**Machine learning y seguridad: deteccion de intrusiones en redes**

1. Intro

IDS: monitoreo y análisis de eventos para detectar incidentes de seguridad.

Aprendizaje automático para detectar anomalías.

1. Recolección de datos

Extraer características de los paquetes (duración de conexiones, tamaño, protocolos, comportamientos estadísticos como bits por segundo). con esto se etiquetan en dos: benigno y maligno (DoS hulkk, ddos, portscan, brute force, webattack, etc).

preprocesamiento

* limpieza de datos (eliminar valores nulos, duplicados, etc) → caracteristicas categoricas (hiperplano para representarlo) → normalizacion () → division de datos

entrenamiento del modelo

hay diferentes modos de entrenamiento, regresion logistica, deep neural networks, etc.

Para evaluarlo usaoms el recall.

Nos fijamos en los falsos positivos (primera fila).

Hay distintos problemas (falsos positivos ABIDS, falta de interpretabilidad, calidad de los datos, ataques advdersariales, envenenamiento de los datos).

* optimizacion d eumbrales y puntuacion de riesgo –< sistemas hibridos → aprendizaje activo y supervision humana → tecnicas de argupamiento y correlacion de eventos

Mercado & Futuro

Se usa en muchas areas e industrias, de medico a bancos, etc.

Splunk y paloalto y dark trace son los principales que lo han adoptado.

Objetio es tener un SOc mas inteligente: aprendizaje por refuerzo (sin necesidad de intervencion humana); aprendizaje federativo, etc.

ejemplo: alta tasa falsos positivos

dark trace se dedica para reducirlos. agruparon y priorizaron eventos, configurarion tthresholds e integraron feedbackd e analistas.

Falta de interpretabilidad en ABIDS

En deep learning tenemos una caja negra que no sabemos 100% como funciona, de modo que es dificil debugarlo. No se sabe por que se clasifican ciertas cosas de la manera que salen. Entonce se tiene que investigar a ciegas.

soluciones:

midelos explicables por diseño (más pequeños con menor precision: arboles, regresion, etc) →

tecnicas de explainable AI (XAI) (LIME, SHAP, Grad-CAM) →

sistemas hbirdos

Mercado & Futuro

Modelos con capas interpretables, ingegracion al ciclo de vida del sistema, auditoria uatomatizada de decisiones, colaboración humano-ai. Se utiliza en ibm, microsoft.

CAso SAS: detectaron reclamaciones y transacciones atipicas, microsoft lo soluciono haciendo que cada alerta rhiciera un reporte de las variables mas influyentes y una explicacion para reducir tiempo de investigacion y obtener mejor confianza por parte del equipo.