Vol.8 No.3 (2024): Journal Scientific MQRInvestigar ISSN: 2588–0659
https://doi.org/10.56048/MQR20225.8.3.2024.275-296

Evaluation of machine learning, random forest and xgboost techniques for trojan detection

Evaluación de técnicas de machine learning, random forest y xgboost para la detección de troyanos

Autores:

Luna-Haro, César
PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR
Ingeniero en Tecnologías de la Información
Estudiante de la Maestría de Ciberseguridad
Ambato – Ecuador
cdluna@pucesa.edu.ec
https://orcid.org/0009-0008-0745-2882

Arellano-Aucancela, Alberto
PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR
Ing. Electrónica y Computación , Msc. en Informática Aplicada
Docente Tutor de Seguridad en Infraestructura de Redes
Ambato – Ecuador
aarellano@pucesa.edu.ec
https://orcid.org/0000-0002-4392-5148

Fechas de recepción: 20-JUN-2024 aceptación: 27-JUN-2024 publicación:15-SEP-2024
https://orcid.org/0000-0002-8695-5005
http://mqrinvestigar.com/

Resumen

El objetivo de este estudio era evaluar técnicas de aprendizaje automático, en concreto Random Forest y XGBoost, para la detección de troyanos. La idea central fue que ambos métodos resultarían eficaces en esta tarea. La metodología consistió en seleccionar características relevantes y utilizar un conjunto de datos representativo. Los resultados demostraron que tanto Random Forest como XGBoost alcanzaron una efectividad del 99% en la detección de troyanos, con una ligera diferencia de XGBoost. Esta diferencia se atribuye a la gestión eficaz y adaptable de la complejidad del conjunto de datos por parte de XGBoost, que optimiza la precisión del modelo, al no registrar falsos positivos. La solidez de estos resultados se ve reforzada por los datos de evaluación recopilados. La importancia de este hallazgo radica en la aplicación con éxito del aprendizaje automático para la detección de amenazas de ciberseguridad, con implicaciones críticas para la seguridad de la información en entornos corporativos. El estudio destaca la eficacia de Random Forest y, en particular, de XGBoost en la detección de troyanos, lo que ofrece información valiosa para la supervisión de amenazas en tiempo real. Los autores sugieren considerar XGBoost como la mejor alternativa en este contexto, subrayando la importancia permanente de explorar y perfeccionar las técnicas de aprendizaje automático para mejorar la seguridad informática

Palabras clave: Detección de troyanos; Machine learning; Random Forest; XGBoost

Abstract

The aim of this study was to evaluate machine learning techniques, namely Random Forest and XGBoost, for Trojan detection. The central idea was that both methods would be effective in this task. The methodology consisted of selecting relevant features and using a representative dataset. The results showed that both Random Forest and XGBoost achieved an 99% effectiveness in detecting Trojans, with a slight difference for XGBoost. This difference is attributed to XGBoost's efficient and adaptive management of the complexity of the dataset, which optimises the accuracy of the model by not registering false positives. The robustness of these results is reinforced by the evaluation data collected. The significance of this finding lies in the successful application of machine learning for cybersecurity threat detection, with critical implications for information security in corporate environments. The study highlights the effectiveness of Random Forest and, in particular, XGBoost in detecting Trojans, which provides valuable information for real-time threat monitoring. The authors suggest considering XGBoost as the best alternative in this context, underlining the continued importance of exploring and refining machine learning techniques to improve IT security.

Keywords: Machine learning; Random Forest; Trojan Detection; XGBoost

Introducción

La era digital en la actualidad ha convertido a la ciberseguridad en un desafío crucial, ya que la tecnología se encuentra presente en las rutinas diarias de empresas y personas. La integridad, confidencialidad y disponibilidad de los sistemas informáticos, han sido amenazados por el aumento de los peligros y amenazas que se encuentran en internet, entre ellos tenemos el malware, que es una forma de troyano de tipo malicioso muy perjudicial y sigiloso, diseñado para operar de forma encubierta. Con el objetivo de robar información confidencial o facilitar otros ataques, los troyanos se infiltran en sistemas y redes. Esta situación resalta la importancia de mantener medidas de seguridad sólidas en este mundo digital que se encuentra en una evolución continua. (Kaspersky, 2021)

Con el paso del tiempo los ciberdelincuentes han encontrado formas para ocultar a los troyanos y de esta manera evitar las defensas convencionales. Con esto, es necesaria la búsqueda de métodos más proactivos y eficaces que ayuden a rastrear y mitigar amenazas, que constantemente van creciendo. (Kamboj et al, 2022)

> Expertos en ciberseguridad han obtenido una ayuda significativa para la detección de troyanos, gracias a la evolución en el campo del aprendizaje automático y la inteligencia artificial; las mismas que han ayudado identificando de una manera más rápida y precisa los patrones maliciosos y actividades sospechosas, con el análisis del comportamiento de programas en tiempo real.

Como solución para combatir a los troyanos y otros riesgos que se presentan en línea ha surgido el Machine Learning, que no es más que el aprendizaje automático de las computadoras mediante información y experiencias previas. Esta evolución ha ayudado mediante la implementación de algoritmos de Machine Learning en ciberseguridad, a mejorar la habilidad de detectar y prevenir los ataques de troyanos en tiempo real. Este progreso es gracias al análisis de grandes volúmenes de datos llamados datasets y al reconocimiento de patrones sutiles, comportamientos anómalos y firmas de malware.

Un dataset, o conjunto de datos, es un depósito rico y organizado de información, meticulosamente recolectado para abordar interrogantes de investigación específicas. La confiabilidad y validez de los resultados que se obtienen durante el análisis, están relacionados con la calidad y la composición del conjunto de datos. La elección cuidadosa de variables, la representatividad de las muestras y la gestión considerada de posibles sesgos son aspectos cruciales en la construcción y utilización de datasets. (Louk y Tama, 2022).

La combinación del aprendizaje automático en la detección de troyanos es ideal para proteger la infraestructura digital de las organizaciones y personas en una época de innovación tecnológica constante y amenazas en línea crecientes.

La seguridad informática ha encontrado un aliado poderoso en la forma de algoritmos de aprendizaje automático ahora llamados ALAU, particularmente XGBoost y Random Forest, los cuales llamaremos XGB y RF, respectivamente. Estos ALAU se destacan en la identificación de comportamientos anómalos en los datos, una habilidad útil para identificar troyanos, que suelen manifestar acciones inusuales en comparación con el software legítimo. A través del análisis de patrones de actividad y el entrenamiento con datos históricos, XGB y RF pueden identificar acciones sospechosas, como accesos no autorizados, transferencias de datos inusuales o cambios en el comportamiento del sistema. Los ALAU pueden tener la capacidad de generar alertas y respuestas en tiempo real, lo que es fundamental para mitigar los efectos dañinos de los troyanos antes de que causen un daño significativo. (Roy et al, 2023)

La importancia de las técnicas XGB y RF radica en la capacidad de hacer predicciones en tiempo real. Los mismos que pueden emitir alertas y actuar al instante con el fin de reducir daños ocasionados por las amenazas detectadas; además pueden ir evolucionando constantemente para detectar nuevas amenazas que con el tiempo van surgiendo. Esto hace que los algoritmos sean flexibles y efectivos a largo plazo. Todo esto ha permitido a los profesionales e investigadores comprender el origen de las alertas, para aportar en las investigaciones de las amenazas y en la toma de medidas adecuadas, gracias a su grado de comprensión o interpretabilidad.
Además, en la detección de amenazas, estos algoritmos tienen la capacidad de reducir la cantidad de falsos positivos. Al modelar el comportamiento normal de un sistema o red, es posible que las alertas sean recibidas en caso de detectar un cambio considerable en su comportamiento. (Sanz, 2022)

El RF, tiene una capacidad predictiva sólida y es versátil. Su función consiste en la creación de múltiples árboles de decisión, los cuales, en conjunto contribuyen de manera única para realizar una predicción final. Cada árbol se construye con un conjunto de datos aleatorio, seleccionados de manera independiente y con reemplazo durante el proceso inicial. Este método, también conocido como "muestra de arranque", aumenta la variabilidad y la diversidad del modelo. (Hu y Szymczak, 2023)

Durante la construcción de cada árbol, se evalúa la mejor característica en cada nodo mediante técnicas como la ganancia de información o la impureza de Gini. La iteración del proceso en diferentes subconjuntos de datos garantiza la independencia entre los árboles, evitando la susceptibilidad al sobreajuste. En la fase de predicción, la salida final se determina mediante una votación ponderada de todos los árboles, proporcionando una predicción robusta y equilibrada. Esta metodología, al mitigar la sobreajustación y mejorar la estabilidad del modelo, subraya la utilidad y aplicabilidad del RF en diversas situaciones predictivas. (Dasari et al., 2022)

La técnica XGB destacada en el aprendizaje automático, se centra en potenciar árboles de decisión secuencialmente para corregir errores previos. A diferencia de enfoques aleatorios, emplea ponderaciones para enfocarse en correcciones más significativas. Durante la construcción de árboles, utiliza funciones de pérdida específicas y regularización para controlar la complejidad del modelo y evitar sobreajuste. Finalmente, combina las predicciones ponderadamente para obtener una salida precisa. Su implementación eficaz lo posiciona como una herramienta poderosa para tareas predictivas complejas. (Doghramachi y Ameen, 2023)

La implementación de los ALAU RF y XGB representan una estrategia para proteger sistemas y redes contra las amenazas cibernéticas. Estos algoritmos pueden detectar comportamientos anómalos, modelar características relevantes y adaptarse a la evolución constante de los troyanos, lo que contribuye a fortalecer las defensas cibernéticas y garantizar la seguridad en un entorno en constante cambio. Realizando una convergencia poderosa en la lucha contra las amenazas cibernéticas.

A pesar del progreso en la aplicación de ALAU a la ciberseguridad, la literatura científica todavía presenta lagunas en la evaluación comparativa de algoritmos para la detección de troyanos. Los estudios previos se han centrado en enfoques individuales o en contextos específicos, lo que ha dejado sin explorar una revisión detallada que compare y contraste la efectividad de algoritmos como RF y XGB en la detección de troyanos. (Sidhu et al., 2019)
El F1-score es una métrica de evaluación que combina precisión y recall en un solo número. Es especialmente útil cuando hay un desequilibrio entre las clases en el conjunto de datos. La fórmula del F1-score se define como la media armónica de precisión y recall y se expresa de la siguiente manera, en la ecuación (1): (Aldhyani y Alkahtani 2023).

$$F1 = \frac{2\ x\ Presicion\ x\ Recall}{Presicion + Recall} \qquad (1)$$

La curva ROC (Receiver Operating Characteristic) es una representación gráfica del rendimiento de un ALAU de clasificación binaria en diferentes umbrales de decisión. Se utiliza comúnmente para evaluar y comparar la capacidad de discriminación de un modelo, cuyas fórmulas se encuentran expresadas en las ecuaciones (2) y (3). (Prakash et al., 2023)

$$FPR = \frac{Falsos\ Positivos}{Falsos\ Positivos + Verdaderos\ Negativos} \quad (2)$$

$$TPR = \frac{Verdaderos\ Positivos}{Verdaderos\ Positivos + Falsos\ Negativos} \quad (3)$$

En contextos relacionados, existen investigaciones realizadas con la aplicación del aprendizaje automático; entre ellos, se puede nombrar al estudio enfocado al pronóstico del consumo de energía eléctrica mediante la comparación entre las técnicas Random Forest y XGBoost, obteniendo un resultado que revela que RF muestra un ajuste superior en términos del error de la raíz cuadrada de la media en las predicciones y las tendencias en comparación con el consumo real (Carrillo et al, 2023). También la comparación de algoritmos de Random Forest y XGBoost en una base de solicitudes de tarjetas de crédito, obteniendo a XGB como el modelo más preciso (Espinosa, 2020)

Lo que aún permanece como una pregunta abierta es la necesidad de una evaluación y comparación de los algoritmos RF y XGB para la detección de troyanos.
Por lo que, el presente estudio tiene como objetivo abordar esta temática en la literatura proporcionando una evaluación de RF y XGB en la detección de troyanos, por medio de la consideración un rango diverso de conjuntos de datos y escenarios. Considerando una gran variedad de datasets, la presente propuesta pretende respaldar esta afirmación proporcionando evidencia de la efectividad de los algoritmos para la identificación de amenazas cibernéticas. Para lo cual, es indispensable el uso de herramientas y entornos de simulación cibernética para recrear escenarios realistas, con el fin de implementar los ALAU seleccionados y realizar evaluaciones comparativas para mejorar la detección de troyanos.

Material y métodos

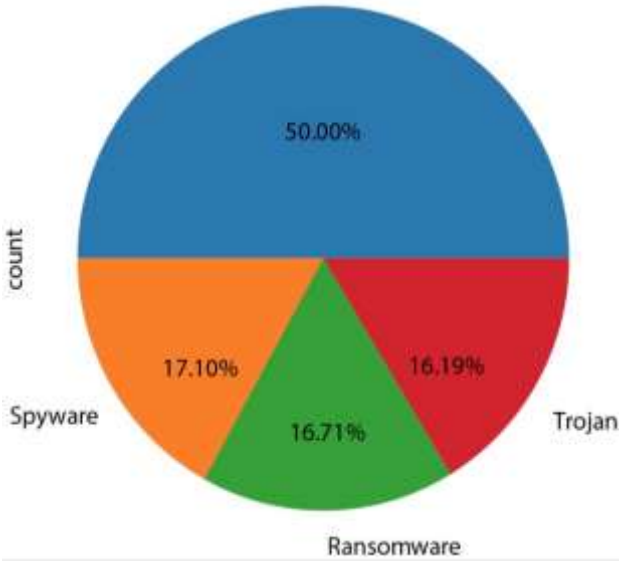
En esta sección, se detalla el proceso metodológico utilizado para evaluar las técnicas de machine learning RF y XGB para detectar troyano. El proceso inicia con la recopilación de datos, seguida del desarrollo de algoritmos con los hiperparámetros necesarios, posteriormente se selecciona el conjunto de datos que tienen

características relevantes para el entrenamiento y evaluación del modelo, y finalmente se ajustan las técnicas RF y XGB, con el objetivo de tener el menor porcentaje de falsos negativos, utilizando el conjunto de herramientas disponibles en el entorno de programación Python.

Recopilación de Datos

La recopilación de datos fue crucial para obtener conjuntos de datos representativos de troyanos y software legítimo para el entrenamiento, validación y pruebas de los modelos de Machine Learning, la precisión del modelo dependió de la calidad de los datos. Fue importante proporcionar información adicional sobre las características específicas de los conjuntos de datos como la información, comportamiento y peso de los archivos, así como justificar la elección del conjunto de datos MalMem2022, ya que de un total de 10 datasets estudiados, este fue el que estuvo mejor estructurado por ello es necesario incluir su relevancia y representatividad en el contexto de la investigación. (Shafin et al., 2023)
En el MalMem2022, en la Figura 1, se muestra una primera evaluación, la cual indica los porcentajes de los tipos de troyanos que existen, siendo los archivos benignos los de mayor afluencia con un porcentaje del 50%.

Figura 1
Primera evaluación del dataset.



Fuente: Luna César, 2024

Posteriormente se realizó un análisis más detallado del dataset en donde se muestra un equilibrio entre datos maliciosos y benignos, en total, comprende 58.596 registros en donde el 50% de los datos malignos se muestran detalladamente en la Tabla 1.

Tabla 1

Análisis del Malware del dataset

Categoría	Familias	Count
Trojan	Zeus	195
	Emotet	196
	Refroso	200
	Scar	200
	Reconyc	157
Spyware	180Solutions	200
	Coolwebsearch	200
	Gator	200
	Transponder	241
	TIBS	141
Ransomware	Conti	200
	MAZE	195
	Pysa	171
	Ako	200
	Shade	220

Note. Carrier et al, 2022

Para validar los datos presentados en la Tabla 1, se siguió un enfoque estándar en el aprendizaje automático, dividiendo el conjunto de datos en dos subconjuntos: uno para entrenamiento y otro para validación. Este proceso implica asignar el 70% de los datos al conjunto de entrenamiento, que se utiliza para ajustar los modelos de manera óptima. Los datos restantes, que constituyen el 30%, se reservan para la validación del rendimiento de los modelos entrenados. Este enfoque se conoce como validación cruzada y garantiza que los modelos sean evaluados de manera imparcial y objetiva, ya que se prueban con datos independientes de aquellos utilizados para el entrenamiento. Esta separación de datos permite verificar la capacidad de generalización de los modelos, es decir, su habilidad para realizar predicciones precisas en datos nuevos y no vistos durante el entrenamiento.

Desarrollo de los Algoritmos

El desarrollo de los algoritmos de Machine Learning fue una fase decisiva, donde se involucra la configuración precisa de hiperparámetros, adaptándolos específicamente

a la detección de troyanos. Por esta razón, se desarrolló un modelo de aprendizaje automático basado en RF y otro en XGB para la detección de troyanos, adaptado al dataset utilizado.

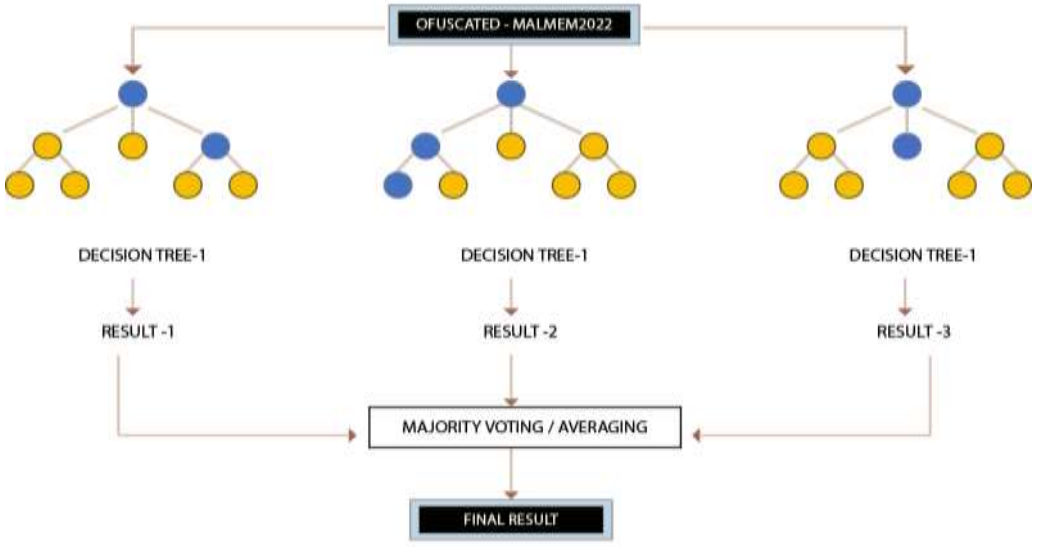
En el desarrollo de los algoritmos propuestos, se aplicó un proceso estructurado de aprendizaje automático. Se realizaron tareas de preprocesamiento de datos, como la codificación de características categóricas, como "malware_category" y "malware_family", que representan diferentes categorías o familias de malware. Estas características categóricas fueron transformadas a un formato numérico para permitir su utilización en los modelos de aprendizaje automático. El proceso de codificación fue esencial para que los algoritmos de aprendizaje automático puedan trabajar con estas características categóricas de manera efectiva. Este paso se llevó a cabo para asegurar que los datos estuvieran en un formato adecuado para el análisis y modelado posterior.

Entrenamiento del Modelo

La etapa de entrenamiento del modelo implicó la aplicación de los algoritmos RF y XGB al conjunto de datos preseleccionado (MalMem2022) como se muestran en las Figura 2 y 3.

Figura 2

Modelo de aprendizaje automático basado en RF

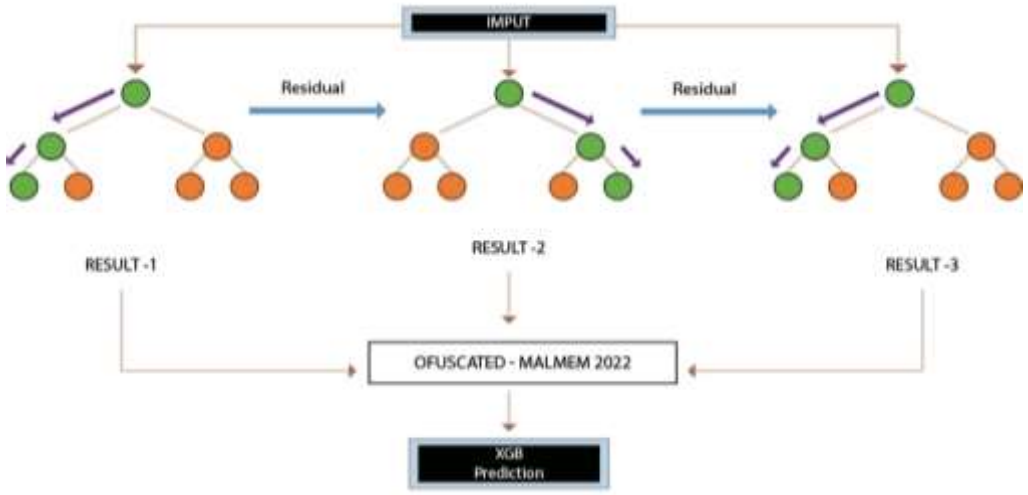


Fuente: Luna César, 2024

La votación por mayoría es una de las técnicas de aprendizaje en conjunto. El concepto detrás de este enfoque implica múltiples árboles de decisión para predecir resultados a partir de nuevos datos; cada árbol presenta su predicción y la decisión final se basa en la predicción más común entre todos los árboles. Esta estrategia hace que cada árbol de decisión aporte una predicción sobre el nuevo dato, y luego el que aparece con más frecuencia se convierte en la elección final. Aunque la votación por mayoría es simple y puede mejorar la precisión al no sobreajustar demasiados datos (ya que combina diferentes fortalezas de los árboles de decisión), generalmente es más vulnerable a valores atípicos (es decir, menos precisos) en comparación con otros métodos conjuntos como el impulso o el embolsado. No obstante, este enfoque tiene diversas aplicaciones prácticas, incluyendo la clasificación, la regresión y la detección de anomalías.

Figura 3

Modelo de aprendizaje automático basado en XGB



Fuente: Luna César, 2024

La Figura 3 representa el esquema de un algoritmo de aprendizaje automático conocido como red neuronal residual (ResNet). Las ResNets son una arquitectura de red neuronal única diseñada para combatir lo que llamamos el problema de desaparición del gradiente que afecta a las redes profundas. Este problema inhibe el flujo fluido de gradientes a través de una red, lo que dificulta una formación eficaz. Sin embargo, las ResNets abordan esto de frente estableciendo lo que llaman conexiones directas (o conexiones de acceso directo) que actúan como un bypass para que estos gradientes atraviesen libremente la red. En este proceso de entrenamiento, los datos de entrada pasan por capas convolucionales y de agrupación, mientras que la última salida convolucional se suma a estos datos iniciales a lo largo de la ruta de

conexión directa. Las ResNets son conocidas por su capacidad para alcanzar una profundidad significativa durante el entrenamiento sin sobreajuste. Esto ha llevado a ResNets a demostrar un gran rendimiento en diferentes tareas de visión por computadora, como clasificación de imágenes o detección de objetos, así como segmentación semántica, a pesar de que dichos modelos son más costosos desde el punto de vista computacional y necesitan un ajuste cuidadoso de los hiperparámetros. Pero en realidad encuentran amplias áreas de aplicación como la detección de troyanos.

Evaluación del Rendimiento

La evaluación del rendimiento se llevó a cabo mediante dos métricas fundamentales: el F1-score y el área bajo la curva ROC (AUC). El F1-score proporcionó una medida equilibrada de la capacidad del modelo para detectar troyanos, considerando tanto falsos positivos como falsos negativos. Por otro lado, el AUC evaluó la capacidad del modelo para distinguir entre troyanos y software legítimo, ofreciendo una perspectiva integral de la efectividad del algoritmo en diferentes umbrales de clasificación. La justificación para la elección de estas métricas estuvo en su capacidad para medir tanto la sensibilidad como la especificidad, aspectos importantes para la evaluación del rendimiento de los algoritmos en la detección de amenazas cibernéticas.

Análisis de Resultados

En la fase de análisis de resultados, se examinó cómo los algoritmos responden a diversas amenazas y se evaluó su desempeño general. Enfocándose en los hallazgos, identificando patrones, debilidades y fortalezas de los algoritmos en relación con los tipos de amenazas específicas enfrentadas en el entorno cibernético.

Resultados

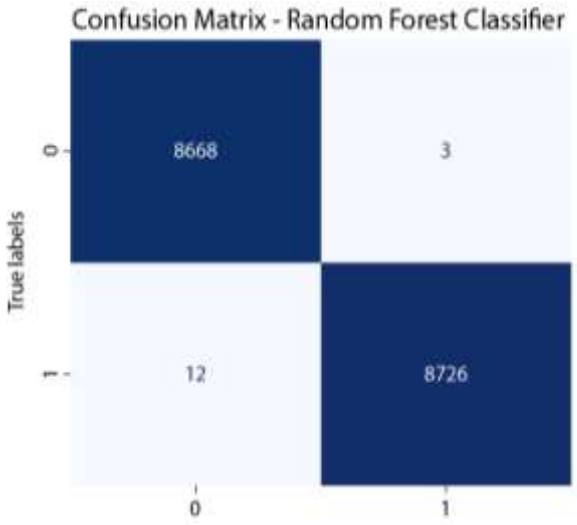
En esta investigación, una vez definida la metodología, primero se recopilaron los datos de los conjuntos de datos mejor estructurados, dando como resultado el dataset llamado MalMem2022, que tiene 58598 datos de los cuales la mitad son benignos y el resto de data maliciosa, posteriormente, fueron desarrollados y configurados los algoritmos de Machine Learning, adaptándolos específicamente a la detección de troyanos. Utilizando este conjunto de datos diverso y representativo de amenazas cibernéticas reales específicamente de troyanos, este proceso permitió evaluar el desempeño de RF y XGB en una variedad de contextos, incluyendo la selección de

hiperparámetros, el preprocesamiento de datos e implementación de validación cruzada.

Los resultados iniciales F1-Score mostraron una alta precisión en la identificación de troyanos, marcando un avance en la seguridad digital y fortaleciendo las defensas contra amenazas cibernéticas. Adicionalmente se analizó la matriz de confusión, como se muestra en la Figura 4, la matriz de confusión refleja un desempeño general positivo del modelo de clasificación, con 8668 verdaderos positivos y 8726 verdaderos negativos teniendo un 99% de efectividad. Esto demuestra una capacidad constante para clasificar instancias tanto positivas como negativas. Sin embargo, se observan tres falsos negativos, sugiriendo una oportunidad de mejora en la identificación de casos positivos. Por otro lado, se registraron doce falsos positivos, lo que indica una disminución leve en la precisión del modelo porque algunas instancias negativas se clasificaron incorrectamente como positivas.

Figura 4

Matriz de confusión con RF



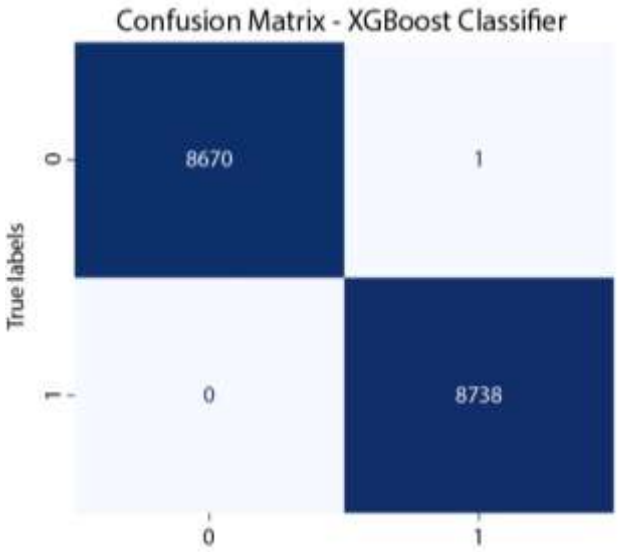
Fuente: Luna César, 2024

Este modelo fue entrenado con un conjunto de datos diversos donde se abordó características y comportamientos que se asocian a actividades maliciosas. La elección de XGB se basó en su capacidad para reconocer patrones. Los resultados del F1 SCORE mostraron una alta precisión en la detección de troyanos, contribuyendo a mejorar las líneas de defensa en contra de las amenazas cibernéticas. La matriz de confusión resultante se representa en la Figura 5, revelando un rendimiento sólido del modelo de clasificación. Con ocho mil seiscientos setenta (8670) verdaderos positivos y ocho mil setecientos treinta y ocho (8738) verdaderos negativos, XGB demostró una alta capacidad para identificar correctamente tanto instancias positivas como

negativas. El único punto de mejora se encuentra en un solo falso negativo, indicando una mínima falla en la clasificación de casos positivos dándonos como resultado un 99% de efectividad. No se registraron falsos positivos, destacando la precisión del modelo al evitar la clasificación incorrecta de instancias negativas. En conjunto, estos resultados respaldaron la eficacia de XGB con un rendimiento positivo y mínimos errores de clasificación.

Figura 5

Matriz de confusión con XGB

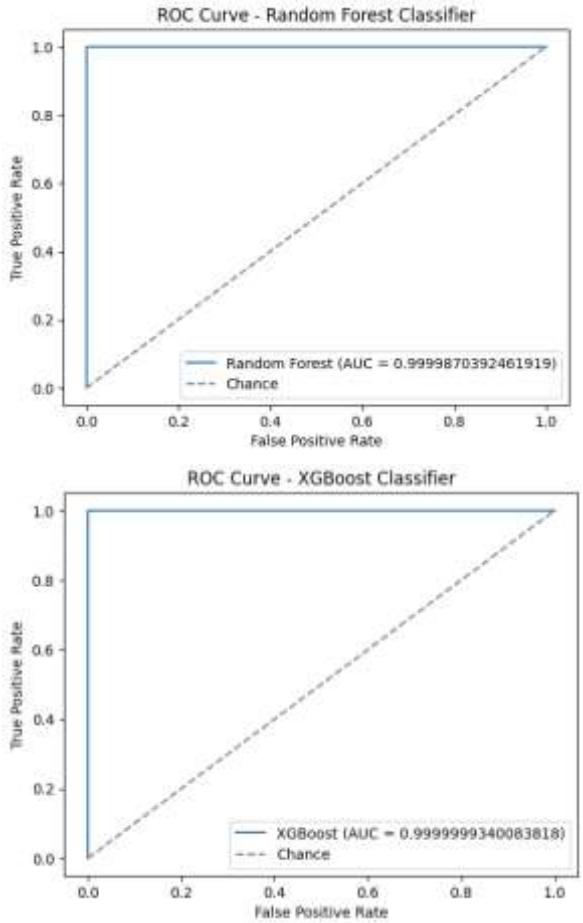


Fuente: Luna César, 2024

Luego, se realizó la evaluación del rendimiento de los algoritmos, como se aprecia en la Figura 6. Utilizando la métrica clave, el área bajo la curva ROC, para evaluar cómo se desempeñaron RF y XGB en la detección de troyanos. Esta evaluación utilizó conjuntos de datos específicos para cada algoritmo, lo que nos permitió abordar la variabilidad en las amenazas cibernéticas.

Figura 6

Curva ROC de RF (izquierda) y Curva ROC de XGB (derecha)



Fuente: Luna César, 2024

Los resultados de la evaluación demostraron la efectividad del 99% tanto de RF como XGB en la detección de troyanos. La precisión de los algoritmos fue positiva. Las curvas ROC también revelaron un rendimiento sólido, y los valores del área bajo la curva indicaron una capacidad para distinguir entre amenazas y software legítimo.
Para elegir el mejor modelo es necesaria comparación entre los dos algoritmos, podemos observar en la Tabla 2, que los resultados sugieren que, aunque no hay una diferencia sustancial entre la efectividad de estos algoritmos para detectar troyanos, se diferencian otros parámetros que están mencionados a continuación.

Tabla 2

Aspecto	Random Forest	XGBoost
Método de aprendizaje	Árboles de decisión.	Algoritmo basado en la Gradiente.

Velocidad de aprendizaje	Generalmente rápido, los árboles se pueden entrenar en paralelo.	Lento, debido a que el entrenamiento es secuencial.
Ajuste	Menos propenso al sobreajuste debido al promedio.	Tiene incorporado parámetros para sobreajuste.
Casos de Uso	Puede usarse en diversos datasets.	Necesita datasets estructurados y organizados.
Hiperparámetros	Necesita menos hiperparámetros.	Necesita más Hiperparámetros cuidadosamente ajustados.
Solidez	Sólido frente a datasets mal estructurados.	Es afectado a los datos ruidosos o mal estructurados.

Note. Roy et al, 2023

Ambos mostraron un rendimiento competitivo en el conjunto de datos y métricas evaluadas. pero basándonos en los criterios de método de aprendizaje, ajuste y regularización, capacidad para manejar interacciones complejas, y el contexto específico de detección de troyanos en datos estructurados, los autores eligen a XGB como el modelo superior. Su precisión, robustez en el manejo de características complejas y capacidad de regularización detallada hacen de XGB la opción más eficaz para la detección de troyanos, a pesar de requerir un ajuste más complejo de hiperparámetros y tener una velocidad de entrenamiento más lenta.

Estos hallazgos respaldaron la idea central de que la implementación de ALAU, en particular RF y XGB, mejoran la detección de troyanos.

Discusión

El avance tecnológico indica que los troyanos, se han convertido en una amenaza constante y desafiante, es así que las amenazas cibernéticas son cada vez más peligrosas y sigilosas. Por lo tanto, la evaluación de las técnicas XGB y RF en los múltiples escenarios, para la detección de troyanos tuvieron un rendimiento sólido, relevante en la ciberseguridad; resultados que fueron importantes para señalar que la

implementación de algoritmos de aprendizaje automático potencia notablemente la detección de troyanos.

Para el entrenamiento y evaluación de los modelos, se empleó la técnica de evaluación cruzada, la cual consiste en dividir los datos en k subconjuntos y realizar varias iteraciones de entrenamiento y validación para asegurar una evaluación confiable del rendimiento del modelo y reducir el riesgo de sobreajuste. La optimización de los hiperparámetros de XGB y RF fueron mediante búsquedas aleatorias, donde la elección aleatoria de un número fijo de combinaciones, proporciono una alternativa eficiente en cuanto al tiempo de computación. (Diaz y Sanyer, 2021)

Al comparar estos modelos se puede establecer que la efectividad del 99% alcanzada tanto para RF como para XGB en la detección de troyanos, concuerda con estudios previos presentes en la literatura que reconocen la eficacia de los algoritmos de aprendizaje automático en el campo de la ciberseguridad. XGB mostró una mayor precisión y capacidad de manejo de interacciones complejas entre características, lo cual se debe a su técnica avanzada de regularización que previene el sobreajuste. Estos resultados se alinean con los hallazgos realizados en el ámbito de aplicación de algoritmos RF y XGB en una base de solicitudes de tarjetas de crédito, donde ambos algoritmos demostraron alta efectividad; y al igual que en nuestro estudio, XGBoost mostró una ligera ventaja debido a la capacidad de manejar datos complejos y realizar una optimización a través de la regularización. La concordancia en los resultados a través de diferentes aplicaciones marca la robustez y versatilidad de XGB, haciendo de esta técnica la mejor opción para tareas de clasificación en diversos contextos (Espinosa, 2020).
De igual manera comparamos con los resultados obtenidos en la investigación sobre el pronóstico del consumo de energía en la provincia de Chimborazo aplicando algoritmos Random Forest y XGBoost, a diferencia de nuestro estudio, la técnica de Random Forest mostró un ajuste superior en términos del error de la raíz cuadrada de la media en las predicciones y las tendencias en comparación con el consumo real, XGB también tuvo un buen rendimiento pero fue ligeramente inferior en algunas métricas específicas (Carrillo et al, 2023).

La relevancia de estos resultados se extiende más allá del ámbito académico. En un entorno empresarial y de seguridad cibernética, poseen aplicaciones prácticas importantes (Shafin et al., 2023). Las organizaciones que buscan mejorar su capacidad de detección de amenazas cibernéticas pueden considerar la implementación de RF o XGB como una estrategia efectiva (Doghramachi y Ameen, 2023). La elección entre estos dos algoritmos puede depender de factores específicos,

como la infraestructura tecnológica y los recursos disponibles. Además, los resultados pueden guiar la toma de decisiones en políticas de ciberseguridad a nivel gubernamental o empresarial. La inversión en la implementación de algoritmos de Machine Learning, respaldada por evidencia sólida de su eficacia, puede contribuir significativamente a la protección de activos digitales y la preservación de la privacidad en línea.

Para aumentar la eficacia operativa en el funcionamiento de los sistemas de seguridad, se han incorporado estos algoritmos. Ya que al reducir la cantidad de tiempo expuesto a troyanos y otras amenazas, se reduce el posible daño y se fortalece la resiliencia de los sistemas de información. Esto es fundamental en los sectores como financiero, de salud y el de infraestructura, donde la rapidez de respuesta es vital. (Kamboj et al, 2022)

En futuras investigaciones, se recomienda estudiar la combinación de RF y XGB con otras técnicas avanzadas, como el aprendizaje profundo y los modelos híbridos para comprobar su funcionamiento en situaciones de producción y en tiempo real, en la detección de amenazas. De igual manera, una dirección con potencial sería, desarrollar sistemas de detección adaptables que puedan cambiar con las amenazas emergentes y aprender de nuevos patrones de ataques.

Conclusiones

La evidencia presentada anteriormente demuestra que la evaluación de las técnicas de Machine Learning, con RF y Extreme XGB para identificar troyanos proporcionó resultados favorables, mostrando una precisión del 99% en la evaluación de los datos. La efectividad de estos modelos entrenados al procesar conjunto de datos extensos y complejos se basa en la detección con precisión de patrones maliciosos. La utilización a largo plazo se fundamenta en la capacidad de enfrentar ciberamenazas que se encuentra en constante evolución y también a adaptarse a comportamientos maliciosos nuevos, con el fin de mejorar la protección en un entorno digital que se ha vuelto cada vez más peligroso y complejo.

Además, la habilidad de estos modelos para lidiar con desbalances en los datos ha mejorado significativamente la efectividad en la identificación de troyanos, especialmente en escenarios con limitadas instancias positivas.

En la comparación del rendimiento entre las técnicas RF y XGB, se notó mínimas variaciones tanto en la precisión como en la sensibilidad, lo que nos indica que la decisión para la elección entre estas técnicas puede variar de acuerdo al tamaño y la

complejidad de los datos específicos. En este sentido se indica que la eficiencia computacional de XGB, destaca en entornos con recursos limitados, esto se debe a su mayor velocidad y menor consumo de memoria.

La implementación correcta de estos modelos en la detección de troyanos destaca su aplicabilidad práctica. Sin embargo, es esencial la evaluación continua y la adaptación a nuevas amenazas para mantener su efectividad en entornos del mundo real.

Finalmente se sugiere que investigaciones futuras se enfoquen en estudiar diferentes algoritmos de Machine Learning con sus respectivas técnicas para que proporcionen beneficios adicionales; también, sería útil la investigación en categorías de peligros cibernéticos y otros entornos de operación. Además, se podría contemplar la adquisición de conocimientos sobre los nuevos modelos de ataque en tiempo real, gracias a la creación de los sistemas de detección adaptables que evolucionan al mismo ritmo que los hacen las amenazas emergentes.

Referencias bibliográficas

Aldhyani, T.H.H. y Alkahtani, H. (2023). Cyber Security for Detecting Distributed Denial of Service Attacks in Agriculture 4.0: Deep Learning Model. Mathematics, 11(233). https://doi.org/10.3390/math11010233

Carrier, T., Victor, P., Tekeoglu, A. y Lashkari, A. (2022). Detecting Obfuscated Malware using Memory Feature Engineering. Proceedings of the 8th International Conference on Information Systems Security and Privacy (ICISSP 2022), 177-188. Doi: 10.5220/0010908200003120

Carrillo, D., Pazuña, W. y Quinatoa, C. (2023). Forecasting Energy Consumption in the Chimborazo Province, Ecuador, Using Random Forest and XGBoost Algorithms. 1st International Conference on Advanced Engineering and Technologies (ICONNIC), 66-72. Doi: 10.1109/ICONNIC59854.2023.10467276

Dasari, H. Pradhyumna Danduboina, B. y Chinna Rao, M. (2022). Malware Prediction Classifier using Random Forest Algorithm. INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN TECHNOLOGY (IJIRT), 9(2). Doi: 10.1109/ROPEC.2018.8661441

Diaz S. y Sanyer, W. (2021) Selección de candidatos para encuestas mediante técnicas de machine learning. Escuela Superior Politécnica del Litoral. 20. Pág. 31. Recuperado de: https://www.dspace.espol.edu.ec/bitstream/123456789/54068/1/T-110310%20SAMUEL%20DIAZ%20Y%20WLADIMIR%20SANYER.pdf

Doghramachi, D. F. y Ameen, S. Y. (2023). Internet of Things (IoT) Security Enhancement Using XGboost Machine Learning Techniques. Computers, Materials & Continua, 77(1). DOI: 10.32604/cmc.2023.041186

Espinosa, J. (2020). Application of Random Forest and XGBoost algorithms based on a credit card applications database. Ingeniería Investigación y Tecnología, 11 (3), 1 – 16. https://doi.org/10.22201/fi.25940732e.2020.21.3.022

Hu, J. y Szymczak, S. (2023). A review on longitudinal data analysis with random forest. Briefings in Bioinformatics, 24(2), 1 – 11. https://doi.org/10.1093/bib/bbad002

Kamboj, A. Kumar, P. K. Bairwa, A y Joshi, S. (2022). Detection of malware in downloaded files using various machine learning models. Egyptian Informatics Journal, 24(2023), 81 – 94. doi.org/10.1016/j.eij.2022.12.002

Kanimozhi, V. y Prem Jacob, T. (2020). Artificial Intelligence outflanks all other machine learning classifiers in Network Intrusion Detection System on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing. ICT Express, 7(2021), 366 - 370. doi.org/10.1016/j.icte.2020.12.004

Kaspersky. (2021). Machine Learning for Malware Detection. Machine learning application specifics in cybersecurity, 4 – 5. Retrieved from: https://media.kaspersky.com/en/enterprise-security/Kaspersky-Lab-Whitepaper-Machine-Learning.pdf

Louk, M.H.L. y Tama, B.A. (2022). Tree-Based Classifier Ensembles for PE Malware Analysis: A Performance Revisit. Algorithms, 15(332). https://doi.org/10.3390/a15090332

Prakash, Ch., Barthwal, A., Avikal, Sh. y Singh, G.K. (2023). FSAS: An IoT-Based Security System for Crop Field Storage. International Journal of Distributed Sensor Networks, 2023, 13. https://doi.org/10.1155/2023/2367167

Roy, P.B., Bhargava, M., Chang, Ch., Hui, E., Gupta, N., Karri, R. y Pearce, H. (2023). A survey of Digital Manufacturing Hardware and Software Trojans. arXiv, Massachusetts Institute of Technology. https://doi.org/10.48550/arXiv.2301.10336

Sanz, S. (2022). Uso De Algoritmos De Machine Learning Para La Detección De Archivos Malware. [Tesis de Maestría, Universidad Nacional de Educación a Distancia]. Recuperado de: http://e-spacio.uned.es/fez/eserv/bibliuned:master-ETSInformatica-CBS-Ssanz/Sanz_Garcia_Sergio_TFM.pdf

Shafin, S. S., Karmakar, G. y Mareels, I. (2023). Obfuscated Memory Malware Detection in Resource-Constrained IoT Devices for Smart City Applications. Sensors, 23(5348). https://doi.org/10.3390/s23115348

Sidhu, S. Mohd, B. J. y Hayajneh, T. (2019). Hardware Security in IoT Devices with Emphasis on Hardware Trojans. Sensor and Actuator Networks, 8(3), 42. doi:10.3390/jsan8030042.

Conflicto de intereses:

Los autores declaran que no existe conflicto de interés posible.

Financiamiento:

No existió asistencia financiera de partes externas al presente artículo.

Agradecimiento:

N/A

Nota:

El artículo no es producto de una publicación anterior.