Financiado por la Unión Europea NextGenerationEU
Gobierno de España · Vicepresidencia Primera del Gobierno · Ministerio de Asuntos Económicos y Transformación Digital · Secretaría de Estado de Digitalización e Inteligencia Artificial
Plan de Recuperación, Transformación y Resiliencia
España | digital 2026
incibe_ Instituto Nacional de Ciberseguridad

Alicia Mayor Bal
presenta el resumen del trabajo

Detección de URLs fraudulentas mediante Machine Learning

Nuria Reyes-Dorta, Pino Caballero-Gil, Carlos Rosa-Remedios
Actas VIII JNIC, atlanTTic, 329-336, (2023).

10 de julio de 2024

SERIE DE RESÚMENES EN ESPAÑOL

Palabras clave: Árbol de decisión, Deep learning, Red Neuronal, Support Vector Machine, URL.

Introducción

Entre el creciente número de ataques cibernéticos en los últimos años se encuentran las URLs fraudulentas, las cuales son cada vez más sofisticadas con el objetivo de engañar al usuario para obtener información confidencial.

Diferentes técnicas de Inteligencia artificial ayudan a desarrollar modelos para la detección de URLs fraudulentas. En este estudio se analiza la aplicación de Machine Learning, específicamente de Deep Learning para lograr una alta precisión en la **detección temprana** de estas URLs.

1. Metodología

Se hace uso de un **dataset ya existente**, el cual, tras un tratamiento para adaptarlo a este estudio (como eliminación de columnas no relevantes), cuenta con 181.916 filas de las cuales 8.618 pertenecen a URLs fraudulentas. Para el entrenamiento se usaron 145.533 datos con 6.959 URL fraudulentas.

A continuación, se comprobó la bondad de los datos en dos casos: sin analizar las URLs se identificaron el 80 % de URL fraudulentas y analizando las URLs, para lo cual se definieron nuevas variables (como la longitud de la URL o la cantidad de '/' que aparecen en ella), se identificaron el 86 % de ellas.

Tras la obtención de estas cifras, se **prepararon los datos** combinando las nuevas **variables** generadas para analizar las URL y algunas de las anteriores:

- ServerType: indica el servidor.
- CompromissionType: indica si el sitio está comprometido.
- PoweredBy: indica la plataforma que subyace al servidor.
- ContentType: indica tipo de codificación.
- ContentLength: indica tamaño del cuerpo del mensaje enviados al destinatario.

Además, se normalizaron los datos para todos los **algoritmos** usados menos para el Árbol de decisión, estos son: Regresión logística, Support Vector Machine y Redes neuronales.

2. Resultados

En primer lugar se analizan los resultados de todos los algoritmos menos los árboles de decisión y, para el caso de Support Vector Machine, se emplearon tres kernels diferentes. Para cada uno de ellos se establece la Precisión, el Recall y el F1-score.

Para URLs no fraudulentas los algoritmos que ofrecen mayor F1-score son el Árbol de decisión y el Support Vector Machine con kernel RBF (Función de Base Radial), ambos con un 99 %. Sin embargo, para URLs fraudulentas el mejor porcentaje de esa misma medida baja, siendo para el Árbol de decisión conun 89 %.

Para las redes neuronales se desarrollaron tres diferentes, variando el número y tipo de capas ocultas. Tras comparar las tres, se replantearon los modelos mediante Dropout (apagar para cada capa un porcentaje de neuronas) al verse sobreajustados.

Tras realizar y analizar la curva ROC, que representa la sensitividad frente a la especificidad de un sistema clasificador, se coloca en **primer lugar el Support Vector Machine**, luego las Redes Neuronales (en concreto una con tres capas ocultas, en dos de las cuales se usa una función sigmoide y en la otra la función de activación 'relu') y por último el Árbol de decisión.

3. Discusión/Conclusiones

El aumento de intentos de ataques mediante URLs fraudulentas hace necesario el continuo crecimiento de investigaciones como la de este estudio.

En función del tipo de dataset y algoritmos usados se pueden mejorar los resultados obtenidos, como se ha visto en los resultados con el Support Vector Machine tras el tratamiento de los datos iniciales. Es por esto, que este trabajo es un **proceso en progreso**, que pretende continuar mejorando la metodología usada para aplicarla tanto a este como otros a problemas de ciberseguridad.

4. Valoración del documento original

En este paper se explican brevemente todos los algoritmos usados en la metodología, lo cual es enriquecedor a la hora de entender el proceso. Además, es interesante como se trata la parte del dataset, añadiendo o eliminando columnas para favorecer la búsqueda de resultados.

Los resultados y conclusiones se exponen de una forma clara, pasando por todos los algoritmos probados y ordenándolos según el área de la Curva ROC.