estratégicas.

2.4.2.1 Dirección *IP*

Una dirección *IP* es una dirección única, exclusiva que identifica un dispositivo ya sea en Internet o en una red local. La abreviatura "*IP*" corresponde a "protocolo de Internet", el cual establece las normas para la estructura de los datos transmitidos en Internet o en la red local.

2.4.2.1.1 Direcciones IP privadas

Todos los dispositivos que se conectan a la red de Internet están asignados con una dirección *IP* privada, abarcando desde computadoras hasta teléfonos, *tablets* y otros dispositivos. En la institución tenemos segmentación de red, lo que nos permite tomar medidas de control de acceso.

2.4.2.1.2 Direcciones IP públicas

Una dirección *IP* pública es la dirección principal vinculada a toda la red. Aunque cada dispositivo conectado tiene su propia dirección *IP*, también están incluidos en la dirección *IP* principal de la red. El proveedor de servicios de Internet (*ISP*) suministra la dirección *IP* pública del enrutador. Los ISP poseen direcciones IP que distribuyen a sus clientes. La dirección IP pública es la que los dispositivos externos a la red utilizarán para identificarla.

2.4.2.2 Seguridad de TI

De acuerdo con *IBM*, la seguridad informática, también conocida como seguridad de la tecnología de la información, consiste en resguardar los sistemas informáticos, redes, dispositivos digitales y datos contra accesos no autorizados, filtraciones de información, ataques cibernéticos y otras actividades maliciosas.

2.4.2.3 Ciberseguridad

Según *Cisco*, esta disciplina se define como la práctica de resguardar sistemas, redes y programas contra ataques digitales. Estos ciberataques generalmente buscan acceder, modificar o destruir

información confidencial, extorsionar a los usuarios o interrumpir la continuidad del negocio. La creciente implementación de medidas de seguridad digital se atribuye al aumento de dispositivos conectados en comparación con la población y a la creatividad en constante evolución de los atacantes.

2.4.2.4 Ciberdefensa

La ciberdefensa engloba un conjunto de estrategias activas, proactivas, preventivas y reactivas. El objetivo de estas acciones es asegurar la integridad y el uso seguro del ciberespacio de un estado específico, protegiéndolo contra amenazas de ciberdelincuentes y enemigos en el ámbito bélico.

2.4.2.5 Herramientas de seguridad

2.4.2.5.1 Firewall

Según Cisco, un *firewall* en una red informática desempeña un papel fundamental al proporcionar seguridad en el perímetro. Supervisa tanto los paquetes de datos entrantes como salientes en el tráfico de red para detectar *malware* y anomalías. Se trata de un dispositivo de seguridad que examina el tráfico de red y toma decisiones sobre permitir o bloquear tráfico específico, según un conjunto definido de reglas de seguridad.

Además, establecen una barrera efectiva entre las redes internas, que son seguras y controladas, y las redes externas, como Internet, en las que la confianza puede ser cuestionable. Pueden adoptar diversas formas, como *hardware*, *software*, servicios basados en *software* (*SaaS*), nube pública o nube privada (virtual).

2.4.2.5.2 Web Application Firewall WAF

Según *fortinet*, un *firewall* de aplicaciones *web* (*WAF*) se configura como un tipo de cortafuegos diseñado para resguardar aplicaciones *web* y *APIs*. Su función principal radica en filtrar, monitorear

y bloquear el tráfico *web* peligroso, así como proteger contra ataques a la capa de aplicación, tales como *DDoS*, inyección *SQL*, manipulación de *cookies*, ataques de scripts en sitios cruzados (*XSS*), falsificación de sitios cruzados e inclusión de archivos.

Dentro de la defensa en la Capa 7 (Aplicación), los *WAF* se enfocan en el tráfico que transita entre las aplicaciones *web* a internet. Su capacidad para identificar y responder a solicitudes maliciosas antes de que las aplicaciones y los servidores *web* las acepten se convierte en un componente de seguridad crucial para las empresas y sus clientes.

Además, el WAF proporciona una mayor seguridad para las aplicaciones desplegadas y expuestas en producción, las cuales cuentan con certificados. Sin embargo, en el ámbito de la seguridad, es crucial realizar una tarea constante de actualización periódica de políticas y planificación de estrategias.

2.4.2.6 Amenazas a la seguridad informática

Todas las organizaciones enfrentan amenazas cibernéticas tanto internas como externas. Estas amenazas pueden ser intencionales, involucrando a delincuentes cibernéticos, o no intencionales, como cuando empleados o contratistas hacen *click* inadvertidamente en enlaces maliciosos o descargan *malware*.

La Seguridad Informática propone abordar esta amplia variedad de riesgos de seguridad, considerando diversos actores de amenazas, sus motivaciones, tácticas y niveles de habilidad.

2.4.2.7 *Log* Informática

Los *logs* son documentos que registran eventos particulares dentro de un sistema; estos eventos abarcan desde actividades cotidianas hasta errores críticos, ofreciendo una pista de información valiosa sobre el rendimiento de un sistema o aplicación en un momento específico.

2.4.2.7.1 Logs de seguridad

Registros vinculados a la seguridad del sistema que registran intentos de acceso no autorizado, modificaciones en permisos y otras actividades de seguridad.

2.4.2.7.2 Monitorización de logs

La supervisión de *logs* se ha vuelto fundamental en el entorno tecnológico actual, permitiendo generar contramedidas o acciones oportunas. Implica la recopilación, análisis y alerta en tiempo real de eventos registrados en los *logs*.

Estos *logs* se pueden seguir, pero la herramienta actual en la institución financiera carece de funcionalidades para el análisis de datos y la implementación de modelos predictivos.

2.4.2.8 Ciencia de datos

De acuerdo con *IBM*, la ciencia de datos fusiona conceptos de matemáticas y estadística, programación especializada, análisis avanzados, inteligencia artificial (*IA*) y aprendizaje automático para revelar información valiosa oculta en los datos de una organización. La información permite tomar de decisiones y la planificación estratégica.

La institución financiera tiene un departamento de análisis de datos, las organizaciones demandan de estos roles, ya que son cruciales en interpretar datos y ofrecer recomendaciones prácticas para mejorar los resultados comerciales.

2.4.2.9 *Python*

Lenguajes de programación como *python* nos facilitan llevar a cabo análisis exploratorios de datos y regresiones estadísticas, estas herramientas de código abierto ofrecen funciones pre integradas para gráficos, *machine learning* y creación de modelos estadísticos.

Python destaca como el más popular y ampliamente utilizado en comparación con otros lenguajes, gracias a su naturaleza dinámica y flexible. También cuenta con numerosas bibliotecas, como *NumPy, Pandas y Matplotlib*, que facilitan el análisis rápido y eficiente de datos.

2.4.2.10 *Datawarehouse*

Según Oracle, un almacén de datos desempeña la función de centralizar y fusionar extensas cantidades de datos provenientes de diversas fuentes. Los *datawarehouses* han sido diseñados específicamente para llevar a cabo consultas y tareas de análisis, y suelen contener grandes volúmenes de datos históricos. La información almacenada en un *datawarehouse* proviene con frecuencia de una variada gama de fuentes, como archivos de registro de aplicaciones o sistemas transaccionales.

2.4.2.11 Extracción, transformación y carga de datos (*ETL*)

Según la descripción de *Microsoft*, la Extracción, Transformación y Carga (*ETL*) es un proceso de canalización de datos utilizado para reunir información de diversas fuentes. Después, se transforman los datos según las reglas de negocio establecidas, para cargarlos en el almacén de datos designado. La fase de transformación en *ETL* se lleva a cabo mediante un motor especializado y a menudo implica el uso de tablas de almacenamiento temporal para preservar los datos durante su transformación antes de ser cargados en su destino final.

2.4.2.12 *Machine learning*

El aprendizaje automático constituye una rama fundamental de la inteligencia artificial (*IA*) y la informática, focalizada en la utilización de datos y algoritmos para simular el proceso de aprendizaje humano y mejorar gradualmente su precisión.

En el ámbito de la ciencia de datos, el aprendizaje automático desempeña un papel crucial, a través de métodos estadísticos, los algoritmos se entrenan para realizar clasificaciones o predicciones, así como para descubrir patrones clave en proyectos de minería de datos. Estos conocimientos posteriormente influyen en la toma de decisiones en aplicaciones y empresas, impactando positivamente en las métricas de crecimiento.

2.4.2.13 Análisis exploratorio de datos

El análisis exploratorio de datos, con sus siglas en inglés *EDA* es una técnica utilizada para examinar y resumir conjuntos de datos. Según *IBM*, los científicos de datos emplean el *EDA* para analizar e investigar conjuntos de datos, resumiendo sus características principales mediante métodos frecuentes de visualización de datos. El *EDA*ayuda a los científicos de datos a descubrir patrones, identificar anomalías, probar hipótesis o validar suposiciones.

El *EDA* se usa para descubrir qué datos pueden revelarse más allá de las tareas formales de modelado o las pruebas de hipótesis. Facilita una comprensión más profunda de las variables en el conjunto de datos y las relaciones entre ellas. Además, posibilita evaluar la idoneidad de las técnicas estadísticas consideradas para el análisis de datos.

2.4.2.14 Análisis predictivo

Según *IBM*, el análisis predictivo representa una rama de los análisis avanzados que realiza predicciones sobre resultados futuros mediante la combinación de datos históricos con modelado estadístico, técnicas de extracción de datos y aprendizaje automático.

Para extraer información significativa de estos datos, los científicos de datos recurren a algoritmos de aprendizaje automático y aprendizaje profundo. Estos algoritmos se utilizan para identificar patrones y realizar predicciones sobre eventos futuros. Entre las técnicas aplicadas se encuentran modelos logísticos y de regresión lineal, redes neuronales y árboles de decisión.

2.4.2.15 Artículos científicos

Según los artículos científicos revisados, los ataques en América Latina, especialmente en Ecuador, han aumentado considerablemente. Es crucial tomar medidas e invertir en seguridad, aunque a menudo se ve limitado por restricciones presupuestarias debido a los altos costos de estas implementaciones. A veces, estas medidas no se consideran tan relevantes hasta que ocurren eventos tanto internos como externos que afectan a la institución.

De acuerdo al trabajo de investigación de ataques Ingrid, C.(2021) las *ciberamenazas* son riesgos que deben ser abordados en las instituciones dentro de su infraestructura tecnológica, para garantizar la disponibilidad de los servicios que brinda.

El aprendizaje automático ofrece la posibilidad de abordar de manera efectiva varios ámbitos, incluyendo la seguridad informática y la prevención de ciberataques. Mediante un análisis efectivo de datos generados tanto históricos como en tiempo real, se busca contribuir a la implementación de contramedidas eficientes y proactivas.

Según Merterun (2023), es crucial cubrir las vulnerabilidades mediante un examen profundo que permita identificar aspectos críticos, descubrir patrones y tendencias, así como entender qué productos están involucrados y cómo pueden influir en el entorno. Es fundamental comprender estos aspectos adecuadamente para formular acciones defensivas efectivas y evitar exponerse a riesgos potenciales.

En las investigaciones realizadas, se recomienda realizar un análisis exhaustivo, libre de sesgos y replicable, utilizando capacidades avanzadas de diversas opciones tecnológicas para el análisis y la visualización de datos. Se sugiere realizar pruebas de diseño e implementación de varias técnicas de machine learning, lo cual permitirá validar y profundizar en la investigación, así como mejorar la comprensión de los datos.

CAPITULO III: DESARROLLO E IMPLEMENTACIÓN

Metodología y técnicas

3.1 Diseño de la Investigación

Usé el método o enfoque cuantitativo, lo que me permitió, mediante análisis exploratorio, validar la hipótesis.

El proyecto fue experimental cuantitativo, abordando todo el ciclo de vida de los datos. Posteriormente, realicé el tratamiento usando CRISP-DM, cubriendo sus fases para la minería de datos (comprensión del negocio, comprensión de los datos, preparación de los datos, modelado, evaluación y despliegue), lo me permitió cumplir con todo el ciclo del proyecto de minería de datos. Esto fue crucial para cubrir los objetivos empresariales, desde la fase de recolección, comprensión, preparación y creación del modelo de datos, sin descuidar la evaluación mediante métricas de ciencia de datos, asegurando que realmente estuviera alineado con los objetivos de la institución y, finalmente, su despliegue en la empresa.

Planeé el experimento utilizando variables independientes X para tratar de demostrar la correlación con la variable de salida Y y sus posibles efectos sobre esta.

Apliqué *machine learning* para tratar de deducir si, desde el mismo origen, se generaban varios eventos de conexión o ataque que resultaban en un nivel de severidad. Esto podría indicar un ataque con intentos de acceso por fuerza bruta, generando un patrón.

3.2 Características de los experimentos

- **Manipulación intencional de una o más variables independientes**

En la fase de "preparación de datos", dejé las variables categóricas utilizando *one hot encoding*(transformando variables categóricas en múltiples columnas binarias) y *ordinal encoding* (implica un orden entre las categorías). Esto me ayudó a prevenir el sesgo que otras técnicas podrían generar.

- **Medición de las variables dependientes.**

La medición de la variable dependiente debe ser adecuada, válida y confiable, mediante correlación y xorroboración de causalidad.

- **Control sobre la situación experimental.**

Esto me permitió, conjunto con los especialistas de seguridad, detectar si las variables independientes afectaban a la variable dependiente y si los ataques estaban catalogados con la severidad adecuada.

3.3 Selección de la muestra

Para el caso planteado, adopté un muestreo probabilístico, tomando una muestra representativa aleatoria del conjunto de datos. Conté con los datos de los últimos dos meses, garantizando el cumplimiento de la política establecida, que asegura la conformidad con las normativas internas y externas.

Usé, entre otras herramientas, la función *sample()* de *pandas*. En el análisis de datos y el aprendizaje automático, es crucial asegurar que los resultados sean reproducibles. Por ello, utilicé el parámetro *random_state*, para tomar muestras aleatorias reproducibles.

Mediante los instrumentos de validación, utilicé la ciencia de datos para detectar el comportamiento de los datos en los *logs* de seguridad y realicé su respectivo tratamiento.

3.4 Metodología

Apliqué técnicas de *machine learning* y visualización de datos en el complejo mundo de la ciberseguridad, proporcionando *insights* clave para anticipar, responder y mitigar las amenazas de manera más efectiva.

- Recolecté los datos, inicialmente muy amplios, que correspondían alrededor de 53 columnas del *log*. Trabajé en darle un formato adecuado, siguiendo las políticas establecidas en conjunto con el equipo de Seguridad Informática.

- El formato establecido fue para almacenar los datos en archivos csv. Reduje las columnas a 16, que fueron las consideradas relevantes por el equipo de Seguridad Informática. Programé el envío de estos archivos por correo cada hora y, posteriormente, los almacené en OneDrive.
- El archivo csv generado se aplicó a un proceso de extracción, transformación y carga (*ETL*) hacia un repositorio de bases de datos PostgreSQL.
- Realicé una conexión a la base de datos accediendo a las tablas que contienen los datos, para obtener un muestreo de estos.
- Realic un análisis exploratorio de datos (*EDA*), permitiendo descubrir características en el conjunto de datos, aportando una mejor comprensión de los datos, sus relaciones ocultas y permitiendo visualizarlas.
- Diseñé un modelo predictivo que me permitió probar la hipótesis, asistiendo al departamento de Seguridad Informática en la realización de acciones y mitigación de riesgos, basándome en datos mucho más concretos.
- Según los resultados, realicé el diseño e implementación de modelos de aprendizaje automático.
- Evalué el rendimiento del modelo utilizando métricas relevantes, como precisión, recall y F1-score.
- Establecí consideraciones éticas en todas las fases del proyecto para garantizar el cumplimiento ético, incluyendo la privacidad y el manejo de información sensible. Esto facilitó la realización de un trabajo transparente mientras se protegían y resguardaban los datos críticos que no debían ser divulgados. Además, evité la inferencia o el sesgo en los datos que podrían llevar a resultados poco confiables.

- Anonimicé los datos proporcionados para el análisis, cumpliendo con la ley de protección de datos y las políticas internas de la institución.

CAPITULO IV: RESULTADOS Y DISCUSIÓN

Usando *machine learning*, obtuve información valiosa de los conjuntos de datos y desarrollé un modelo predictivo que permitió identificar patrones anómalos en los *logs* de herramientas de seguridad. Esto mejoró los tiempos y la capacidad de detección temprana de posibles amenazas. Además, proporcioné una evaluación detallada de la efectividad del modelo, comparándola con los análisis manuales realizados por los especialistas de seguridad utilizando enfoques tradicionales.

Verifiqué el error si se trataba de varias conexiones del mismo origen y de la misma IP, utilizando criterios de análisis exploratorio de datos en varios escenarios y una muestra probabilística representativa y aleatoria. Visualicé y documenté estos datos para tratar de demostrar si los ataques provenían de un país específico.

Realicé un contraste de hipótesis para validar la hipótesis nula o alternativa, aceptando o rechazando la hipótesis de acuerdo con diferentes factores, incluso cambiantes.

Durante todo el proceso de desarrollo del proyecto, busqué asegurar que hubiera retroalimentación constante. Esto permitió mejorar en cada una de las etapas, garantizando que todo el proceso fuera útil para la institución y que contribuyera a alcanzar los objetivos empresariales.

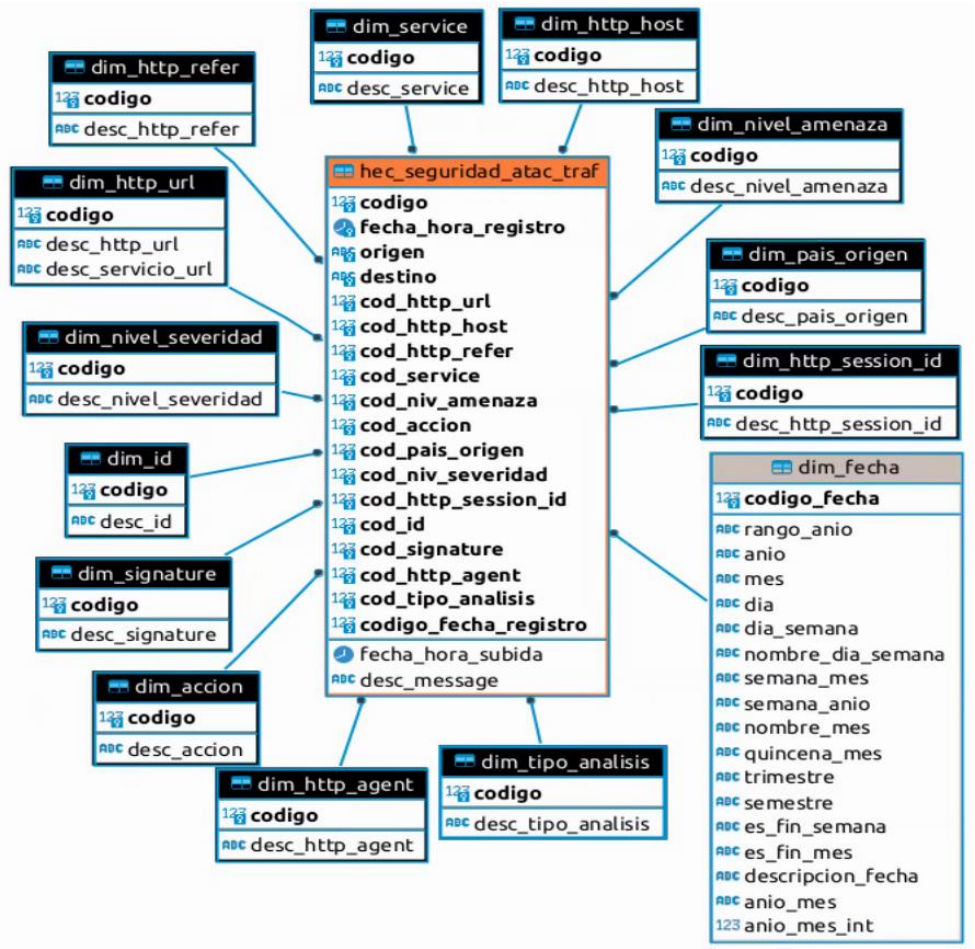
4.1 Modelo de datos

Desarrollé un modelo estrella que se creó en la base de datos en el ambiente de *data warehouse*. Esta estructura permitió almacenar los datos de seguridad, los cuales se cargaron desde los *logs* hacia el motor de base de datos.

Creé una tabla central de hechos llamada *hec_seguridad_atac_traf* y sus diferentes dimensiones: servicio, detalles de *host*, nivel de amenaza, país de origen, detalles de la sesión, dimensión de fechas, tipo de análisis, datos del agente, acción, firma, nivel de severidad, referencia *HTTP y URL*.

Ilustración 1

Modelo de datos de logs



4.2 Carga de datos

Proceso ETL (Extracción, Transformación y Carga)

Mediante el uso de herramientas como *Pentaho Data Integration*, los datos catalogados que se encuentra en formato csv (*Comma-separated values*), elabore el *ETL* para la carga de logs a la base de datos PostgreSQL.

Configuración del Entorno

- Configure *Pentaho Data Integration (PDI)*
- Creé un nuevo esquema dedicado para el proyecto en *PostgreSQL*

Diseño del Proceso ETL en PDI

Extracción (*Extract*)

- Creé una nueva transformación en *PentahoDataIntegration* (*Spoon*).
- Añadí un paso de *CSV Input*.
- Configuré el paso para que leyera el archivo *CSV* proporcionando la ruta del archivo.
- Definí las propiedades del archivo, como el delimitador (coma) y las comillas.
- Especifiqué las columnas y sus tipos de datos.

Transformación (*Transform*)

Limpieza de datos

- Añadí pasos como *select values*, para seleccionar y renombrar columnas.
- Utilicé pasos como *replace in string* o *string operations*, para limpiar o transformar los datos, como eliminar caracteres no deseados, formatear fechas, etc.

Validación de datos

- Implementé pasos de validación para asegurarme de que los datos cumplieran con los requisitos del destino *DataValidator*.
- Filtré registros inválidos utilizando pasos como *filter rows* y los redirigí a un archivo de error.

Carga (*Load*)

Conexión a PostgreSQL

- Configuré la conexión a la base de datos *PostgreSQL*, proporcionando los detalles necesarios (host, puerto, nombre de la base de datos, usuario y contraseña).

- Probé la conexión para asegurarme de que podía conectar a la base de datos.

Carga de datos

- Configuré el paso *tableoutput* para insertar los datos en la tabla de destino en *PostgreSQL*.
- Especifiqué la tabla de destino y mapeé las columnas de origen con las columnas de destino.
- Configuré opciones adicionales como insertar filas o actualizar filas existentes, según las necesidades.

Ejecución y Monitoreo

Ejecución de la transformación

- Guardé la transformación y la ejecuté en *Spoon* para verificar que los datos se cargaban correctamente en *PostgreSQL*.

Automatización y programación

- La transformación necesitaba ejecutarse regularmente, por lo que creé un trabajo (*job*) en *Spoon*.
- Añadí el paso de transformación al trabajo.
- Configuré el trabajo para que se ejecutara según un cronograma usando el paso *Scheduler*.
- Optimicé del rendimiento, para que la carga de datos estuviera optimizada para manejar grandes volúmenes de datos.

Manejo de Errores

Gestión de errores

- Configuré pasos de gestión de errores para capturar registros que no se podían insertar en *PostgreSQL*.

- Redirigí estos registros a una tabla específica para su revisión.

4.3 Almacenamiento en PostgreSQL

En el esquema *jdw_seguridades*, creé las tablas planteadas en la elaboración del modelo de datos. Estas tablas contienen datos estructurados y reducidos mediante catálogos, manteniendo y precautelando su integridad.

A pesar de las estrategias de optimización de uso en almacenamiento, la tabla de hechos tienen un tamaño de 29Gb con miles de registros almacenados.

Table 1

Tabla del esquema de seguridad y su tamaño

NOMBRE TABLA	**PESO**
hec_seguridad_atac_traf	29G
dim_id	2,6G
dim_message	409M
dim_http_url	361M
dim_http_refer	4,3M

4.4 Conjunto de datos

Trabajé con un conjunto de datos en un rango de fechas del 03 de marzo al 10 de marzo del 2024, que posee 807.964 instancias y 20 variables, revisaremos sus detalles a continuación.

Debido al cumplimiento de la protección de datos y por recomendación de seguridad informática, excluí el campo *destination* en la presentación de los datos. Este campo es crítico debido a que contiene información sensible, trabajé bajo esta premisa para asegurar la privacidad y seguridad de los datos.

El conjunto de datos es del *firewall*, equipo de seguridad de frontera, presenta características fundamentales que permite conocer las dimensiones de los ataques y el comportamiento de los diferentes aplicativos, presenta un volumen de datos considerable para el análisis sobre el objetivo y la severidad de los ataques.

Table 2

Variables del conjunto de datos

Nombre	Descripción
date_time	Nos indica la fecha y hora del evento, permitiendo conocer y comprender las tendencias.
source	Característica que nos permite identificar la dirección IP "protocolo de Internet" del origen, permite conocer el epicentro de las actividades maliciosas.
destination	Característica que nos permite identificar la dirección IP "protocolo de Internet" del destino e identificar los aplicativos dentro de estos.
http_url	Almacena el código de la información de la URL, haciendo referencia a la URL accedida en el evento.
http_host	Almacena el código de la información de host, podemos tener claramente identificado el host

	involucrado en el ataque.
http_refer	Almacena la referencia de la url, apuntador de referencia describiendo la URL accedida.
service	Almacena el servicio que se usa para conexión, podría ser http o https.
threat_level	Característica que almacena el nivel de amenaza de un ataque, describe pistas de la gravedad del ataque.
action	Almacena la acción tomada por el equipo, indicadores como las contramedidas adoptadas de acuerdo con el plan de acción.
src_country	Almacena el pais de origen de la conexión, permitiendo mapear geográficos los epicentros de los ciberataques.
severity_level	Almacena el nivel de severidad, describe pistas de la severidad que provoca el ataque, es la variable para predecir.
http_session_id	Almacena la identificación de http de la conexión.
signature_id	Almacena la firma de la conexión, mediane las firmas únicas podemos conocer el esquema de las amenazas cibernéticas.
agent	Almacena el URL del agente de conexión,

	podemos identificar el agente la URL usada en el evento.

4.4.1 Análisis variable objetivo, SEVERITY_LEVEL

En esta sección se realizó un análisis de la proporción de los tres niveles de severidad de los ataques en función de las fechas disponibles. A continuación, se muestra la distribución total:

Table 3

Niveles de severidad de ataques

Nivel de Severidad	**Total**
Bajo	62.299
Medio	718.286
Bajo	27.379

Se analizó junto con el equipo de seguridad informática la opción de convertir la variable severity_level a binaria, la sugerencia fue mantener los tres rangos (*High, Medium, Low*), ya que nos indicaron que lo mejor es no generalizar.

Datos totales del mes de marzo 2024

Table 4

Severidad totales de marzo 2024

Nivel de Severidad	**Total**
Bajo	985.352
Medio	8.631.102
Alto	361.323

Table 5

Análisis de datos únicos y nulos

Nombre	**Código**	**Total**
cod_service	1	9.977.777
cod_http_agent	1	9.977.777
desc_message	NA	9.977.777
cod_signature	1	8.994.606

Con los datos obtenidos de la base de datos, realicé un análisis para conocer su estado en el último mes. Observé que el volumen de datos es alto y, tras el filtrado, obtuve un *top* de niveles medio, bajo y alto. Con estos resultados, el equipo de seguridad informática tuvo más insumos para revisar y tomar las acciones correspondientes.

Además, analicé los datos de manera general para preparar análisis posteriores y asegurar que el modelo puede tener un buen *performance*. Identifiqué características de valor único que no aportan

valor y, tras revisarlas en conjunto con el equipo de seguridad, confirmamos que la presentación de los datos era correcta.

Datos de severidad del mes, total por día

Agrupé la suma de las cantidades de ataques por severidad y por fecha. A continuación,se muestran las cantidades para los últimos días de marzo 2024.

Table 6

Severidad del mes de marzo de 2024 por día

Fecha	Severidad	Cantidad
29/3/2024	1	9.387
29/3/2024	2	21.811
29/3/2024	3	185.718
30/3/2024	1	11.466
30/3/2024	2	28.650
30/3/2024	3	211.359
31/3/2024	1	11.823
31/3/2024	2	27.747
31/3/2024	3	234.224

Agrupé los datos para analizar la severidad de los ataques durante el mes. Observé un comportamiento con los tres tipos de severidad, siendo el código 3 el que representa la severidad más alta, catalogada como tipo *medium*.

Table 7

Análisis de severidad del mes de marzo de 2024

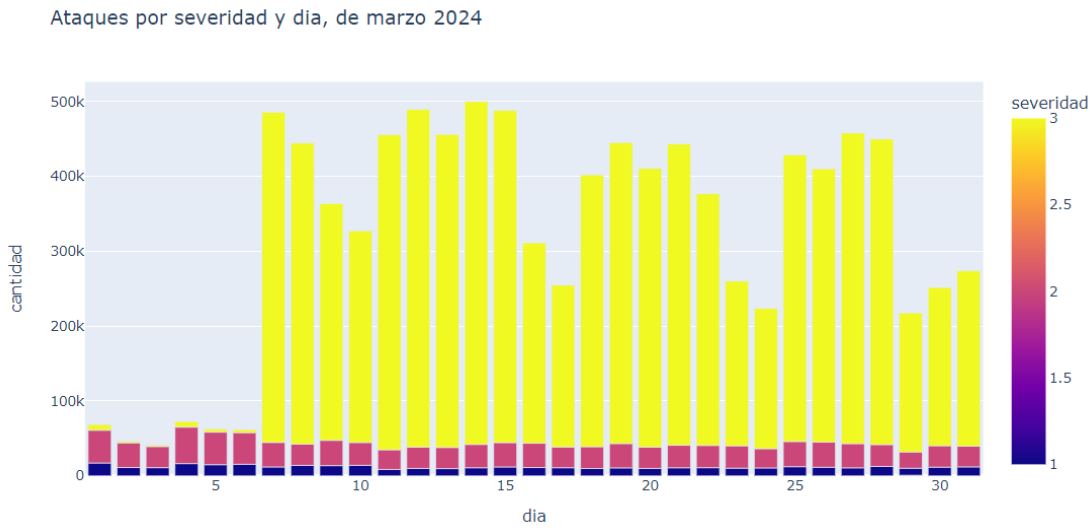
Severidad	Cantidad
1	361.323
2	985.352
3	8.631.102

Generé un gráfico de barras interactivo para ayudar en la interpretación de los datos, mostrando la severidad de los ataques por día y cantidad. Además, se pudo evidenciar que, en marzo, desde el día 7 en adelante, la severidad de tipo 3 varió considerablemente.

Se revisó con el equipo de seguridad que los cambios se presentan debido a modificaciones en nuevo aplicativo, por lo cual si bien son altos, se concluyó que de momento son adecuado. Estos datos les permitirá realizar los ajustes correspondientes.

Ilustración 2

Ataques por severidad y día, del mes de marzo del 2024



4.5 Tratamiento de los datos

Evalué si existían valores nulos en el conjunto de datos y no encontré ninguno. Esto se debió a todo el flujo previo desde sus fases de *ETL*, con la generación de catálogos y el cuidado del ciclo de vida de los datos.

4.5.1 Columnas irrelevantes

Eliminé las columnas que no aportaban valor, incluyendo las columnas categóricas con un solo nivel y las columnas numéricas con un solo valor. Las columnas categóricas con más de un subnivel no fueron eliminadas.

Table 8

Análisis de variables irrelevantes

Nombre	**Subniveles**
date_time	73.674
source	14.619

http_session_id	7.417
http_url	4.570
http_refer	.2053
src_country	27
signature_id	11
threat_level	6
destination	4
action	3
severity_level	3
htpp_host	2
resgistration_date	2
agent	1
service	1

De toda la *data* cargada, observé que algunas características tenían valores únicos por lo tanto, no aportaban valor significativo. Realicé esta revisión en conjunto con el equipo de seguridad informática para asegurar su correcta aplicación.

Posteriormente, eliminé las columnas no relevantes, y creé un nuevo conjunto de datos preservando únicamente con las variables relevantes.

Table 9

Resultados de variables relevantes

Nombre	**Subniveles**
date_time	73.674

source	14.619
http_session_id	7.417
http_url	4.570
http_refer	2.053
src_country	27
signature_id	11
threat_level	6
destination	4
action	3
severity_level	3
htpp_host	2
resgistration_date	2

4.5.2 Características generales de las variables

En la siguiente sección, analicé las variables por separado, presentando cada característica y su conteo correspondiente. La frecuencia obtenida me permitió una mejor interpretación de los datos.

Pude analizar la cantidad de datos por valor de categoría y sus porcentajes. Los ordené para conocer los valores más frecuentes, dependiendo del volumen de datos en cada característica.

Según el análisis, determiné que todas las variables son categóricas; no tenemos variables numéricas. Además, contabilicé la cantidad de filas, la cantidad de categorías y la frecuencia de estas. No coloqué todas las variables, solo aquellas que nos ayudan a cumplir nuestros objetivos. Las demás estarán en el anexo.

4.5.2.1 Origen

La variable que indica la IP de origen del ataque es una característica con muchos valores. Realicé un conteo para identificar las direcciones IP más frecuentes en el origen de los ataques, destacando

el top de direcciones IP.

Table 10

Top de direcciones IP del origen del ataque

IP Origen	Cantidad	Porcentaje
131.196.114.117	5.320	0,006584
157.100.61.154	3.291	0,004073
201.182.241.1	3.275	0,004053
45.189.58.22	2.695	0,003336
186.66.130.220	2.145	0,002655
200.24.159.38	2.057	0,002546
157.100.3.105	1.861	0,002303
157.100.68.139	1.853	0,002293
45.188.229.1	1.851	0,002291
128.201.161.75	1.804	0,002233

4.5.2.2 Source Country

El país de origen de los ataques nos proporcionó un buen indicio para crear listas negras. Evidencié que el porcentaje más alto de ataques provenía de Ecuador y, para una mejor comprensión, catalogué los países por nombre en lugar de utilizar códigos.

Table 11

Análisis de ataque por país de origen

País Origen	**Cantidad**	**Porcentaje**
Ecuador	722.198	89,38
Reserved	44.801	5,54
United States	31.788	3,93
Spain	3.823	0,47
Italy	1.055	0,13
Colombia	891	0,11
Chile	777	0,09
United Kingdom	724	0,09
Peru	623	0,08
Otros	928	0,15

4.5.2.3 http_host

Variable que permite identificar los *hosts* más atacados. Realicé un conteo de los *hosts* más accedidos y registré su frecuencia. Esto me permitió identificar cuáles fueron los objetivos más frecuentes de los ataques y evaluar la necesidad de reforzar la seguridad.

Se encontró un porcentaje alto de accesos en el host jarvi2, y en mucha menor medida en el host loja1. Esto permitió al equipo de Seguridad Informática tomar acciones sobre los aplicativos que están en estos *hosts*.

Table 12

Análisis de hosts atacados

Http_host	Cantidad	Porcentaje
jarvi2	715.660	88,58
loja1	92.304	11,42

4.5.2.4 Threat level

El nivel de amenaza me permitió conocer cómo se han presentado los ataques y si esto coincide con lo esperado. Además, revisé en conjunto con los especialistas de Seguridad Informática si se estaban tomando las acciones adecuadas en respuesta a estos ataques.

Observé que, según el número de casos y su frecuencia, la mayoría de las amenazas son de nivel dos (substancial), seguidas de niveles informacional, moderado, crítico, bajo y severo. Esta información es muy útil para tomar contramedidas adecuadas.

Table 13

Análisis del nivel de amenaza

Nivel Amenaza	Cantidad	Porcentaje
Substancial	739.097	91,48
Informational	60.692	7,51
Moderate	4.586	0,57
Critical	2.143	0,27
Low	1.378	0,17
Severe	68	0,008

4.5.2.5 Action

La acción es la contramedida ejecutada en los diferentes eventos de ataque. Este análisis me ayudó a conocer la tendencia y a revisar si era adecuada o si necesitaba ajustes.

De la información analizada, observé que la mayoría de las acciones eran de tipo *alert* con un 91%, seguidas de un 7.6% de acciones tipo *Erase* y, finalmente, un mínimo de 0.8% de acciones tipo *alert deny*.

Table 14

Análisis de la acción ejecutada

Acción	**Cantidad**	**Porcentaje**
Alert	738.967	91,46
Erase	62.070	7,68
Alert_Deny	6.927	0,86

4.5.2.6 Severity level

El nivel de severidad es una característica clave, como vimos anteriormente es nuestra variable objetivo, acá la analizamos por separado.Tenemos una escala de valores acorde al nivel severidad del ataque, clasificada por *'High', 'Medium', 'Low'*.

Del análisis de la *data* podemos ver la tendencia a eventos de severidad media en el *top*, seguido de baja y finalmente alta.

Table 15

Análisis del nivel de severidad

Nivel de Severidad	**Cantidad**	**Porcentaje**
Medium	718.286	88,9

Low	62.299	7,71
High	27.379	3,39

4.5.2.7 Signature id

El *Signature ID* es una característica muy importante que almacena la firma de la conexión. Al analizar las firmas únicas, pude comprender mejor el esquema de las amenazas cibernéticas.

Esto proporcionó información valiosa para mejorar las defensas y desarrollar estrategias más efectivas contra posibles ataques. Identificar y entender estas firmas permitió al equipo de seguridad anticipar patrones de amenazas y fortalecer las medidas de protección.

Table 16

Análisis de la firma de conexión

Firma Conexión	**Cantidad**	**Porcentaje**
1	745.014	92,2
2	60.692	7,51
3	1.378	0,17
5	720	0,09
28	79	0,009
38	34	0,004
39	34	0,004
31	9	0,001
30	2	0,0002
8	1	0,0001

4.5.3 Visualización de variables

Realicé un análisis detallado por variables, identificando comportamientos y tendencias relevantes. Para optimizar el código, creé funciones que mejoraron la eficiencia y reutilización.

Presenté varias opciones de visualización al equipo de Seguridad Informática y, finalmente, diseñé gráficos interactivos. Grafiqué las características más relevantes y aquellas que seguían buenas prácticas, teniendo en cuenta el volumen de datos y la diversidad de instancias por variable.

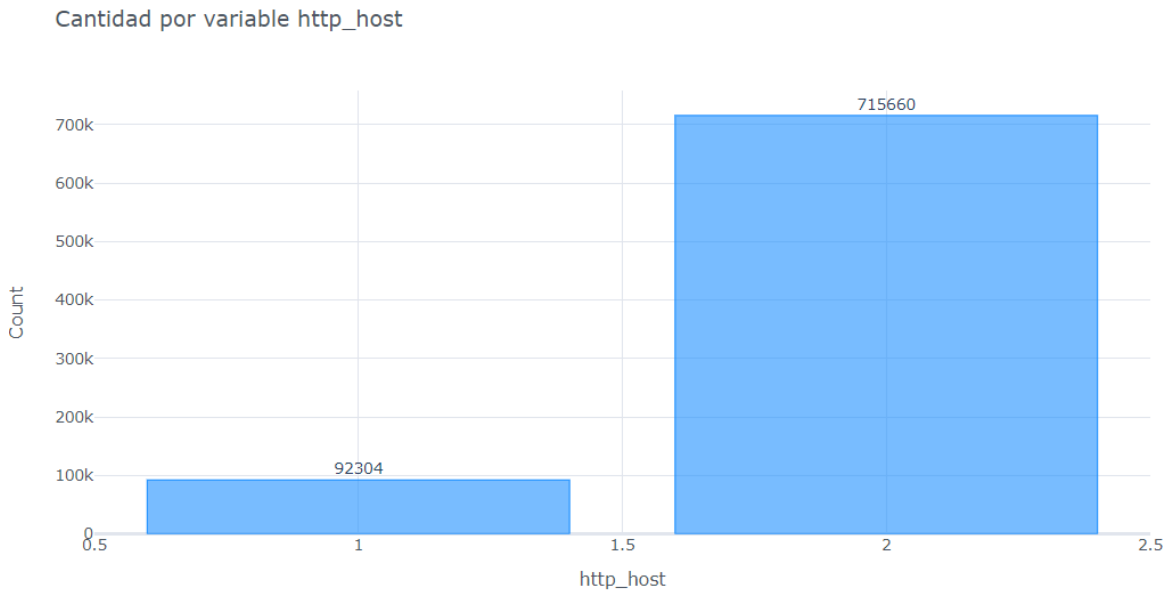
4.5.3.1 http_host

Este gráfico de la variablede *host* fue muy relevante, ya que me permitió visualizar los ataques por *host*, permitiendo identificar rápidamente los más atacados.

Obtuve un porcentaje alto de accesos en el *host jarvi2*, y en mucha menor medida en el *host loja1*. Esto permitió al equipo de Seguridad Informática tomar acciones sobre los aplicativos que estaban en estos hosts.

Ilustración 3

Ataques por host



4.5.3.2 threat_level

Realicé un análisis gráfico del nivel de amenaza y personalicé el gráfico para que se adaptara a las necesidades de Seguridad Informática. De esta manera, pude revisar las tendencias de los niveles de amenaza de forma gráfica, facilitando la identificación de patrones y hosts más afectados.

Table 17

Tabla del nivel de amenaza

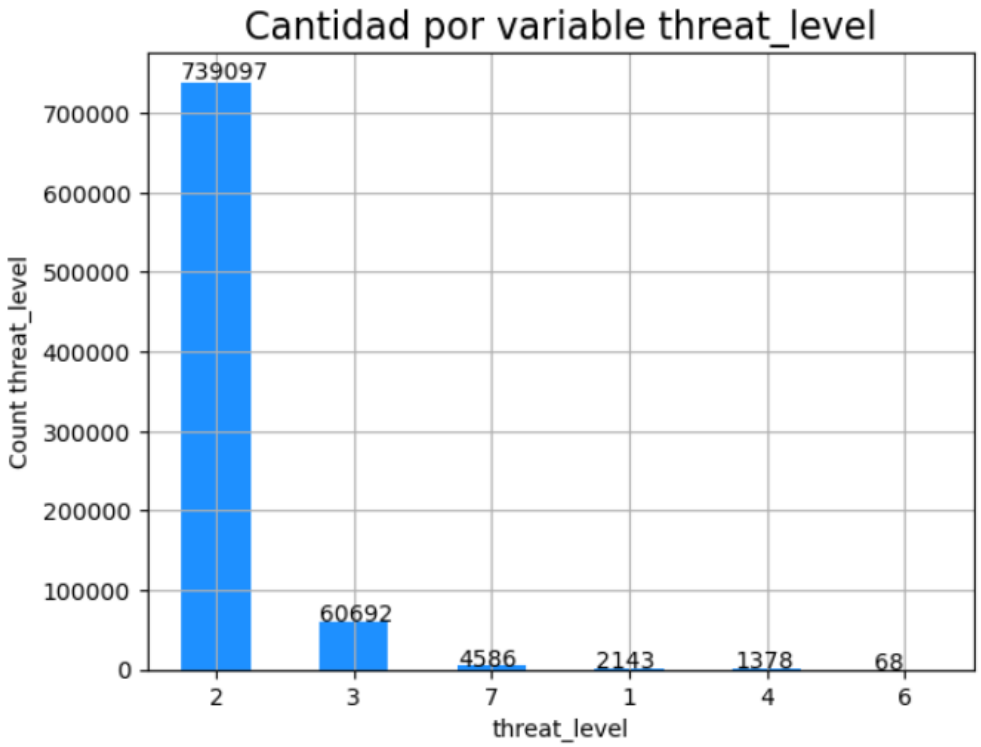
Código	Nivel de Amenaza
1	Critical
2	Substancial
3	Informational
4	Low
5	NA

6	Severe
7	Moderate

Del análisis anterior se observó que la tendencia estaba marcada en el código 2 "Substancial" y en menor medida en los otros niveles de amenaza.

Ilustración 4

Tendencias según el nivel de amenaza



4.5.3.3 Action

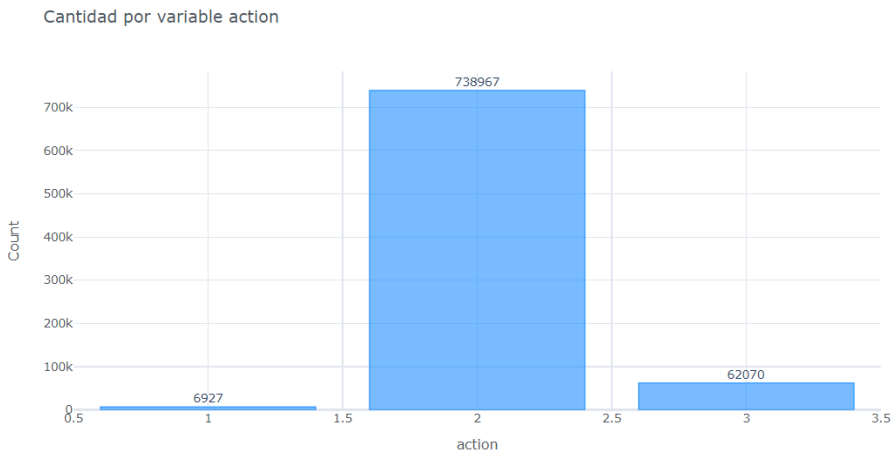
Realicé un análisis de las contramedidas adoptadas en función del ataque. Este proceso me permitió evaluar la efectividad de las medidas implementadas. Al revisar los datos, pude identificar qué acciones fueron más efectivas contra tipos específicos de ataques y optimizar las defensas para futuras incidencias.

Podemos contrastar el análisis anterior donde la tendencia estaba en acción "*Alert*" y en menor

medida en las otras acciones. Referencia de códigos *(1 “Alert_Deny”,2 “Alert”,3 “Erase”).*

Ilustración 5

Tendencias según la acción



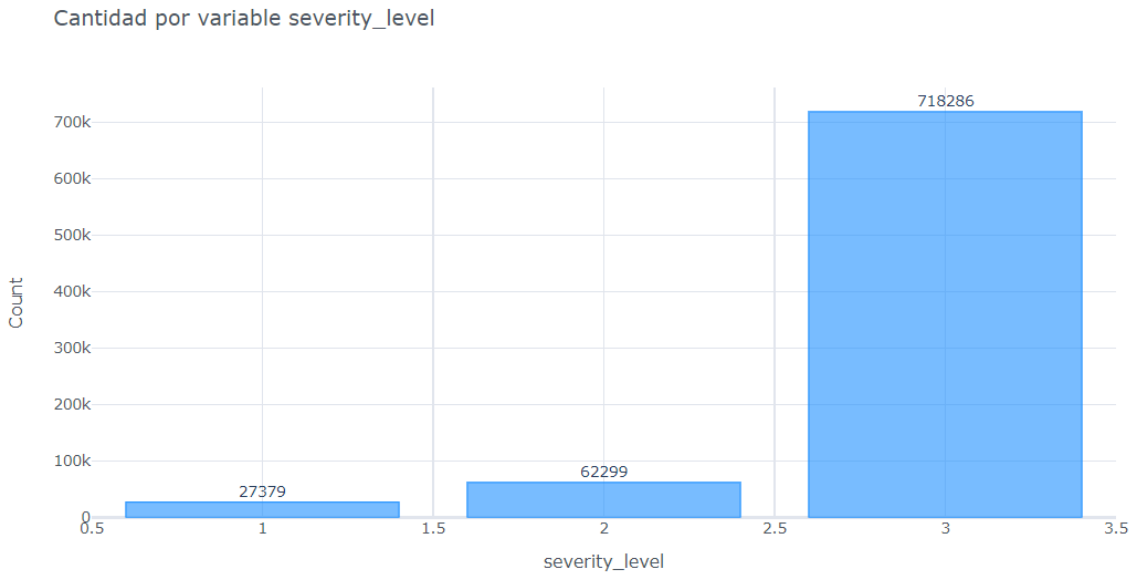
4.5.3.4 severity_level

En el análisis preliminar de la variable objetivo *"severity_level"*, revisé la tendencia de los datos. Realicé una visualización para el análisis del nivel de severidad, lo que me permitió identificar patrones clave y comprender mejor la distribución de los niveles de severidad.

De manera gráfica, pude contrastar lo anterior, donde la tendencia mostró que los eventos de severidad "media" estaban en el top, seguidos de los de severidad "baja" y, finalmente, los de severidad "alta". Esta visualización me permitió identificar claramente la distribución de los niveles de severidad y enfocar los esfuerzos de mitigación.

Ilustración 6

Tendencias según el nivel de severidad



4.5.3.5 Análisis de tendencias

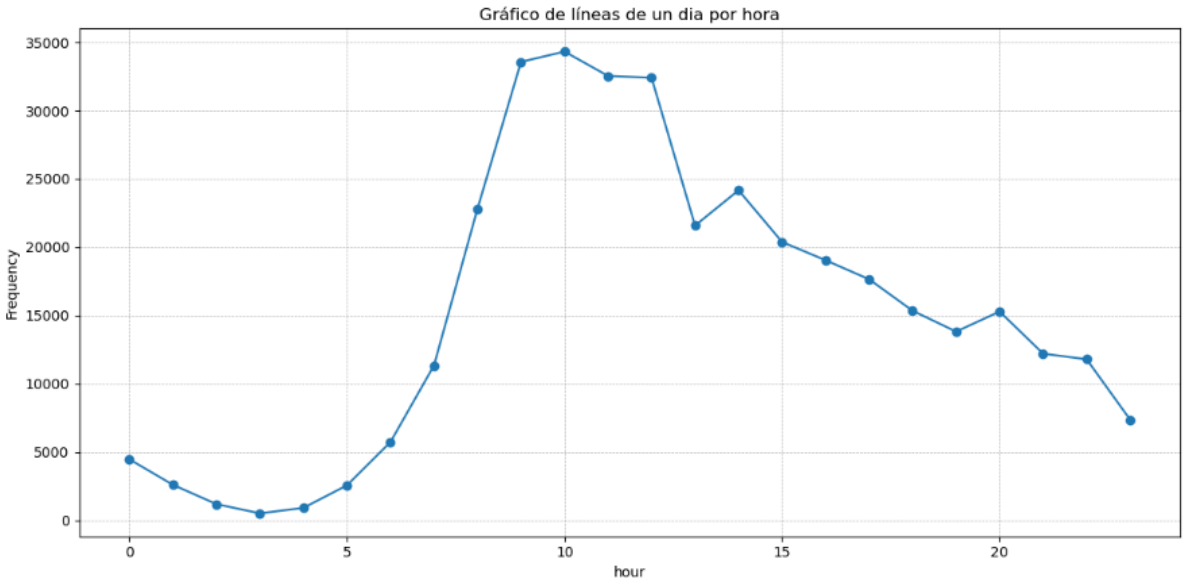
Para un análisis más profundo por fecha, fue necesaria la creación de nuevas características. Agregué características descriptivas para la fecha, incluyendo año, mes, día, hora, entre otras.

Estas me permitieron profundizar el análisis y generar más indicios para alcanzar los objetivos. Adicionalmente, proporcionaron datos que no se podían obtener del análisis anterior.

La siguiente gráfica de líneas fue de mucha utilidad, ya que me permitió conocer los ataques y su frecuencia por hora. Esto resultó muy útil para identificar patrones en los ataques. Observé que el día 09 de marzo del 2024, desde las 08:00 hasta las 11:30, se registró una mayor frecuencia de ataques.

Ilustración 7

Ataques y su frecuencia de un día por hora



Generé una tabla *pivote* que me permitió, obtener los valores en este caso por fecha. Tomé como índice el mes y como columnas el día, además utilicé una función de agregación para contar. De esta manera, pude presentar los ataques por mes y día.

Table 18

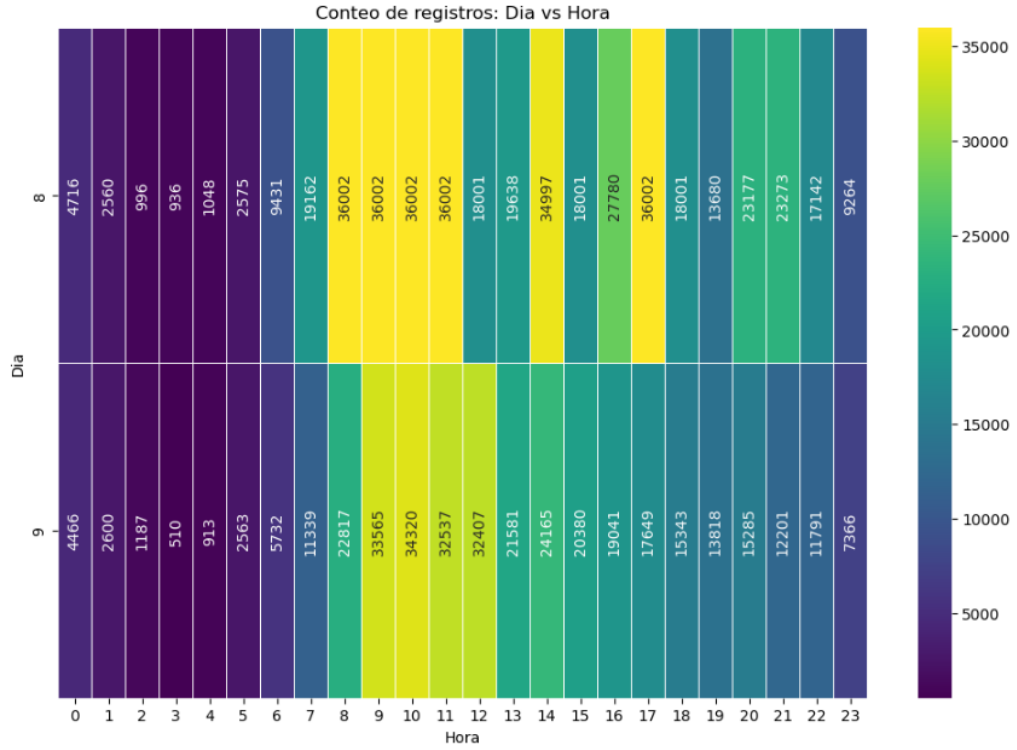
Ataques por mes y día

Día	Mes	Cantidad
8	Marzo	444.388
9	Marzo	363.576

Se puedo evidenciar las horas donde se presentaron más ataques, lo que permitió tomar acciones y ajustar la severidad de acuerdo con los casos.

Ilustración 8

Tendencias de ataques de dos días por hora



La gráfica tipo mapa de calor fue de mucha utilidad, ya que me permitió conocer los ataques por hora en un día en particular. Esto resultó muy útil para identificar dónde se concentraban los ataques y observar las horas de mayor frecuencia.

4.5.4 Análisis bivariado

Me permitió analizar la variable objetivo en relación con cada variable independiente.Utilicé tablas de contingencia para comprender mejor la relación de la severidad respecto a las variables independientes, mostrando la distribución de frecuencia de las categorías de una variable en relación con las categorías de otra variable.

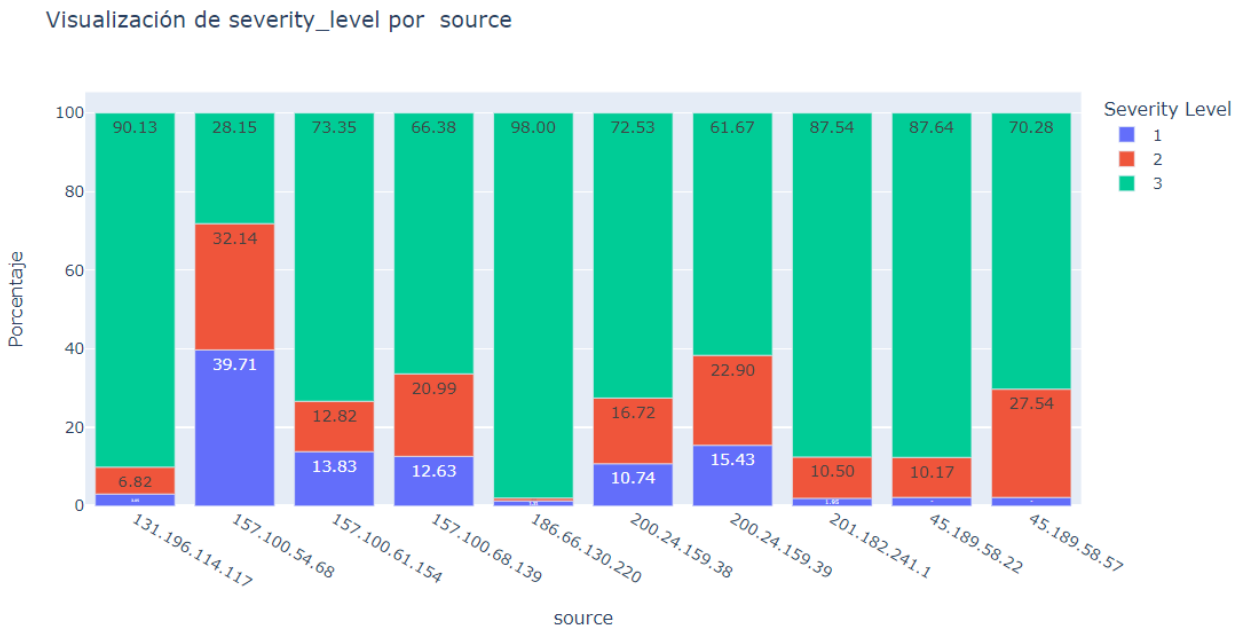
4.5.4.1 severity_level – source

Examiné los datos del origen de los ataques, específicamente desde las direcciones IP de origen, y pude identificar el top diez de IPs donde se generó la mayoría de los incidentes. Posteriormente,

clasifiqué estos incidentes según su nivel de severidad: “Alto”, “Medio” y “Bajo”. Para cada dirección IP del top diez, calculé y visualicé el porcentaje correspondiente a cada nivel de severidad, lo que me permitió entender mejor cómo se distribuyen los ataques en términos de gravedad desde cada fuente.

Ilustración 9

Nivel de severidad de los ataques por dirección IP de origen



4.5.4.2 severity_level - http_url

Con los datos de la variable *http_url*, pude identificar el top 10 de incidentes según su nivel de severidad, clasificados como “Alto”, “Medio” y “Bajo”. Posteriormente, realicé las visualizaciones correspondientes para ilustrar mejor la distribución de estos niveles de severidad.

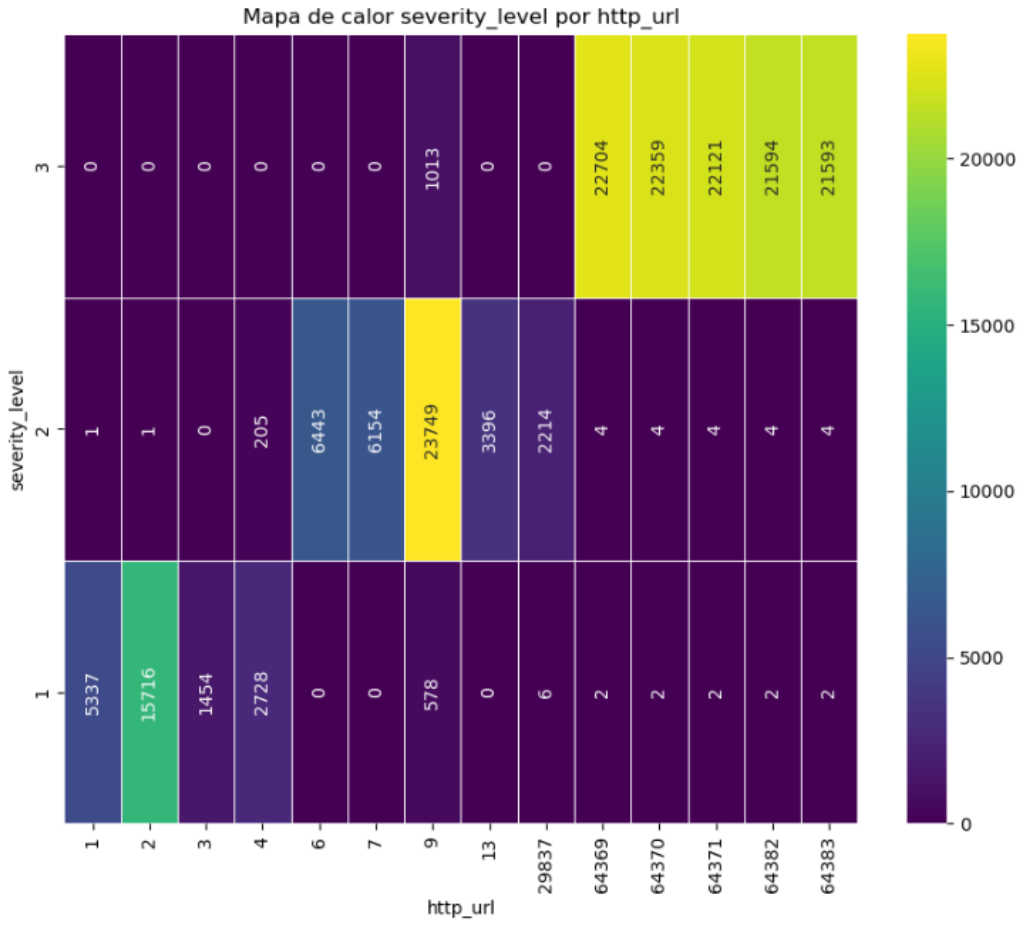
Generé un mapa de calor utilizando los datos de la variable “http_url” para observar las concentraciones de incidentes por nivel de severidad. Noté que la mayor frecuencia de incidentes se encontraba en el código 9 de *http_url* con un nivel de severidad bajo. Además, observé que había una notable concentración de incidentes en los códigos (64369, 64370, 64371, 64382, 64383) con

un nivel de severidad “3” medio.

Además, se encontró que la menor frecuencia de incidentes estaba en los códigos (1,2,3,4,6,7,13,29837) con un nivel de severidad 3 medio y también el código (9) que presentaban un nivel de severidad “Bajo (2)”.

Ilustración 10

Mapa de calor de la variable http_url por nivel de severidad



4.5.4.3 severity_level - http_host

En este análisis, identifiqué cómo varían los niveles de severidad en relación con las diferentes *http_host*. Observé que para el host con código 1 (“loja1”), el nivel de severidad “Bajo” era predominante, mientras que para el host con código 2 (“jarvi2”), los incidentes con nivel de

severidad “Medio” eran más frecuentes. Estos hallazgos me permitieron realizar visualizaciones específicas que mostraban claramente las tendencias y diferencias entre los hosts en relación con los niveles de severidad de los incidentes.

Table 19

Análisis por nivel de severidad y http_host

	http_host	
severity_level	loja1	jarvi2
High	29,41	0,03
Low	67,37	0,02
Medium	3,22	99,95

4.5.4.4 severity_level - threat_level

En mi análisis, extraje el top 10 de amenazas desde la variable *threat_level*, que clasifica las amenazas en siete categorías: 1: “Crítica”, 2: “Sustancial”, 3: “Informativa”, 4: “Baja”, 5: “No Aplicable”, 6: “Severa”, y 7: “Moderada”. Para cada una de estas categorías, evalué la distribución de incidentes según tres niveles de severidad: “Alto”, “Medio” y “Bajo”. Posteriormente, realicé visualizaciones para ilustrar la frecuencia y severidad de las amenazas en cada nivel de “threat_level”.

La figura muestra el gráfico de barras generado, donde se observó que:

- Para el *threat_level* 2, el 97.07% de las incidencias correspondieron a una severidad media (nivel 3), indicando que la mayoría de las incidencias en este nivel no eran críticamente urgentes.
- Se detectó una notable variación en la distribución de la severidad entre diferentes niveles de amenaza. Específicamente, los niveles de amenaza 3 y 4 presentaron una alta proporción de

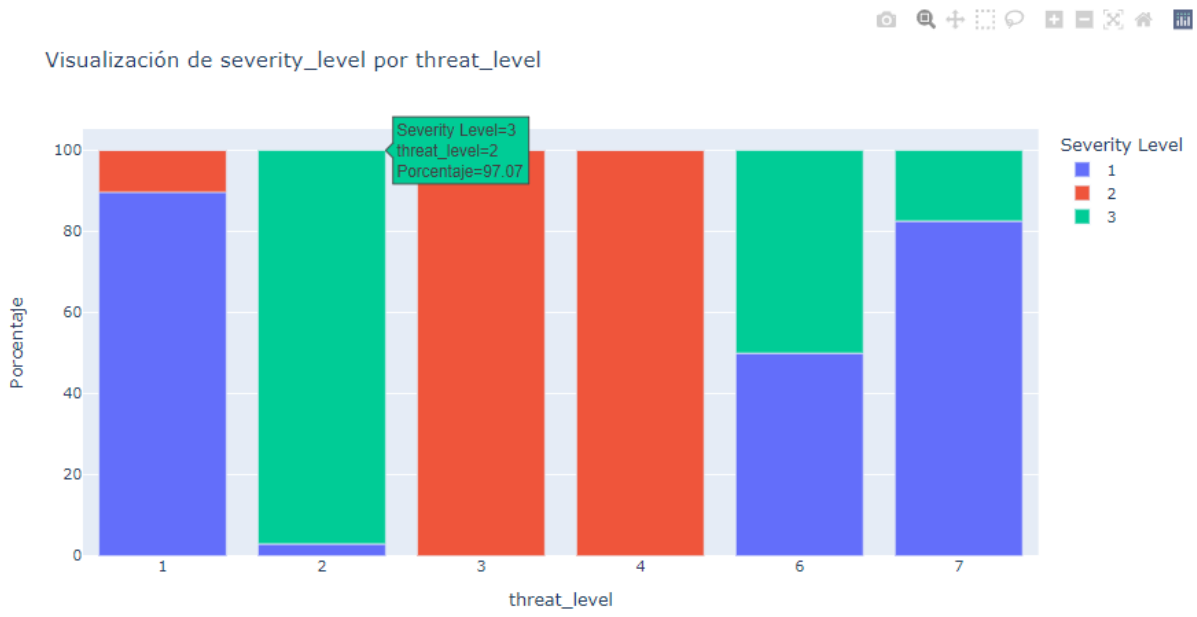
incidencias de severidad baja, mientras que el nivel 6 mostró predominancia del nivel de severidad media.

- Algunos niveles de amenaza, como el 5, no estaban representados en el gráfico, lo que podría sugerir la ausencia de incidencias bajo estos parámetros durante el período analizado.

Este análisis me permitió identificar tendencias específicas y patrones de severidad de las amenazas, lo que es crucial para entender dónde concentrar los esfuerzos de mitigación y respuesta.

Ilustración 11

Nivel de severidad de los ataques por nivel de amenaza



4.5.4.5 severity_level – action

Revisé los datos y obtuve el top 10 de acciones que incluían “Alert_Deny”, “Alert” y “Erase”, clasificadas respectivamente como 1, 2 y 3. Luego, evalué la distribución de estos eventos según los niveles de severidad: “Alto” (1), “Bajo” (2) y “Medio” (3). Tras recopilar y organizar la información, procedí a realizar visualizaciones para ilustrar claramente la relación entre las acciones

tomadas y los niveles de severidad asociados con cada una.

La figura presenta el gráfico de barras que ilustra la distribución porcentual de los niveles de severidad para cada acción:

- En la acción “Alert_Deny” (1), el 84.54% de las incidencias reportadas tenían un nivel de severidad “Alto”.
- La acción“Alert”(2) mostró una predominancia del nivel de severidad “Bajo”, con un pequeño porcentaje correspondiente a los niveles “Alto”.
- Para la acción “Erase” (3), se observó que todas las incidencias estaban clasificadas con un nivel de severidad “Medio”.

Este análisis me permitió identificar patrones y tendencias significativas en la respuesta a las amenazas, lo que es esencial para comprender la efectividad de las medidas de seguridad implementadas.

Ilustración 12

Nivel de severidad de los ataques por acción

4.5.4.6 severity_level - src_country

Examiné los datos relacionados con el origen de las amenazas, utilizando la variable “src_country”, que clasifica los países de origen de las amenazas en categorías numeradas, incluyendo 1 para “United States”, 2 “Ecuador”, 3 “Spain”, 6 “Reserved”, 13 “Italy”, 24 “Chile” y 25 “Colombia”. Posteriormente, clasifiqué las amenazas según los niveles de severidad: “Alto” (1), “Bajo” (2) y “Medio” (3). Después de organizar los datos, realicé visualizaciones que mostraban la relación entre el país de origen de las amenazas y su nivel de severidad.

La tabla de contingencia que generé reveló la frecuencia de cada combinación de país de origen y nivel de severidad. Un hallazgo notable fue que el valor más alto se registró entre el nivel de severidad “Medio” (3) y el país de origen “Ecuador” (2). La observación sugiere que Ecuador es una fuente significativa de amenazas de severidad media, lo que indicaría patrones específicos de seguridad o de ataques más relevantes de este origen.

Este análisis permitió obtener una visión detallada de cómo las amenazas de diferentes severidades están distribuidas geográficamente, proporcionando información valiosa para estrategias de Seguridad Informática y políticas de respuesta ajustadas a las características regionales de las amenazas cibernéticas.

Table 20

Tabla de contingencia, con la frecuencia de cada combinación del país de origen y nivel de severidad

src_country	1	2	3	6	13	24	25
severity_level							
1	1767	24977	134	250	12	19	41
2	756	61222	72	206	1	11	8
3	29265	635999	3617	44345	1042	747	842

Table 21

Tabla de contingencia, con la frecuencia de cada combinación del país de origen y nivel de severidad representada por porcentajes

src_country	1	2	3	6	13	24	25
severity_level							
1	5.56	3.46	3.51	0.56	1.14	2.45	4.6
2	2.38	8.48	1.88	0.46	0.09	1.42	0.9
3	92.06	88.06	94.61	98.98	98.77	96.14	94.5

4.5.4.7 severity_level - signature_id

Examiné los datos y obtuve el top 10 de las firmas de conexión a través de la variable *signature_id*, clasificando los incidentes según los niveles de severidad: “Alto” (1), “Bajo” (2) y “Medio” (3). Posteriormente, realicé visualizaciones para mostrar la relación entre las firmas de conexión y los niveles de severidad.

La tabla de contingencia que generé mostraba las frecuencias de cada combinación de nivel de severidad y firma de conexión. Un hallazgo notable fue que el valor más alto se observó en el nivel

de severidad “Medio” (3) para la firma de conexión 1. Este resultado indicaba que la firma de conexión 1 estaba predominantemente asociada con incidentes de severidad media, lo que sugería una tendencia de riesgos moderados que requerían vigilancia, pero no eran críticamente urgentes, según lo contrastado por Seguridad Informática.

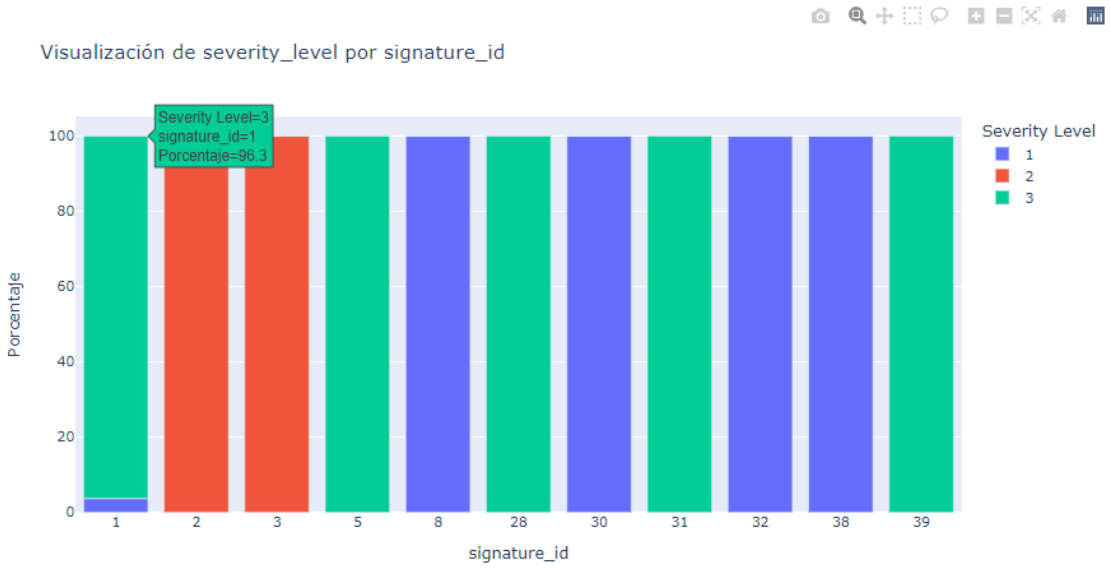
Table 22

Tabla de contingencia, con la frecuencia de cada combinación de firmas de conexión y nivel de severidad

signature_id	1	2	3	5	8	28	30	31	32	38	39
severity_level											
1	27341	0	0	0	1	0	2	0	1	34	0
2	229	60692	1378	0	0	0	0	0	0	0	0
3	717444	0	0	720	0	79	0	9	0	0	34

Ilustración 13

Nivel de severidad de los ataques por firmas de conexión



Este gráfico mostró la distribución del nivel de severidad asociado a diferentes identificadores de firma *signature_id*, clasificados en tres niveles de severidad: alto (1), bajo (2) y medio (3). Cada barra representaba un *signature_id* específico, y su color indicaba el porcentaje de cada nivel de severidad registrado para ese identificador.

Observé lo siguiente:

- En la mayoría de los *signature_id*, como los identificados por las barras para los números 1, 5, 28, 31, 39, noté una predominancia del nivel de severidad “Medio”. Esto indicaba que la mayoría de los incidentes detectados por estas firmas no eran extremadamente críticos, pero tampoco eran insignificantes.
- La severidad alta también estaba presente en algunos identificadores de firma, lo que sugería una tendencia hacia incidentes de gravedad alta en las detecciones realizadas por estas firmas.
- La firma con signature_id 1 mostraba un porcentaje muy alto (96.3%) para el nivel de severidad “Medio”. Esto podría indicar una especialización de esta firma en capturar amenazas o incidentes que, aunque no eran extremadamente urgentes, requerían atención y seguimiento. Corroboramos con los análisis anteriores que en el signature_id 1 concentra la cantidad más alta de valores.

Esta visualización fue útil para identificar qué firmas estaban detectando amenazas más graves y cuáles estaban más orientadas a amenazas de menor gravedad. Esta información fue crucial para ajustar las respuestas de seguridad y priorizar recursos donde más se necesitaban.

4.5.5 Modelos

Dado que el conjunto de datos de seguridad tenía una variable objetivo tipo categórica llamada "nivel de severidad", correctamente etiquetada como multiclase, utilicé un modelo supervisado de clasificación.

La regresión logística es un modelo de clasificación adecuado para problemas de clasificación multiclase (multinomial). Este modelo me permitió predecir la probabilidad de cada clase de severidad basándome en las características de entrada del conjunto de datos.

Usé bibliotecas como *scikit-learn* en *python*, que tenían implementaciones de este modelo. Realicé el modelo, llevando a cabo el proceso completo de entrenamiento, pruebas y evaluación, posteriormente determiné si era el más adecuado para este problema.

Las variables, según el análisis, son categóricas,buscando mejorar la precisión del modelo y su capacidad para hacer predicciones, realicé la codificación adecuada antes de la construcción del modelo.

Para las variables nominales, donde las categorías no tenían un orden intrínseco, utilicé *One-Hot Encoding*. Esto me ayudó a transformar las variables categóricas en múltiples columnas binarias. Cada categoría se convirtió en una característica separada, con valor 1 para indicar su presencia y 0 para su ausencia.

Para las variables ordinales, donde las categorías tenían un orden natural, utilicé *Ordinal Encoding*. La codificación ordinal asignó un valor entero único a cada categoría, implicando un orden específico entre ellas.

La codificación anterior era la recomendable, pero trajo consigo un problema de alta dimensionalidad. Para reducir la dimensionalidad y garantizar la elección de las variables adecuadas, consideré la frecuencia de cada variable como un parámetro dinámico dentro del código. Esto me permitió enfocarme en las variables más relevantes y reducir el número de columnas generadas por la codificación *One-Hot Encoding*, mejorando así la eficiencia del modelo y su capacidad para hacer predicciones precisas.

Utilicé la biblioteca *scikit-learn*de *python*, para dividir el conjunto de datos en subconjuntos de entrenamiento y prueba de manera eficiente. Esto ayudó a que el modelo pudiera ser evaluado adecuadamente con datos no vistos durante el entrenamiento.

Para lo cual, consideré los siguientes parámetros para el particionamiento del dataset:

- *test_size* = 0.3: Este parámetro especificó el tamaño del subconjunto de prueba, asignando el 30% de los datos para pruebas y el 70% para entrenamiento.
- *random_state* = 42: Este parámetro aseguró que la división de los datos fuera reproducible, estableciendo una semilla que garantizó que la misma división de datos se pudiera obtener en futuras ejecuciones del código.

4.5.5.1 Creación del modelo

Creé el modelo de regresión logística con las siguientes opciones:

$$LogisticRegression(multi\_class = 'multinomial', solver = 'lbfgs', max\_iter = 1000, random\_state = 42)$$

- **multi_class,** definió cómo el algoritmo debía manejar la clasificación multiclase. En este caso, “multinomial” indicó que el modelo utilizaría la regresión logística multinomial.
- **solver,** definió el algoritmo de optimización a utilizar. “lbfgs” fue el algoritmo de optimización de *Broyden-Fletcher-Goldfarb-Shanno* limitado.
- **max_iter,** definió el número máximo de iteraciones para el algoritmo de optimización.
- **random_state,** controló la semilla aleatoria utilizada para inicializar el estado interno del algoritmo, asegurando reproducibilidad.

Procedí a entrenar el modelo, realizar predicciones y evaluar su desempeño.

4.5.5.2 Entrenamiento del modelo

Utilicé el método fit para entrenar el modelo de regresión logística con los datos de entrenamiento.

- $log\_reg\_model.fit(X\_train\_logreg, y\_train\_logreg)$

4.5.5.3 Predicciones

Utilicé el método *predict* para realizar predicciones sobre el conjunto de prueba y el conjunto de entrenamiento.

$$y\_pred\_train\_logreg = log\_reg\_model.predict(X\_train\_logreg)$$

$$y\_pred\_test\_logreg = log\_reg\_model.predict(X\_test\_logreg)$$

4.5.5.4 Evaluación del modelo

Evalué el desempeño del modelo calculando la precisión y generando un reporte de clasificación para ambos conjuntos, de entrenamiento y de prueba. Esto me permitió entender qué tan bien estaba funcionando el modelo en ambos contextos.

Usé *accuracy_score*, el cual permitió calcular la precisión del modelo, que es la proporción de predicciones correctas sobre el total de predicciones. Con *classification_report*, generé un informe detallado que incluía métricas como precisión, *recall*, y *F1-score* para cada clase; en ambos conjuntos de datos.

Evaluación en el conjunto de entrenamiento:

$$accuracy\_train\_logreg = accuracy\_score(y\_train\_logreg, y\_pred\_train\_logreg)$$

$$reporte\_train\_logreg = classification\_report(y\_train\_logreg, y\_pred\_train\_logreg)$$

Evaluación en el conjunto de prueba:

$$accuracy\_test\_logreg = accuracy\_score(y\_test\_logreg, y\_pred\_test\_logreg)$$

$$reporte\_test\_logreg = classification\_report(y\_test\_logreg, y\_pred\_test\_logreg)$$

Estos pasos me permitieron entrenar el modelo de regresión logística, realizar predicciones con los datos de prueba y evaluar su desempeño de manera efectiva.

Posteriormenteevaluaremos su desempeño de los datos de evaluación y prueba en términos de precisión, *recall*, *F1-score* y exactitud.

4.5.5.5 Resultados:

Table 23

Métricas de evaluación del conjunto de entrenamiento

```
Accuracy en entrenamiento: 0.9960730160863123
Classification Report en entrenamiento:
              precision    recall  f1-score   support

           1       0.90      1.00      0.95     19284
           2       1.00      1.00      1.00     43640
           3       1.00      1.00      1.00    502650

    accuracy                           1.00    565574
   macro avg       0.97      1.00      0.98    565574
weighted avg       1.00      1.00      1.00    565574
```

Table 24

Métricas de evaluación del conjunto de pruebas

```
Accuracy en prueba: 0.9961095754775362
Classification Report en prueba:
              precision    recall  f1-score   support

           1       0.90      1.00      0.95      8095
           2       1.00      1.00      1.00     18659
           3       1.00      1.00      1.00    215636

    accuracy                           1.00    242390
   macro avg       0.97      1.00      0.98    242390
weighted avg       1.00      1.00      1.00    242390
```

Interpretación

4.5.5.5.1 Entrenamiento

Los resultados en el conjunto de entrenamiento muestran que el modelo tiene una alta precisión, *recall* y *F1-score* para todas las clases. La exactitud (*accuracy*) es de aproximadamente 99.6%, lo que indica que el modelo clasificó correctamente la mayoría de las muestras de entrenamiento. Los

valores de *macro* y *weightedaverage* también son altos, lo que sugiere un buen rendimiento general del modelo.

4.5.5.5.2 *Pruebas*

Los resultados en el conjunto de prueba son muy similares a los del conjunto de entrenamiento, con una exactitud (*accuracy*) también de aproximadamente 99.6%. Esto indica que el modelo generaliza bien y no está sobreajustado (*overfitted*) a los datos de entrenamiento. Las métricas de precisión, *recall* y *F1-score* para cada clase en el conjunto de prueba son casi idénticas a las obtenidas en el conjunto de entrenamiento, lo que refuerza la robustez del modelo.

En conclusión, el modelo de regresión logística entrenado mostró un excelente rendimiento tanto en los datos de entrenamiento como en los de prueba, con métricas de precisión, *recall* y *F1-score* muy altas en todas las clases. La alta precisión global en ambos conjuntos sugiere que el modelo es adecuado para este problema de clasificación multiclase y está bien equilibrado sin evidencias de sobreajuste.

Interpretación de resultados del conjunto de pruebas

El modelo predictivo presenta un buen rendimiento lo que indica que puede generalizar bastante bien, tiene una exactitud muy alta (99.61%); los resultados en precisión, *recall y F1-Score* casi perfectos en la mayoría de las clases, especialmente en las clases “2” y “3”.

La clase 1 “*High”* tiene una precisión ligeramente inferior (0.90) sin embargo, el *recall* perfecto para esta clase sugiere que todas las instancias verdaderas de la clase “1” fueron identificadas correctamente. Además, un *F1-Score* muy alto (0.95). Esto sugiere que el modelo es muy efectivo en la clasificación de las clases 2 y 3 y bastante efectivo en la clasificación de la clase 1, aunque con algunos falsos positivos.

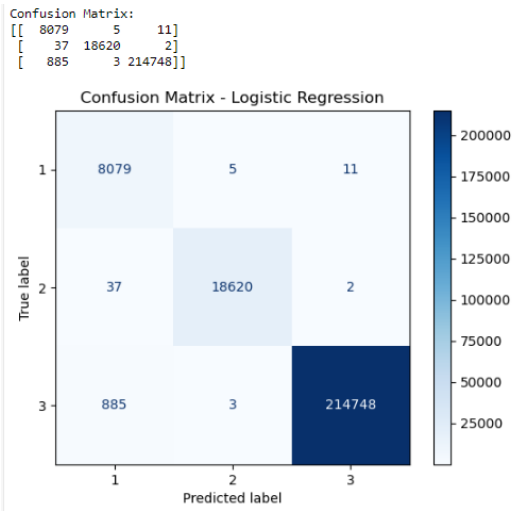
En general de acuerdo con las métricas de evaluación del modelo, parece ser muy efectivo, con un desempeño notablemente alto en todas las clases evaluadas.

4.5.5.6 Matriz de confusión del conjunto de pruebas

Realicé una matriz de confusión para analizar el rendimiento del modelo de clasificación. Esta matriz permitio revisar el desempeño del algoritmo mostrando el número de predicciones correctas e incorrectas, por cada clase.

Ilustración 14

Matriz de confusión del modelo de regresión logística, del conjunto de pruebas



4.5.5.6.1 Interpretación.

La matriz de confusión del modelo de regresión logística, cada celda en la matriz de confusión representa el número de instancias clasificadas en una clase particular. Esta matriz visualizó el rendimiento del modelo de clasificación multiclase comparando las etiquetas de clases previstas y reales para el conjunto de datos.

Etiquetas

1: Clase 1:**High**

2: Clase 2:**Low**

3: Clase 3:**Medium**

Table 25

Matriz de confusión

	Prediccion1	**Prediccion2**	**Prediccion3**
Etiqueta1	8.079	5	11
Etiqueta2	37	18.620	2
Etiqueta3	885	3	214.748

Esta matriz visualizó el rendimiento del modelo de clasificación multiclase, comparando las etiquetas de clases previstas y reales para el conjunto de datos.

El clasificador de regresión logística identificó correctamente los 8.079 casos de clase *high*, clasificó erróneamente 37 casos como *low*, al igual que 885 casos como *medium*.Mientras que clasificó correctamente 18.620 casos de clase *low*, clasificó erróneamente 5 como *high*, al igual que 3 casos como *medium*.Además, clasificó correctamente 214.748 casos de clase *medium*, clasificó erróneamente 11 como *high*, al igual que 2 casos como *low*.

Precisión

La métrica de precisión es utilizada para poder saber qué porcentaje de valores que se han clasificado como positivos son realmente positivos.

Recall (Sensibilidad)

La métrica de *recall*, también conocida como el ratio de verdaderos positivos, es utilizada para saber cuantos valores positivos son correctamente clasificados.

F1 Score

Esta es una métrica muy utilizada en problemas en los que el conjunto de datos a analizar está desbalanceado, combina la precisión y el *recall*, para obtener un valor mucho más objetivo.

Fórmulas:

- **Precisión (Precision)**

$$Precision = \frac{TP}{TP + FP}$$

- **Sensibilidad(Recall)**

$$Recall = \frac{TP}{TP + FN}$$

- **F1-Score**

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Resultados

- **Clase 1 High:**

  Precisión: 89.78%

  Sensibilidad: 99.80%

  F1-score: 94.55%

- **Clase 2 Low:**

  Precisión: 99.96%

  Sensibilidad: 99.79%

  F1-score: 99.87%

- **Clase 3 Medium:**

  Precisión: 99.99%

  Sensibilidad: 99.59%

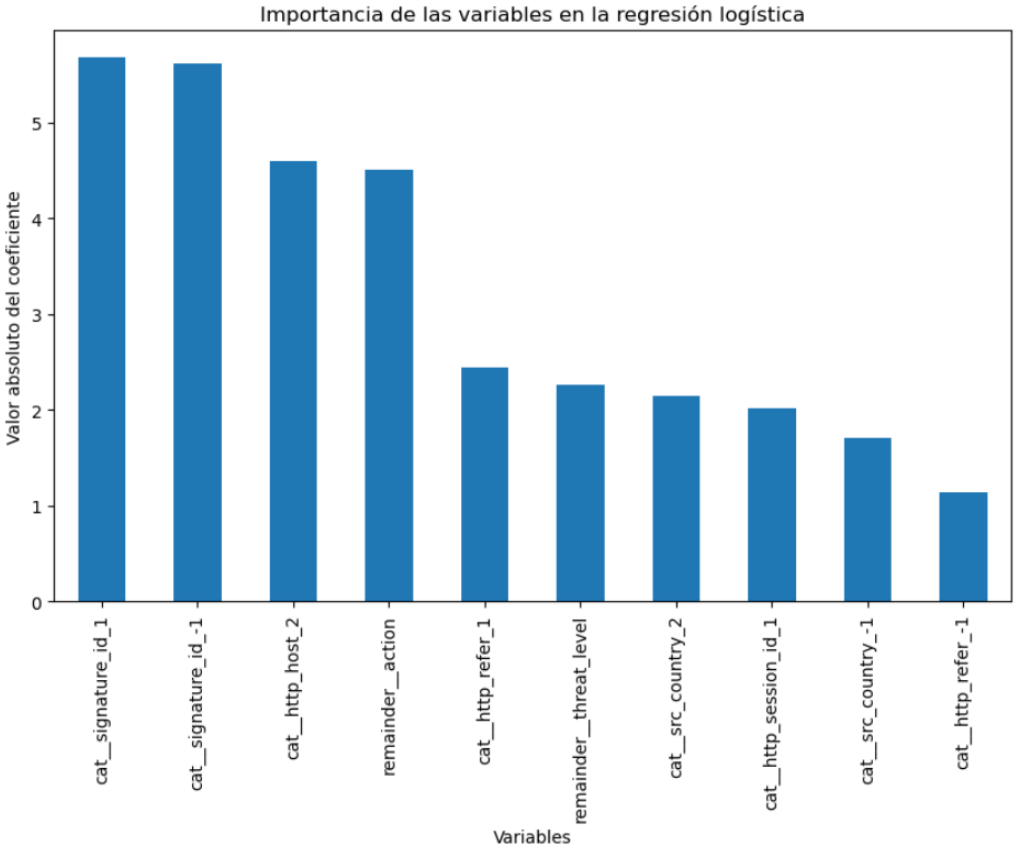
  F1-score: 99.79%

4.5.5.7 Observaciones generales

- Clase Medium (3) tiene la mayor cantidad de verdaderos positivos (214.748), lo que sugiere que la mayoría de las instancias pertenecen a esta clase.
- La precisión y sensibilidad para la clase *Low* (2) son muy altas, con valores cercanos al 100%, lo que indica un excelente rendimiento en la predicción de esta clase.
- La clase High (1) tiene un número relativamente alto de falsos positivos en comparación con los verdaderos positivos, lo que indica que el modelo tiene un poco más de dificultad para predecir correctamente esta clase.
- El modelo tiene una precisión y sensibilidad notablemente altas en general, peropude observar una mayor tasa de desaciertos en la clase *High*(1), esto puede deberse a al desbalanceo de las clases.
- Para este problema la regresión logística resulta una herramienta valiosa y efectiva para la clasificación de la severidad de ataques en sistemas de seguridad del firewall.

Ilustración 15

Importancia de las variables en el modelo de regresión logística



El gráfico muestra la importancia de las variables del modelo de regresión logística, basado en los valores absolutos de los coeficientes.

Las variables principales que tienen los valores absolutos más altos de los coeficientes son las más influyentes en la predicción del modelo, en este caso son cat_signature_id_1 y cat_signature_id_-1 como las más importantes, seguidas por cat_http_host_2 y remainder_action. Estas características son las que más contribuyen al resultado del modelo, tanto en sentido positivo como negativo.

Observé que la mayoría del peso del modelo está concentrado en un pequeño número de características, lo cual indica la distribución de la importancia, lo que es habitual en modelos donde algunas variables son mucho más informativas que otras. En este caso, después de las primeras 4 características, la importancia de las siguientes disminuye significativamente.

Las variables al final del gráfico, tienen valores absolutos de coeficiente, mucho menores. Lo cual indica que estas variables tienen un impacto mínimo en el resultado de la predicción, o en algunos casos, podrían estar casi irrelevantes en comparación con otras.

Este análisis permitió entender cuáles son las variables que el modelo considera más relevantes para predecir el resultado, además proporciona una base para tomar decisiones sobre posibles mejoras o simplificaciones en el modelo.

CAPITULO V: CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

- El modelo predictivo que desarrollé presentó un buen rendimiento, con una alta precisión. Las métricas de rendimiento (precisión, *recall y F1-Score*) fueron excelentes para todas las clases, lo que indicó que el modelo podía detectar ataques con gran exactitud. Esto me permitió cumplir satisfactoriamente con el objetivo general de detectar ataques a herramientas de seguridad perimetral.
- Realicé correctamente la carga de datos, quedando automatizada. Esto garantizó la disponibilidad de los datos para su análisis y modelado. Además, la alta precisión del modelo se debió a la calidad de los datos, que eran accesibles y completos. Esta combinación de datos bien preparados y un modelo adecuado permitió obtener resultados precisos y confiables en las predicciones.
- El preprocesamiento de los *logs* fue efectivo. Elaboré los *ETL* para la carga de los logs en la base de datos, precautelando su integridad. Los datos fueron limpiados, transformados y preparados adecuadamente para el análisis y la modelación, lo que dio buenos resultados para la predicción del modelo.

- La identificación de patrones significativos permitió al área de Seguridad Informática tomar contramedidas, realizar ajustes y comprender mejor el comportamiento de los datos, ya que el modelo pudo clasificar con alta precisión.
- El modelo clasificó correctamente los ataques en diferentes niveles de severidad, como se reflejó en las métricas de rendimiento de cada clase. Esto indicó que el agrupamiento de comportamientos según la severidad fue adecuado.

- El análisis exploratorio, combinado con la aplicación de un modelo de regresión logística en el proyecto, permitió identificar de manera efectiva patrones de comportamiento anómalos en los eventos registrados. Las características más influyentes determinadas por el modelo indicaron que ciertos tipos de firmas y host específicos tienen una correlación significativa con eventos que podrían representar amenazas de seguridad. Resaltando la importancia de estos factores en la detección temprana de incidentes de seguridad.
- El modelo de regresión logística construido no solo mostró un buen rendimiento en el conjunto de entrenamiento, sino que también mantuvo su eficacia al ser probado en datos no vistos previamente. Este desempeño consistente sugiere que el modelo tiene una alta capacidad de generalización, lo cual es crucial para su aplicación en entornos reales donde se analizan logs continuamente. La alta precisión del modelo en la detección de amenazas refuerza su utilidad como una herramienta complementaria para la toma de decisiones en la gestión de la seguridad.

5.2 RECOMENDACIONES

- Probar otros modelos, para este caso práctico, el modelo supervisado permitió generalizar bastante bien usando la regresión logística, de acuerdo con las métricas presentadas. También probé con "*Random Forest*", que generó resultados similares. Sin embargo, preferí quedarme con el modelo más simple.

- Continuar con el mantenimiento del repositorio de almacenamiento para asegurar que los datos de logs de seguridad estén siempre disponibles y actualizados.
- Implementar un procedimiento estándar de preprocesamiento que incluya limpieza, normalización y transformación de los datos, asegurando la calidad y consistencia de los datos futuros, garantizando todo el ciclo de vida de los datos.
- Continuar explorando y actualizando los patrones detectados, incluyendo nuevas variables o características que puedan mejorar aún más la detección de ataques. Implementar análisis continuos para detectar posibles cambios en los patrones de ataque.
- Realizar auditorías periódicas para verificar la eficacia de los niveles de severidad y ajustar según sea necesario. Asegurando que los criterios para determinar la severidad de los ataques se actualicen conforme a nuevas amenazas y comportamientos.
- Aunque el modelo actual ha mostrado un buen rendimiento, es fundamental considerar la actualización periódica de los datos y el modelo en sí mismo. Dado que las amenazas de seguridad evolucionan con el tiempo, es recomendable realizar un reentrenamiento del modelo con logs más recientes para mantener su eficacia.
- Es adecuado implementar un sistema de monitoreo continuo del desempeño del modelo en el ambiente productivo. Esto implica evaluar regularmente métricas clave como la precisión, la sensibilidad (recall), la especificidad y la tasa de falsos positivos. Al detectar cualquier deterioro en el rendimiento del modelo, se podrán realizar ajustes oportunos, como la incorporación de nuevas características relevantes que puedan haber surgido debido a cambios en los patrones de ataque.

CAPITULO VI: BIBLIOGRAFÍA

Curso de seguridad informática para profesionales de TI. (n.d.). Fortinet. https://www.fortinet.com/lat/training/cybersecurity-professionals

Anatomía de un ataque. (2023, 10 febrero). Cisco. https://www.cisco.com/c/es_mx/products/security/what-is-cybersecurity.html

Dagnino, J. (2014). INFERENCIA ESTADÍSTICA: PRUEBAS DE HIPÓTESIS. *Rev Chil Anest*, 43, 125–128. https://www.sachile.cl/upfiles/revistas/54e6379f16fa4_10_inferencia-2-2014_edit.pdf

Security Events log page | Administration Guide. (2024). Fortinet.com. https://docs.fortinet.com/document/fortigate/7.4.0/administration-guide/876272/security-events-log-page

De Máster en Big Data, E. C. R. P. la O. del T. (n.d.). *Análisis y desarrollo de modelos predictivos con redes neuronales para Web Application Firewall*. https://redi.anii.org.uy/jspui/bitstream/20.500.12381/461/1/20200924-MBD-234817-146966-201820.pdf

Firewall de próxima generación (NGFW).(n.d.). https://www.fortinet.com/lat/products/next-generation-firewall

Extracción, transformación y carga de datos (ETL). (s/f). Microsoft.com. https://learn.microsoft.com/es-es/azure/architecture/data-guide/relational-data/etl

raunakjhawar. (n.d.). *Extracción, transformación y carga de datos (ETL) - Azure Architecture Center*.Learn.microsoft.com. https://learn.microsoft.com/es-es/azure/architecture/data-guide/relational-data/etl

Sourtech. (2021). *Los ciberataques en Latinoamérica han aumentado un 24% este año*. Forbes Ecuador. https://www.forbes.com.ec/lifestyle/los-ciberataques-latinoamerica-han-aumentado-24-

ano-n7801

Ciberseguridad y Ciberdefensa: Perspectiva de la situación actual en el Ecuador. (2022). *Revista Tecnológica Ciencia Y Educación Edwards Deming.* https://doi.org/10.37957/rfd.v6i1.88

Villacís, R. P. C. (2022). Ciberseguridad y Ciberdefensa: Perspectiva de la situación actual en el Ecuador. *Revista Tecnológica Ciencia y Educación Edwards Deming*, *6*(1). https://doi.org/10.37957/rfd.v6i1.88

Network analytics for large & complex networks. (s/f). Fortinet. https://www.fortinet.com/products/management/fortianalyzer

What is a firewall? (2023, 6 octubre). Cisco. https://www.cisco.com/c/en/us/products/security/firewalls/what-is-a-firewall.html#~types-of-firewalls

¿Qué es la seguridad informática? (s/f). Ibm.com. https://www.ibm.com/mx-es/topics/it-security

Merterun. (2023, 5 junio). *Cybersecurity Threat analysis. Kaggle*. https://www.kaggle.com/code/merterun/cybersecurity-threat-analysis

What is a WAF? (s/f). Palo Alto Networks. https://www.paloaltonetworks.com/cyberpedia/what-is-a-web-application-firewall

¿Qué son los logs y para qué sirven? (2022, mayo 10). *KeepCoding Bootcamps.* https://keepcoding.io/blog/que-son-logs-y-para-que-sirven/

¿Qué es el análisis predictivo? (s/f). Ibm.com. https://www.ibm.com/es-es/topics/predictive-analytics

Qué es una dirección IP: definición y explicación. (2023, diciembre 19). latam.kaspersky.com. https://latam.kaspersky.com/resource-center/definitions/what-is-an-ip-address

Serra, J. (2022, junio 8). *Ciberdefensa: Qué es y cuál es su importancia.* Ciberseguridad. https://ciberseguridadtips.com/ciberdefensa/

FortiGuard Labs presenta reporte de ciberataques en América Latina. (s/f). Fortinet. https://www.fortinet.com/lat/corporate/about-us/newsroom/press-releases/2022/fortiguard-labs-reporte-ciberataques-america-latina-2021

Fortinet informa que América Latina fue el objetivo de más de 360 mil millones de intentos de ciberataques en 2022. (2022). Fortinet. https://www.fortinet.com/lat/corporate/about-us/newsroom/press-releases/2023/fortiguard-labs-reports-destructive-wiper-malware-increases-over-50-percent

Abad, W. A. (2020). CIBERATAQUES: DESAFÍOS EN EL CIBERESPACIO. *Revista de la Academia del Guerra del Ejército Ecuatoriano, 13*(1), 13. https://doi.org/10.24133/age.n13.2020.11