ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

Modelo predictivo para la detección y categorización de ciberataques basado en el análisis de tráfico de red

Presentado por:

- Mateo Medranda
- Moisés Benalcázar
- Bryan Quispe

PLANTEAMIENTO DEL PROBLEMA

En los últimos años el crecimiento de la digitalización en las organizaciones ha tenido un aumento significativo de los ciberataques en el sector empresarial. según reportes de datos, en el año 2023 hubo un incremento del 63% de ciberataques en comparación con el año 2022 y los ataques más utilizados según (OWAS, 2021) son: Pérdida de Control de Acceso, Fallas Criptográficas, Inyección, Diseño Inseguro, Configuración de Seguridad Incorrecta.

Sin embargo, la detección temprana de estos patrones de ataques no se ajusta a patrones comunes registrados como amenazas, creando una brecha de seguridad en las organizaciones y permitiendo estos ataques que pasen desapercibidos. Por esto, se requiere desarrollar un modelo predictivo más robusto que no solo anticipe la probabilidad de un ataque, sino que también lo categorice de manera efectiva, mejorando la capacidad de detección y respuesta ante distintos tipos de amenazas. Esto permitiría a las organizaciones fortalecer su seguridad y prevenir ataques antes de que se materialicen.

OBJETIVOS

GENERALES

Desarrollar un modelo predictivo capaz de anticipar y clasificar ciberataques, utilizando técnicas de aprendizaje automático para analizar patrones en el tráfico de red, con el fin de mejorar la detección temprana de amenazas y apoyar la toma de decisiones en ciberseguridad.

ESPECIFICO

- Identificar los patrones en el tráfico de red que permitan la detección de un ciberataque.
- Usar técnicas de machine learning para detección de ciberataques.
- Implementar un modelo de clasificación para detectar ciberataques.
- Utilizar métricas de evaluación para medir el rendimiento del modelo implementado.

PREGUNTAS DE INVESTIGACIÓN

01 ¿Cuáles son los patrones más comunes en el tráfico de red cuando sufren un ciber ataque?

02 ¿Que técnicas de machine learning se aplican a la detección de patrones en tráfico de red?

03 ¿Qué métricas se pueden utilizar para evaluar el desempeño adecuado de un modelo de detección de ciberataques?

04 ¿Que tecnicas se aplicaron para resolver el problema?

REVISIÓN DE LITERATURA

01

Se incluyeron artículos en inglés del área de Computer Science, tipo article, que abordaran la detección o clasificación de intrusiones mediante modelos, IA o machine learning

02

Se incluyeron artículos que estuvieran indexados con palabras clave específicas como Intrusion Detection, IDS, Deep Learning, etc

03

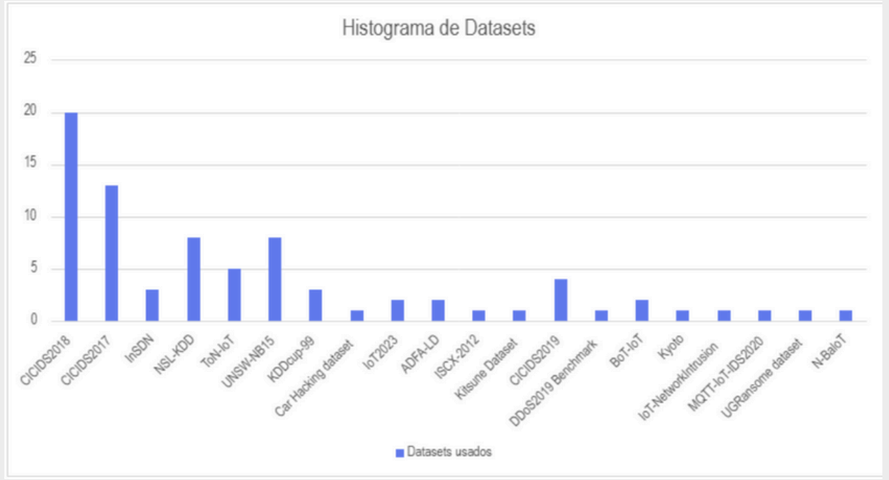
Se excluyeron documentos que no fueran artículos científicos, que no utilizaran técnicas de IA, que no trataran detección de intrusiones en ciberseguridad, que pertenecieran a otras áreas temáticas o que estuvieran escritos en un idioma distinto al inglés.

2020-2025

¿QUÉ SE OBTUVO DE LA RSL?

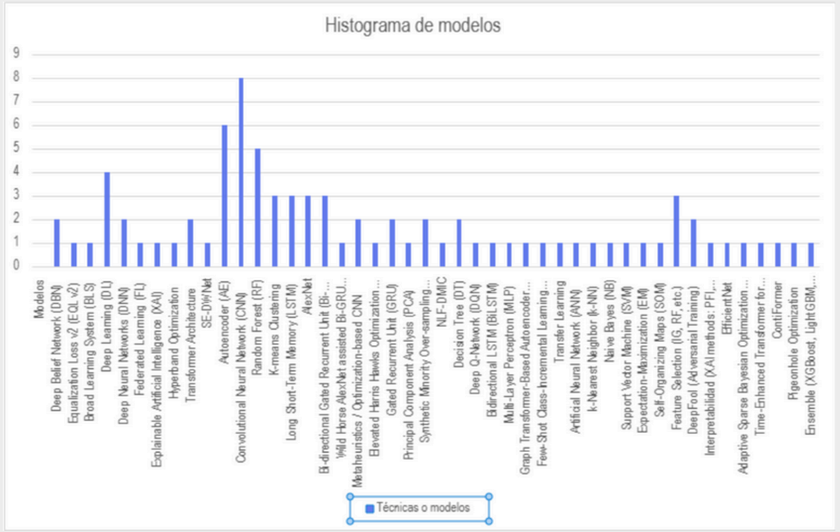
DATASETS MÁS USADOS

- CICIDS2018
- CICIDS2017
- NSL-KDD
- UNSW-NB15



MÉTRICAS

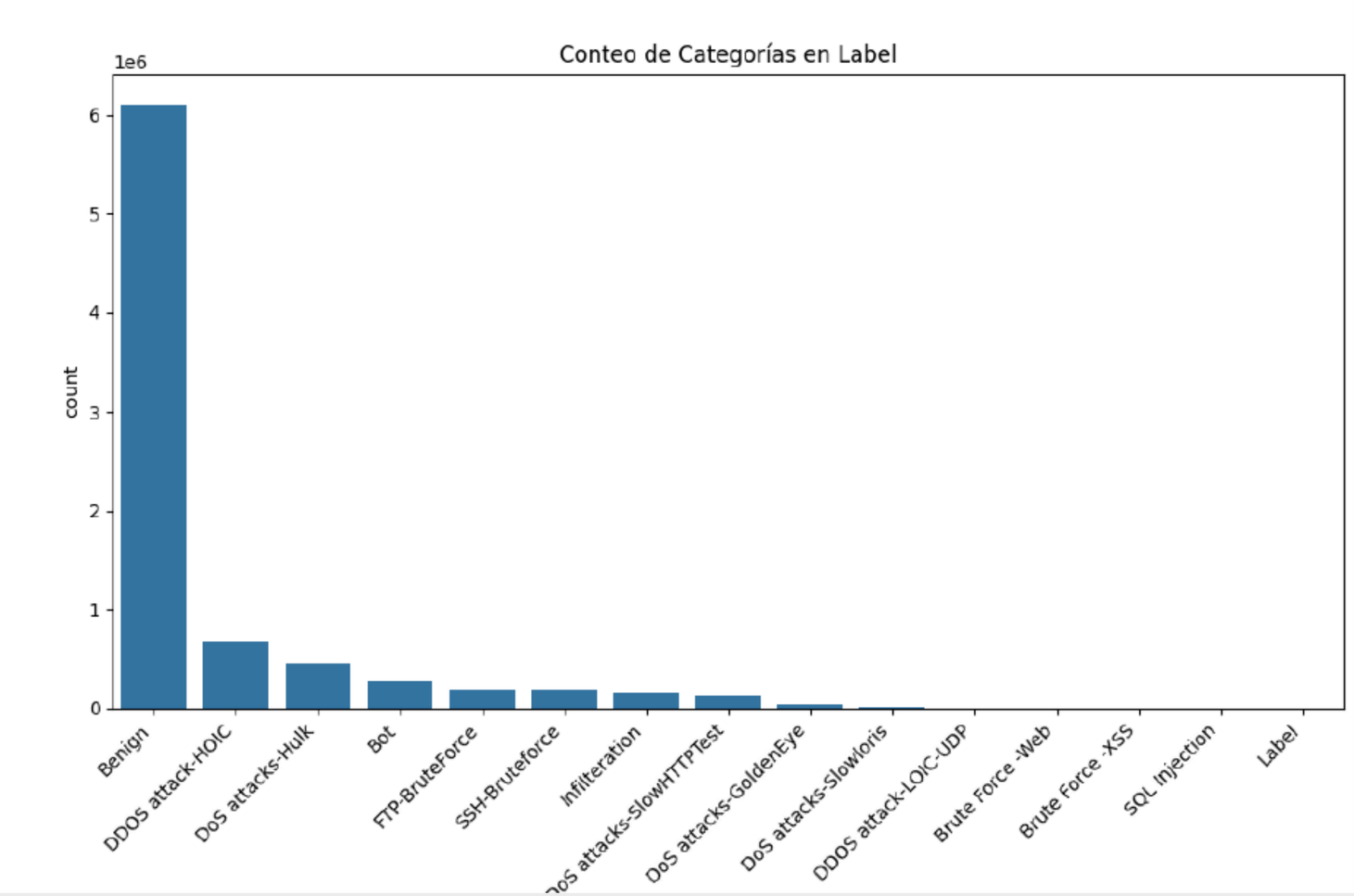
- Accuracy
- Recall
- Tasa de Falsos
- F1 Score
- Training Time



MODELOS MÁS USADOS

- Redes neuronales convolucionales (CNN)
- Autoencoder (AE)
- Random Forest (RF)
- Deep Learning (DL)
- Clustering

DATA (CICIDS2018)



ESPE

UNIVERSIDAD DE LAS FUERZAS ARMADAS

INNOVACIÓN PARA LA EXCELENCIA

REVISIÓN INTEGRATIVA DE LITERATURA

Presentado por:

- Mateo Medranda
- Moisés Benalcázar
- Bryan Quispe

¿QUÉ ES?

Una revisión integrativa de literatura es un estudio que recopila, compara y analiza investigaciones previas para ofrecer una visión completa del estado actual de un tema. Aunque surgió en las ciencias de la salud y sociales, es de carácter mixto ,combina enfoques cualitativos y cuantitativos, y puede aplicarse en computación para identificar tendencias, métodos, avances y vacíos de investigación.

-Para comenzar un proyecto de tesis.
-Para fundamentar un marco teórico.
-Para comprender tecnologías emergentes.
-Para proponer mejoras basadas en evidencia.

FASES DE UNA RIL

FASE 1: ELABORAR LA PREGUNTA ORIENTADORA

Una pregunta clara y específica permite una revisión de forma adecuada, por lo que debe estar enfocada en lo que ya se sabe y lo que se necesita

FASE 2: BÚSQUEDA O MUESTREO EN LA LITERATURA

No basta con una sola base de datos, también se debe buscar en referencias de artículos encontrados, y en revistas, buscando diferentes resultados

Muestra representativa

FASES DE UNA RIL

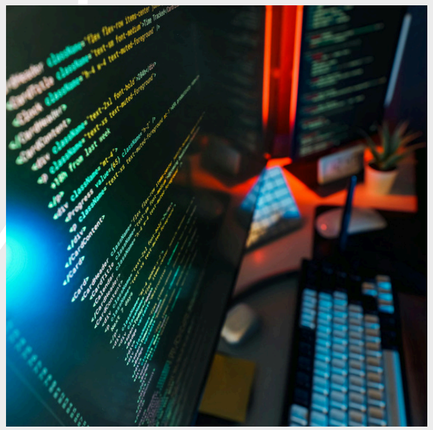
FASE 3: RECOLECCIÓN DE DATOS

Se tiene una plantilla estandarizada para obtener la información importante de cada estudio lo que garantiza consistencia y minimiza errores

FASE 4: ANÁLISIS CRÍTICO DE LOS ESTUDIOS INCLUIDOS

Se debe evaluar el rigor de cada estudio, verificando los puntos fuertes y débiles de cada uno, en ocasiones es necesaria la experiencia del investigador

FASES DE UNA RIL



FASE 5: DISCUSIÓN DE LOS RESULTADOS

Se buscan patrones y temas comunes, comparando también con inconsistencias y lagunas que puede haber, contrastando a su vez con lo que se definió al principio.

FASE 6: PRESENTACIÓN DE LA REVISIÓN INTEGRATIVA

Mediante un informe se deben presentar los resultados, lo que permite a quien lo lea seguir los pasos y definir que nuestro proceso de búsqueda fue adecuado y completo.

Suele ser más usado en el área social y de salud

RECOLECCIÓN

Campo	Dato
Título del Artículo	
Autores	
Fuente	( ) Journal ( ) Conferencia ( ) Workshop ( ) Literatura Gris (Tesis/Reporte)
Año de Publicación	
Idioma	( ) Inglés ( ) Español ( ) Portugués ( ) Otro
DOI / URL	

Entorno	Selección
Ámbito	( ) Académico (Laboratorio/Univ) ( ) Industrial (Empresa Real) ( ) Open Source (Comunidad) ( ) Mixto (Consorcio Industria-Academia)
Dominio de Aplicación	(Ej. Finanzas, Salud, E-commerce, Educación, IoT...)
Ubicación Geográfica	(País donde se realizó el estudio)

Ítem	Opciones / Datos a Extraer
1. Enfoque de Investigación	**( ) Cuantitativo** *(Métricas, Análisis Estadístico, Experimentos)* **( ) Cualitativo** *(Entrevistas, Observación, Teoría Fundamentada)* **( ) Mixto**
2. Diseño del Estudio	**Investigación Empírica:** ( ) Experimento Controlado ( ) Estudio de Caso (Case Study) ( ) Encuesta (Survey) ( ) Action Research **Investigación No Empírica:** ( ) Propuesta de Solución/Herramienta (sin validación robusta) ( ) Revisión de Literatura ( ) Reporte de Experiencia
3. Muestra / Unidad de Análisis	
4. Caracterización de la Muestra	
5. Intervención / Tecnología	**Objeto de Estudio:** (Ej. Nueva metodología Ágil, Herramienta de Testing, Algoritmo de IA)
6. Análisis de Datos	( ) Estadística Descriptiva (Promedios, Porcentajes) ( ) Estadística Inferencial (Tests de hipótesis, p-value) ( ) Codificación Temática (Para datos cualitativos)
7. Rigor Metodológico	( ) ¿Se describen claramente los criterios de inclusión/exclusión? ( ) ¿Se menciona explícitamente el control de sesgos (bias)? ( ) ¿Están justificadas las conclusiones con los datos presentados?

Sección	Descripción
Principales Hallazgos	(Resumen de los datos numéricos o cualitativos. Ej: "La herramienta redujo los errores en un 15%").
Análisis Estadístico	¿Hubo pruebas de significancia (p-value, t-test)? ( ) Sí ( ) No
Limitaciones Declaradas	(Ej. Muestra pequeña, solo probado en Java, entorno simulado).

APLICACIONDE UNA RIL

FASE 1: ELABORAR LA PREGUNTA ORIENTADORA

"¿Cómo detectar y categorizar ciberataques a partir del análisis de patrones en tráfico de red usando machine learning?" + las 4 RQ.

Preguntas de investigación

FASE 2: BÚSQUEDA O MUESTREO EN LA LITERATURA

- Período: 2020-2025
- Bases consultadas: IEEE Xplore, ScienceDirect, SpringerLink, ACM Digital Library
- Palabras clave: "Intrusion Detection", "Network Traffic Analysis", "Cyberattack Detection", "Machine Learning", "Deep Learning", "CICIDS2017", "CICIDS2018", "UNSW-NB15"
- Criterios de inclusión/exclusión
- Muestra representativa obtenida: 30 artículos

APLICACIONDE UNA RIL

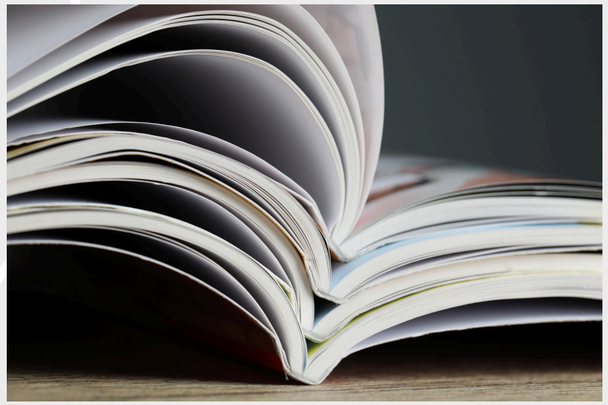
FASE 3: RECOLECCIÓN DE DATOS

Usamos la plantilla estándar + Zotero, Excel y Google Forms para extraer de forma ordenada: título, dataset, modelo, métricas, hallazgos y limitaciones de cada artículo.

FASE 4: ANÁLISIS CRÍTICO DE LOS ESTUDIOS INCLUIDOS

- Fortalezas: Datasets realistas (CICIDS) y modelos Deep Learning con >95% de precisión.
- Debilidades: Mucho desbalance de clases, pruebas solo en laboratorio, poca atención a detección temprana (escaneo/reconocimiento) y alto costo computacional.

APLICACIONDE UNA RIL



FASE 5: DISCUSIÓN DE LOS RESULTADOS

- Datasets populares: CICIDS2017/2018, NSL-KDD, UNSW-NB15
- Modelos dominantes: Deep Learning (CNN, LSTM, Autoencoders) y Random Forest
- Métricas más usadas: Accuracy, Precision, Recall, F1, Detection Rate
- Hay brecha clara o no, lo que justifica nuestro modelo predictivo.

FASE 6: PRESENTACIÓN DE LA REVISIÓN INTEGRATIVA

Publicar los resultados de nuestra RIL en una revista científica indexada (Scopus / WoS / Latindex) para contribuir al conocimiento en ciberseguridad.

BIBLIOGRAFÍA

- Garrido, V., y López-Gil, M. (2021). Cómo hacer una revisión integrativa de la literatura científica. Pirámide.
- Whittemore, R., y Knafl, K. (2005). The integrative review: Updated methodology. Journal of Advanced Nursing, 52(5), 546–553.
- Whittemore, R., y Knafl, K. (2017). La revisión integradora: Metodología actualizada (L. Torres, Trad.). Enfermería Intensiva, 28(4), 184–191.
- Torraco, R. J. (2016). Writing integrative literature reviews: Using the past and present to explore the future. Human Resource Development Review, 15(4), 404–428.
- Botella, J., y Gambara, H. (2020). Hacer investigación en psicología y ciencias afines: De la idea al informe. Pirámide. (Capítulo 10 dedicado a revisiones sistemáticas e integrativas).
- Manterola, C., y Zavando, D. (Eds.). (2022).

GRACIAS