Classificação Automática de Alertas de Segurança num SOC

Tomás Ladeiro Domingues | 1231443

01

PROBLEMA

Descrição do problema a apresentar

02

OBJETIVOS

Objetivos a serem alcançados

03

WORK BREAKDOWN STRUCTURE (WBS)

Planeamento do projeto

04

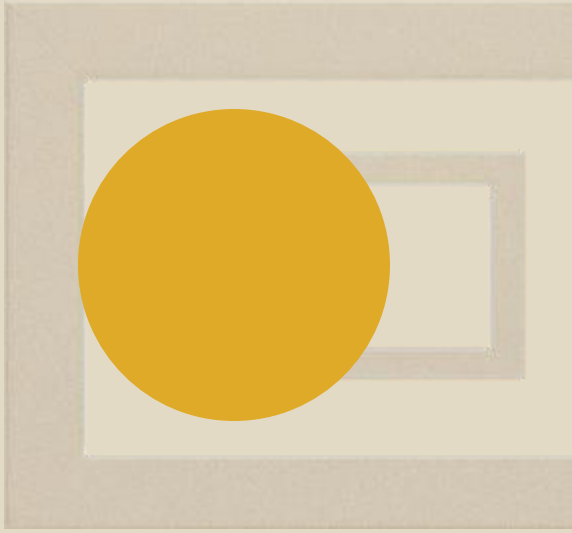
ESTADO DA ARTE

Introdução teórica e trabalhos existentes relacionados com o problema

05

FERRAMENTAS DE TRABALHO

Comparação e análise de possíveis ferramentas a utilizar.

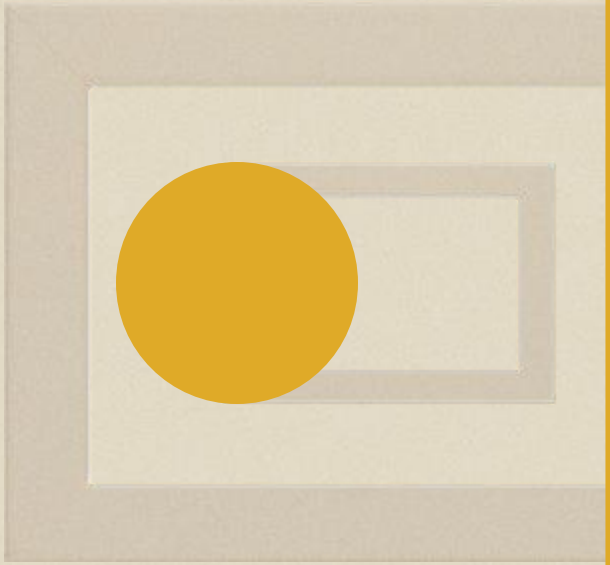


INTRODUÇÃO

A gestão eficiente de alertas de segurança é um desafio constante nos Centros de Operações de Segurança (SOC).

O volume elevado de alertas e a necessidade de identificar ameaças rapidamente tornam o processo de triagem inicial fundamental para a eficácia de um SOC.

Este trabalho tem como objetivo principal desenvolver um programa independente capaz de automatizar a triagem de alertas de segurança em um ambiente SOC.



PROBLEMA

Análise Manual

A triagem inicial de alertas depende da inspeção manual, o que torna o processo demorado e suscetível a erros humanos.

Tempo de Resposta

A análise manual aumenta o tempo de deteção (MTTD) e o tempo de resposta (MTTR), permitindo que ameaças permaneçam indetetáveis por mais tempo.

Sobrecarga de Alertas

Os analistas lidam com um volume excessivo de alertas, dificultando a priorização e aumentando o risco de não detetar ameaças críticas.

PROBLEMA

Perda de Ameaças

A abordagem reativa aumenta o risco de perder ameaças importantes devido à sobrecarga de alertas e falhas na triagem.

Fadiga dos Analistas

A sobrecarga de trabalho provoca fadiga nos analistas, reduzindo a eficácia e a precisão na triagem dos alertas.

Fragilidade na Segurança

A análise manual compromete a postura de segurança da organização, tornando-a mais vulnerável a ataques e riscos de brechas.

Objetivos do Estudo

Automatizar a triagem de alertas de segurança em SOCs.

Pesquisa e Análise

Realizar uma revisão de literatura sobre automação e Machine Learning em SOCs.

01

- Identificar ferramentas e métodos de automação usados em SOCs.
- Estudar modelos de ML aplicados à triagem de alertas.

Integração com SIEM

Desenvolver e ligar o modelo de ML ao SIEM e outras fontes de dados.

02

- Compilar datasets para treinar o modelo de Machine Learning.
- Implementar ingestão de dados em tempo real a partir de um SIEM.

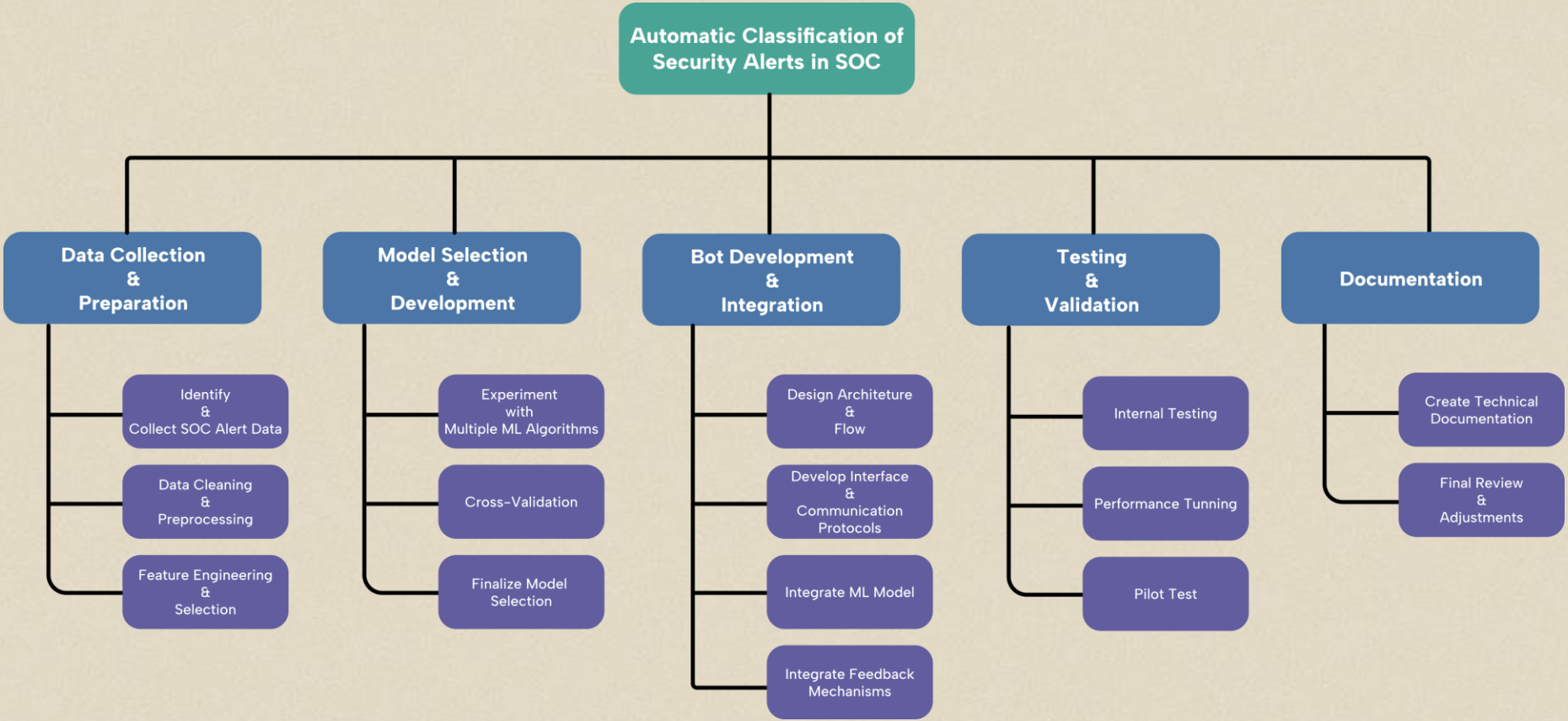
Avaliação de Modelos

Testar e avaliar a eficácia do modelo de ML.

03

- Criar um dashboard para feedback dos analistas sobre o modelo.
- Reduzir falsos positivos em triagem de alertas de segurança.

WBS


Automatic Classification of Security Alerts in SOC
Data Collection & Preparation
Identify & Collect SOC Alert Data
Data Cleaning & Preprocessing
Feature Engineering & Selection
Model Selection & Development
Experiment with Multiple ML Algorithms
Cross-Validation
Finalize Model Selection
Bot Development & Integration
Design Architeture & Flow
Develop Interface & Communication Protocols
Integrate ML Model
Integrate Feedback Mechanisms
Testing & Validation
Internal Testing
Performance Tunning
Pilot Test
Documentation
Create Technical Documentation
Final Review & Adjustments

ANÁLISE TEÓRICA - SOC

01

PREVENIR

Prevenir incidentes com medidas proativas (scans e análises).

MONITORIZAR

Monitorizar, detetar e analisar intrusões.

02

SOC

04

INFORMAR

Informar stakeholders sobre incidentes, tendências e ameaças emergentes.

RESPONDER

Responder a incidentes confirmados e coordenar recursos.

03

ANÁLISE TEÓRICA - SIEM

RECOLHA

Recolha e agregação de logs (métodos push/pull).

ANÁLISE

Análise e deteção de ameaças em tempo real.

RELATÓRIOS

Relatórios e visualização para investigações detalhadas.

Gerenciamento e Correlação de Eventos de Segurança

NORMALIZAÇÃO

Uniformiza logs de diferentes fontes para consistência.

CORRELAÇÃO

Liga eventos dispersos para detetar padrões maliciosos.

AUTOMAÇÃO

Gera alertas e notifica em tempo real sobre incidentes.

ANÁLISE TEÓRICA - ML

MACHINE LEARNING

SUPERVISIONADA

Usa dados rotulados para treinar modelos, como Random Forest.

NÃO SUPERVISIONADA

Agrupa dados sem rótulos, útil para identificar padrões desconhecidos e anomalias em grandes volumes de dados.

REFORÇO

Modelos aprendem continuamente interagindo com o ambiente, adaptando-se a padrões emergentes.

ANÁLISE TEÓRICA - ML

DESAFIOS

Overfitting ocorre quando o modelo memoriza dados específicos; underfitting, quando não captura padrões essenciais. Ambos comprometem resultados.

DEPENDÊNCIAS

Modelos baseiam-se em dados históricos para aprendizado; datasets limitados ou desatualizados afetam a capacidade de detectar novas ameaças.

ANÁLISE TEÓRICA – ML MODELOS

ÁRVORES DE DECISÃO

Simples, interpretáveis, mas propensas a overfitting.

CLASSIFICADORES ENSEMBLE

Random Forest e XGBoost; maior precisão e robustez.

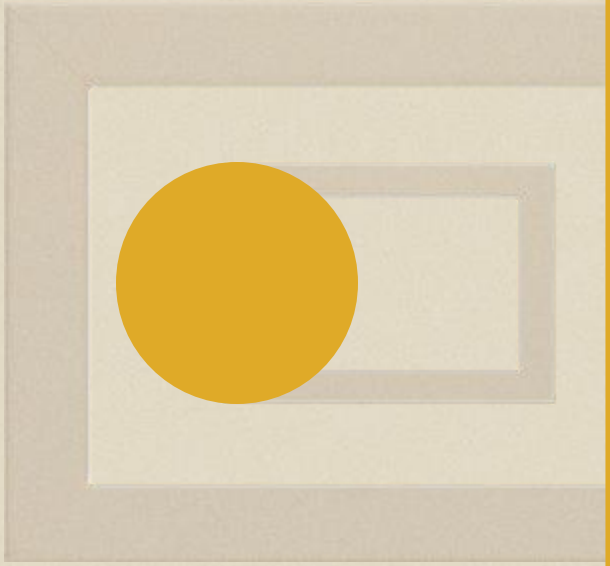
Naive Bayes

Rápido e eficiente, mas assume independência entre variáveis.

OUTROS TRABALHOS

A partir de trabalhos existentes ligados com os objetivos deste projeto, podemos retirar:

1. A necessidade de automatizar a triagem de alertas em SOCs
2. Modelos híbridos apresentam elevada precisão e baixo FPR
3. Integração de algoritmos de ML em ambientes SIEM
4. Técnicas de rastreamento de fluxo de informações
5. Modelos baseados em aprendizagem por reforço
6. Filtros probabilísticos são eficazes em ameaças específicas



FERRAMENTAS - SIEM

QRADAR

- Integração profunda com o ecossistema IBM, incluindo SOAR e ferramentas de automação.
- Deteção avançada de ameaças baseada em anomalias.
- Arquitetura modular e escalável, ideal para empresas de todos os tamanhos.
- Suporte a infraestruturas modernas, incluindo ambientes híbridos e nativos da nuvem.

SPLUNK

- Focado em análises em tempo real e processamento de grandes volumes de dados.
- Ideal para caça a ameaças e gestão de conformidade.
- Extensa integração com ferramentas de terceiros e suporte a personalização.
- Oferece dashboards predefinidos e alertas personalizados no Splunk Enterprise Security (ES).

LOGRHYTHM

- Fortemente orientado para conformidade (GDPR, HIPAA, PCI-DSS).
- Deteção de ameaças baseada em machine learning e anomalias comportamentais.
- Gestão do ciclo de vida de ameaças com fluxos de trabalho estruturados.
- Integração fluida com ferramentas externas e dashboards centralizados.

FERRAMENTAS - TICKETING

QRADAR SOAR

- Automação com playbooks para respostas consistentes e rápidas.
- Gestão de casos com rastreamento centralizado de incidentes, evidências e comunicações.
- Integração com EDRs e serviços na nuvem para maior eficiência na resolução de incidentes.

JIRA

- Ferramenta intuitiva e amplamente utilizada para gestão de incidentes em SOCs
- Ideal para equipas já familiarizadas com o ecossistema JIRA.
- Sistema de tickets fácil de usar para criação, atribuição e atualização de alertas.
- Integrações extensas com ferramentas de SIEM, mas funcionalidades de automação limitadas.

SERVICENOW

- Solução versátil para gestão de serviços e incidentes em TI e cibersegurança.
- Módulo Security Incident Response (SIR) automatiza triagem e escalonamento de incidentes.
- Painéis interativos para monitorização e relatórios em tempo real.

FERRAMENTAS - ML

SCIKIT-LEARN

- Biblioteca focada em algoritmos clássicos, como Random Forest e SVM.
- Ideal para tarefas de classificação, regressão e clusterização.
- Integração com pandas e NumPy para manipulação de dados eficiente.
- Ferramenta robusta para prototipagem rápida, mas limitada para deep learning

TENSORFLOW

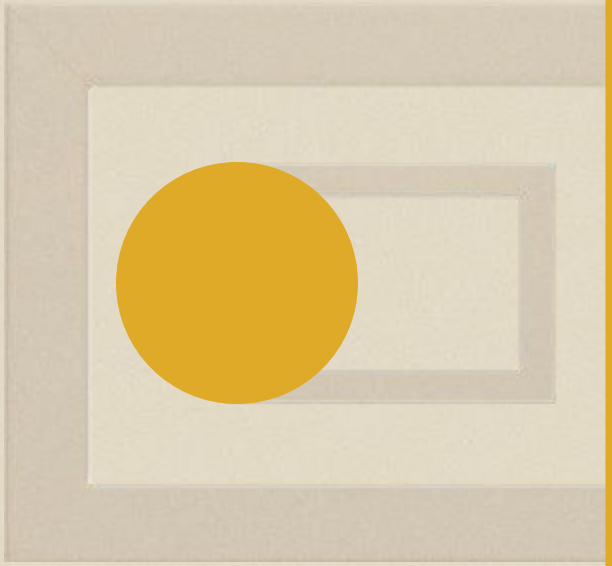
- Desenvolvido pelo Google, suporta machine learning e deep learning em larga escala
- Excelente para produção em ambientes móveis, edge computing e na nuvem.
- Ferramenta TensorBoard para monitorização de desempenho e ajustes

PYTORCH

- Destaca-se em processamento de linguagem natural e visão computacional.
- Integração com ONNX e NumPy para maior compatibilidade e portabilidade.
- Versátil para pesquisa e experimentação, mas requer experiência adicional para produção.

Benefícios da Automação no SOC

- Redução da carga de trabalho dos analistas:
- Melhoria no MTTD e MTTR:
- Prioritização automática de alertas críticos:
- Fortalecimento da postura de segurança:



Diferenciação do Projeto

- Integração com dados dinâmicos e sistemas existentes
- Desenvolvimento de um dashboard personalizado
- Adaptabilidade a novas ameaças
- Alinhamento com necessidades reais do SOC