Analisis del aporte del aprendizaje de maquinas a la seguridad de la informacion

name author[1] *

1- School of First Author - Dept of First Author
Address of First Author's school - Country of First Author's school
2- School of Second Author - Dept of Second Author
Address of Second Author's school - Country of Second Author's school

Abstract. Todo dispositivo informÃ¡tico tiene sus propios registros de seguridad. Al juntar miles de dispositivos que intervienen en las comunicaciones de computadores personales, bases de datos, servidores web, dispositivos de red, firewalls, etc., se genera un gran volumen de registros con informaciÃ³n.

1 Introduccin

el volumen de datos digitales debido al uso de tecnologias en organizaciones, lo que incrementa el riesgo de ataques informaticos. Las organizaciones consideran la information como un activo valioso y buscan protegerla con tecnologia eficiente y economica. Los delincuentes informaticos se vuelven mÃ¡s sofisticados, lo que impulsa el avance de la industria en la lucha contra las amenazas digitales. La Inteligencia Artificial y el Machine Learning son herramientas de ultima generacion que ayudan a contrarrestar estas amenazas. El Machine Learning utiliza algoritmos matematicos para analizar informacion y prever comportamientos similares a los humanos. Dada la evolucion constante de los ataques informaticos, las herramientas de analisis y deteccion, con intervencion humana y analisis de datos, permiten un tratamiento proactivo de los casos en las organizaciones que utilizan tecnologias de vanguardia para estos fines.

1.1 MATERIALES Y MeTODOS

En esta investigacion, se emplearon diversos materiales y metodos para explorar como el Machine Learning, una rama de la Inteligencia Artificial (IA), puede ser una herramienta efectiva en el fortalecimiento de la seguridad de la informacion en las organizaciones. Los materiales fundamentales incluyeron informacion recopilada de diversas fuentes, asi como conjuntos de pruebas que sirvieron como base para el analisis. A continuacion, se detallan los materiales y metodos utilizados en esta investigacion: Los materiales utilizados fueron Informacion Recopilada, Conjunto de Pruebas, Datos de Entrenamiento y Herramientas de Machine Learning, los metodos implemmentados fueron Optimizacion de Algoritmos, Analisis Predictivo y Clasificacion y Evaluacion.

*This is an optional funding source acknowledgement.

1.2 RESULTADOS

Se destaca la dependencia generalizada de organizaciones y personas en la tecnologia de la informacion para alcanzar objetivos empresariales y llevar a cabo actividades cotidianas. Sin embargo, esta dependencia tambien conlleva una serie de amenazas y vulnerabilidades en los entornos informaticos. Las amenazas pueden ser tanto internas como externas, representando la probabilidad de eventos o acciones que podrian causar danos a la informacion. Por otro lado, las vulnerabilidades se refieren a debilidades en la tecnologia o procesos relacionados con la informacion, intrinsecas a los sistemas de informacion o a la infraestructura que los sostiene. Acontinuacion se representan los tipos de amenazas:

1.2.1 Spyware

Codigo malicioso cuyo principal objetivo es recoger informacion sobre las actividades de un usuario en un computador.

1.2.2 Troyanos, virus y gusanos

Son programas de codigo malicioso, que de diferentes maneras se alojan en los computadores con el proposito de permitir el acceso no autorizado a un atacante.

1.2.3 El virus

tiene como objetivo principal ser destructivo, danando la informacion de la maquina.

1.2.4 Phishing

Es un ataque del tipo ingenieria social.

1.2.5 Spam

Recibo de mensajes no solicitados, principalmente por correo electronico.

1.2.6 Botnets

Son mÃ¡quinas infectadas y controladas remotamente, que se comportan como âzombisâ

1.2.7 Trashing

Un metodo cuyo nombre hace referencia al manejo de la basura.

1.2.8 Ransomware

Es un tipo de programa danino que restringe el acceso a determinadas partes o archivos del sistema infectado.

1.3 APRENDIZAJE SUPERVISADO

En algoritmos de Machine Learning (ML) supervisados, los datos de entrenamiento tienen etiquetas conocidas como spam/no-spam o precios de acciones. Un modelo se entrena haciendo predicciones y corrigiendolas cuando son incorrectas, continuando hasta alcanzar la precision deseada en los datos de entrenamiento. Este enfoque se utiliza para problemas de clasificacion y regresion, con ejemplos de algoritmos como Regresion Logistica y Red Neuronal de Propagacion Hacia Atras.

1.4 APRENDIZAJE NO SUPERVISADO

En algoritmos de aprendizaje no supervisados, los datos de entrada no tienen etiquetas ni resultados conocidos. Se crea un modelo deduciendo las estructuras en los datos, ya sea extrayendo reglas generales mediante procesos matematicos para reducir redundancia o organizando datos por similitud.

1.5 APRENDIZAJE SEMI-SUPERVISADO

En los algoritmos de aprendizaje semisupervisados, los datos de entrada incluyen ejemplos etiquetados y no etiquetados. A pesar de tener un problema de prediccion definido, el modelo debe aprender las estructuras de los datos y hacer predicciones.

2 ALGORITMOS AGRUPADOS POR SIMILITUD

Los algoritmos de aprendizaje automatico se agrupan segun sus funciones, como los metodos basados en arboles y los inspirados en redes neuronales.

2.1 Algoritmos de regresion

La regresion se ocupa de modelar la relacion entre variables que se refina iterativamente utilizando una medida de error en las predicciones hechas por el modelo.

2.2 Algoritmos basados en instancias

Los algoritmos basados en instancias son un tipo de algoritmos de aprendizaje automatico que se basan en la memorizacion de ejemplos especificos del conjunto de datos de entrenamiento.

2.3 Algoritmos basado en la regularizacion

Los algoritmos basados en la regularizacion son tecnicas utilizadas en el campo del aprendizaje automatico y la estadistica para prevenir el sobreajuste (overfitting) de modelos.

2.4 Algoritmos basado en arbol de decision

Los algoritmos basados en arbol de decision son metodos de aprendizaje automatico que se utilizan tanto en problemas de clasificacion como en problemas de regresion.

2.5 Algoritmos bayesianos

Los algoritmos bayesianos se basan en el teorema de Bayes, que es un principio fundamental en la teoria de la probabilidad.

2.6 Algoritmos de aprendizaje de reglas de asociacion

Los algoritmos de aprendizaje de reglas de asociaciÃ^3*nsontÃ©cnicasutilizadasenminerÃ-adedatosparadescubrirpatronesyrelacionesinteresantesenconjuntosdedatos*.

2.7 Algoritmos de agrupamiento

Los algoritmos de agrupamiento, tambien conocidos como algoritmos de clustering, son tecnicas de aprendizaje automatico utilizadas para dividir un conjunto de datos en grupos o clusteres basandose en la similitud entre los elementos.

2.8 Algoritmos de redes neuronales artificiales

Los algoritmos de redes neuronales artificiales son modelos computacionales inspirados en el funcionamiento del cerebro humano.

2.9 Algoritmos de aprendizaje profundo

Los algoritmos de reLos algoritmos de aprendizaje profundo, tambien conocidos como algoritmos de deep learning en inglÃ©s, son un conjunto de tÃ©cnicas de aprendizaje automatico basadas en redes neuronales artificiales con multiples capas interconectadas.

3 DISCUSIoN

El avance del Machine Learning desde sus inicios en los anios 50 ha sido notable, pero no esta exento de preocupaciones. Las crecientes aplicaciones de esta tecnologia han suscitado inquietudes sobre la privacidad y la seguridad de los datos, asi como sobre la opacidad de los modelos de Machine Learning, que a menudo operan como "cajas negras". Ademas, existe la preocupacion por la automatizacion del empleo y el posible impacto negativo en la economia. A pesar de su progreso, estas cuestiones eticas y sociales subrayan la necesidad de un uso responsable del Machine Learning en beneficio de la sociedad.

References

[1] C. H. Tarazona. *Amenazas informaticas y Seguridad de la informacion.* DerecPenaly-CriminXXIX, 2007.

[2] R. Duda, P. Hart y D. Stork *Pattern classification.*New York: Wiley Sons, 2001.

[3] P. Domingos*The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World (Ed. rev).* New York: Basic Books, 2015.

[4] Brownlee.*A tour of Machine Learning Algorithms, Melbourne: Jason Brownlee.*2013