MACHINE LEARNING: APLICABILIDADE EM MONITORAMENTO DE REDES
MACHINE LEARNING: APPLICABILITY IN NETWORK MONITORING

Amadeu José Marchi, FATEC Americana, Segurança da informação,
amadeu.marchi@fatec.sp.gov.br
Jonas Bodê, FATEC Americana, Segurança da informação,
jona.bode@fatec.sp.gov.br
Mauricio Zazeri Fonseca, a FATEC Americana, Segurança da informação,
mauricio.fonseca01@fatec.sp.gov.br

Resumo

Almejando uma maior eficiência, agilidade e produtividade, diversas organizações tendem a ampliar suas conectividades e sistemas não apenas às redes internas, mas também à Internet. Tal ampliação é capaz de se provar nociva à segurança da informação e sistemas daquela organização; com isso, a aplicação de sistemas de gerenciamento e monitoramento de redes tem se tornado cada vez mais necessária. Porém, com o aumento das requisições e acessos a um dado sistema, este monitoramento pode tornar-se cada vez mais complexo. O presente texto busca analisar a efetividade da instauração de algoritmos de *machine learning* ou, em português, aprendizado de máquina em sistemas de monitoramento e gerenciamento de redes de computadores com o intuito de se tornar um aliado destes. Este artigo apresenta também a conceitualização de *machine learning* e de seus diferentes tipos e técnicas, além de expor um modelo prático da aplicação de um algoritmo treinado em um sistema de monitoramento com a finalidade de analisar a sua utilidade, aplicabilidade e efetividade na automação da análise constante e no auxílio à segurança em redes de computadores e na tomada de decisões perante o gerenciamento da rede.

Palavras-chave: Aprendizado de máquina, monitoramento, gerenciamento, segurança em redes de computadores.

Abstract

Looking for greater efficiency, agility, and productivity, several organizations tend to expand their connectivity and systems not only to internal networks but also to the internet. Such expansion can prove detrimental to the security of the organization's information and systems; as a result, the implementation of network management and monitoring systems has become increasingly necessary. However, with the increase in requests and access to a given system, this monitoring can become more complex. This text aims to analyze the effectiveness of implementing Machine Learning (ML) algorithms in computer network monitoring and management systems, in order to become an ally to them. This text also presents the conceptualization of Machine Learning and its different types and techniques, in addition to showcasing a practical model of applying a trained algorithm in a monitoring system to analyze its utility, applicability, and effectiveness in automating continuous analysis and aiding in computer network security and decision-making in network management.

Keywords*: Machine learning, monitoring, management, network computing security.*

1. Introdução

Com o aumento do uso da Internet e a crescente dependência de sistemas informatizados, a segurança da informação tem se tornado uma preocupação cada vez mais presente em diversas áreas. O monitoramento de redes é uma das medidas mais eficazes para garantir a segurança dos sistemas e prevenir ataques cibernéticos.

No entanto, o monitoramento de redes pode se tornar um desafio para empresas e organizações, especialmente quando se trata de redes de grande porte. Nesse contexto, a utilização de técnicas de *machine learning* (ML) tem se mostrado uma solução promissora para aprimorar a eficácia do monitoramento de redes.

O *machine learning* é uma das áreas da inteligência artificial que tem como objetivo desenvolver algoritmos capazes de aprender com os dados e tomar decisões sem serem explicitamente programados para isso. No contexto da segurança da informação, essas técnicas podem ser aplicadas para detectar ameaças, identificar padrões de comportamento suspeitos e prever ataques antes que eles aconteçam.

O *machine learning* também pode ser utilizado para automatizar o processo de detecção de ameaças e monitoramento de rede, reduzindo a carga de trabalho dos profissionais responsáveis pela segurança da informação. Isso permite que eles se concentrem em tarefas mais estratégicas, como análise de dados e tomada de decisões, em vez de se dedicar a tarefas repetitivas e de baixo valor agregado.

Outro benefício do uso de técnicas como estas no monitoramento de rede é a capacidade de lidar com grandes volumes de dados em tempo real. Essas técnicas são capazes de processar e analisar uma grande quantidade de informações em segundos, permitindo que os profissionais de segurança da informação ajam rapidamente em caso de ameaças ou ataques cibernéticos.

Diante desses benefícios, o uso de técnicas de *machine learning* para o monitoramento de rede tem se tornado cada vez mais popular entre empresas e organizações de todos os tamanhos.

Segundo Michael Chui *et al.* (2018) o uso de sistemas autônomos inteligentes é inevitável, uma vez que 28,5 bilhões de dólares foram destinados para *machine learning* em todo o mundo durante o primeiro trimestre de 2018. As maiores empresas do mundo, como Google, IBM, Apple, entre outras, já utilizam essa tecnologia há anos e aprimoram cada vez mais seus usos. A automatização de serviços se torna cada vez mais comum. No entanto, é importante ressaltar que essa tecnologia não deve ser utilizada como uma ferramenta final em um serviço, trabalho ou projeto.

O ChatGPT, um programa de *machine learning* desenvolvido pela OpenAI, por exemplo, é altamente eficaz em várias áreas do conhecimento, encurtando horas de pesquisa e auxiliando profissionais em seu dia a dia de trabalho. No entanto, esse *software* não substitui a função humana. Ele é eficaz como uma ferramenta de apoio, mas como gerador de um produto completo, pode apresentar diversas falhas.

Portanto, o uso de aprendizado de máquina no monitoramento de redes será uma ferramenta auxiliar para os profissionais de segurança, ajudando na detecção de possíveis falhas ou vulnerabilidades na rede. É importante destacar que a tecnologia deve ser utilizada em conjunto com a capacidade humana de análise e interpretação de dados para que sejam obtidos os

melhores resultados.

Assim, o objetivo deste trabalho é demonstrar o uso de técnicas de *machine learning* para auxiliar no monitoramento de redes como uma solução eficaz para a segurança da informação. E, como objetivos específicos:

- Desenvolver modelos inteligentes para auxiliar no monitoramento de redes.
- Simular uma rede de computadores de uma empresa.
- Exemplificar a aplicabilidade desta tecnologia no âmbito da segurança da informação.
- Analisar a eficiência da Inteligência Artificial (IA) para o monitoramento.

2. Referencial Teórico

2.1 Machine Learning

O *machine learning* é uma técnica de análise de dados que permite que um sistema computacional aprenda e melhore com base em exemplos e experiências passadas, sem a necessidade de ser explicitamente programado para isso. De acordo com Zhou (2021), *o machine learning* pode ser definido como:

> [...]Uma abordagem de computação baseada em algoritmos que permitem que um sistema aprenda e melhore com base em dados e experiências anteriores, sem a necessidade de programação explícita. Esses algoritmos permitem que o sistema identifique padrões, extrapole insights e tome decisões com base nas informações disponíveis.

Zhou (2021) destaca que existem várias técnicas de *machine learning*, incluindo aprendizado supervisionado, aprendizado não supervisionado e aprendizado por reforço. Cada técnica é adequada para diferentes tipos de problemas e dados. O aprendizado supervisionado, por exemplo, é usado quando há um conjunto de dados rotulados disponíveis para treinar o modelo. Já o aprendizado não supervisionado é utilizado quando não há rótulos disponíveis e o objetivo é identificar padrões e estruturas no conjunto de dados. O aprendizado por reforço, por sua vez, é usado para ensinar um agente a tomar decisões em um ambiente dinâmico e complexo, portanto o presente projeto baseou-se majoritariamente no desenvolvimento de um modelo de aprendizado supervisionado, uma vez que o modelo poderia ser treinado por meio de dados rotulados no intuito de classificar ou identificar padrões em novos dados não vistos.

2.2 Monitoramento e Gerenciamento De Redes

Para Black (2008), a constante evolução tecnológica, idem a usuário e aplicação, impulsiona o crescimento de redes computacionais tanto internas quantos externas, alavancando a necessidade de monitorar estas redes a fim de manter-se íntegro e correto o seu funcionamento.

Ademais, segundo LEE, LEVANTI e KIM (2014) a evolução das redes de computadores tem levado a uma demanda cada vez maior por ferramentas de monitoramento mais sofisticadas e eficientes. É apresentado pelos autores uma revisão das técnicas de monitoramento de redes, incluindo os métodos passivos e ativos, e destacam as principais características e desafios de cada um desses métodos. Os métodos passivos de monitoramento de redes envolvem a captura de pacotes de dados que trafegam na rede. Esses pacotes são analisados para fornecer informações sobre o tráfego de rede, incluindo o volume de dados, o tempo de resposta, a qualidade do serviço, entre outros. Esse tipo de monitoramento é eficiente e não interfere no tráfego da rede, mas pode gerar uma grande quantidade de dados que precisam ser processados e armazenados. Já os métodos ativos de monitoramento de redes envolvem o envio de pacotes de teste pela rede para avaliar seu desempenho e sua disponibilidade. Esse tipo de monitoramento pode ser mais eficiente em detectar problemas de rede em tempo real, mas pode gerar tráfego adicional nela e afetar o desempenho desta.

Esse grupo de pesquisadores destaca que, independentemente do método utilizado, é importante que as ferramentas de monitoramento de redes sejam capazes de fornecer informações em tempo real, de forma precisa e confiável. Além disso, é importante que essas ferramentas sejam capazes de identificar anomalias e alertar os administradores de rede sobre possíveis problemas. Perante tais afirmações, as capacidades destacadas tornaram-se a base métrica para o desenvolvimento prático do projeto, agindo como uma lista de requisitos cujos quais deveriam ser atendidos.

A equipe de pesquisa também discute as tendências futuras no monitoramento de redes, incluindo o uso de técnicas de Inteligência Artificial e aprendizado de máquina para analisar e identificar padrões de tráfego de rede. Eles destacam que essas técnicas podem melhorar a eficiência e a precisão do monitoramento de redes e permitir uma detecção mais rápida e precisa de possíveis problemas.

Portanto, é a partir da ideia discutida pelos autores que o presente artigo buscou indagar a efetividade de se aplicar tais técnicas computacionais no intuito de se analisar sua efetividade e viabilidade em constante aplicação do monitoramento contínuo de redes computacionais.

2.3 Ferramentas de monitoramento e gerenciamento de redes.

Segundo Black (2008), o gerenciamento de redes permite controlar os recursos desta além de identificar e prevenir possíveis problemas em seu funcionamento e ou gerência. A realização do monitoramento preza pelas características de segurança, desempenho, configuração, contabilização e falhas sendo efetuado a partir da utilização de *softwares* tanto

disponíveis sob licença de software livre quanto ferramentas proprietárias.

As ideias de Belentani *et al.* (2018) assimilam-se às de Black (2008), uma vez que, segundo estes, é necessário um gerenciamento eficaz de redes de computadores a partir da crescente disseminação destas associado a integração densa de componentes e o elevado número de ferramentas para o auxílio de gestão.

Podem ser citadas algumas ferramentas populares dedicadas à essa atividade como, por exemplo, Zabbix, Manage OP Engine e Nagios. Para Ramos, Nascimento e Bischoff (2022), a utilização destes ou de outros softwares já existentes pode não ser a melhor solução para uma organização, com a justificativa de que:

> O desenvolvimento de uma ferramenta personalizada permite que sejam incorporadas funções a ela, atendendo assim a demandas específicas de cada empresa

Portanto, o uso de ferramentas comerciais pode ser limitado ou pouco flexível, enquanto a criação de uma ferramenta própria permite customização, escalabilidade e maior controle sobre a solução, o que é de suma importância para o desenvolvimento deste projeto uma vez que seu intuito não é se aprofundar nas especificidades e singularidades das ferramentas comerciais.

3. Materiais e Métodos

O presente artigo utiliza a modalidade de uma pesquisa aplicada objetiva visando a implementação simulada de um cenário para a aplicação de um algoritmo baseado no aprendizado de máquina como ferramenta de auxílio para um sistema de monitoramento de redes em uma rede empresarial também simulada.

3.1.1 Requisitos de Hardware

No intuito de suprir a necessidade de máquinas servidoras com alto desempenho e dispositivos de rede específicos para um dos cenários de teste a tratar, foi-se utilizada apenas uma máquina física configurada para suprir máquinas virtuais por meio da utilização de um *software* específico descritos a seguir.

3.1.2 Requisitos de Software

A ambientação da infraestrutura simulada foi baseada, como previamente citado, em um ambiente de virtualização utilizando o *software* Oracle VM VirtualBox versão 6.1 disponibilizado sob GNU (General Public License), Licença Pública Geral.

Já para o desenvolvimento da Inteligência artificial foi utilizado o Google Colab Notebook

3.2.1 Metodologia para definição da arquitetura de Rede

Primeiramente, foi projetada uma topologia de rede que simula uma infraestrutura empresarial, incluindo servidores, estações de trabalho e dispositivos de rede para que se tornasse possível criar massas de dados com base em cenários modestamente inspirados em estruturas reais.

3.2.2 Geração e pré-processamento de conjuntos de dados

Para a aplicação do modelo tornou-se necessária a obtenção de uma massa de dados que pudesse ser classificada e utilizada para o treinamento deste. Evidentemente era necessária uma quantidade grande de dados para que o treinamento do algoritmo pudesse se tornar o mais efetivo e preciso possível. Após a obtenção dos dados, os mesmos seriam enviados para a plataforma *Google Colab Notebook* para manipulação utilizando a linguagem de programação *Python* por meio do uso também da biblioteca *Pandas*, transformando os arquivos com os dados dos protocolos em *Data Frames*.

Pensando nos diferentes cenários que poderiam ser abordados para esta coleta, a primeira geração de dados foi realizada a partir da gravação de dados aleatórios e sem quaisquer cenários, padrões, ou extrações reais, apenas inserindo assim manualmente as informações em um arquivo *.csv*. Para o treinamento do modelo com base nesses dados, foram criadas no *software*, *excel*, as colunas de endereço de origem e destino do pacote, porta de origem e destino, protocolo, tempo de conexão, tamanho do pacote e uma coluna *label* para servir de classificação daquele pacote, nesse cenário criado podendo ser classificado nesta como *"Normal"* ou *"Attack"*.

Para um segundo cenário de testes, foi aplicado a partir da definição anteriormente realizada da arquitetura de rede um cenário de máquinas virtuais que se comunicariam e realizariam requisições e transferências de pacotes entre si. Toda essa comunicação seria então registrada a partir do uso da ferramenta *TCP Dump*, para então ser tratada e enviada para um arquivo também *.csv* para a aplicação no treinamento de outro treinamento de um modelo. A comunicação dessa máquina para este estudo se baseou integralmente no uso do protocolo ICMP e na ferramenta HPING3, primeiramente realizando transferências de pacotes de tamanhos exorbitantes por essa ferramenta para classificá-los como *"Attack"* e posteriormente realizando apenas *Pings* comuns e classificando-os como *"Normal"*.

Com a arquitetura de rede definida anteriormente, o tráfego de rede foi simulado e capturado. Essa captura de tráfego foi então processada para extrair características relevantes, como padrões de comunicação, protocolos predominantes e volume de tráfego por dispositivo. Os dados capturados foram submetidos a etapas de pré-processamento, incluindo normalização, tratamento de valores ausentes e transformação de dados categóricos em formatos numéricos adequados para alimentar os modelos de *machine learning.*

3.2.3 Seleção e Treinamento dos Modelos

Foram selecionados diversos algoritmos de *machine learning*, como redes neurais, árvores de decisão e máquinas de vetores de suporte como possíveis candidatos à utilização. Dentre estes, o modelo escolhido para a aplicação prática foi o modelo de aprendizado de máquina supervisionado, por se adequar melhor com o objetivo deste artigo. Partindo da ideia de classificar as linhas do *Data Frame* por um *label* entre *"Attack"* e *"Normal"*.

3.2.4 Avaliação dos Modelos

Por fim, os modelos foram avaliados por meio do conjunto de dados de validação, utilizando métricas como acurácia, precisão, *recall* e F1-*score*.

A acurácia é a métrica que mede a proporção de previsões corretas em relação ao total de previsões feitas pelo modelo, proporcionando uma visão geral da qualidade das previsões.

A unidade de medida de precisão avalia a proporção de previsões positivas corretas em relação ao total de previsões positivas, destacando a capacidade do modelo de evitar falsos positivos em situações de monitoramento de redes.

O *recall*, mede a proporção de previsões positivas corretas em relação ao total de casos positivos reais, enfatizando a capacidade do modelo de identificar todos os casos verdadeiros, minimizando os falsos negativos no contexto aplicado.

Para obter uma avaliação com um melhor equilíbrio utilizou-se o *F1-score*, que combina precisão e recall em uma única pontuação, calculada como a média harmônica entre essas duas métricas. Isso é especialmente útil quando é necessário encontrar um equilíbrio entre a precisão e o recall, algo comum em problemas de monitoramento de redes, onde um compromisso adequado entre detecção precisa e minimização de alarmes falsos é essencial.

4. Resultados e Discussões

4.1. Obtenção de dados para treinamento do algoritmo

A obtenção de dados para o treinamento do modelo pode ser considerada como sendo a parte mais importante e mais complexa para que as taxas de precisão deste estivessem o mais próximo possível de 1 (cem por cento). Como a aplicação prática deste artigo volta-se a um ambiente simulado, a capacidade dos dados utilizados para o treinamento do modelo se espelharem fielmente às situações enfrentadas nos ambientes empresariais reais acaba por ser comprometida, dificultando a estimativa da precisão de acerto do modelo treinado caso fosse aplicado em um ambiente real. Portanto, para a aplicação de um modelo treinado com algoritmos de aprendizado de máquinas a um ambiente prático real, é fundamental que o mesmo seja treinado a partir de dados também reais.

4.2. Comportamento de um modelo treinado a partir de dados completamente randômicos

Como mencionado anteriormente, o primeiro cenário testado teve como base o treinamento de um modelo a partir de um conjunto de dados gerados manualmente, inserindo-os em um arquivo de extensão *.csv* sem quaisquer critérios, padrões ou cenários. É interessante observar que mesmo assim a pontuação de precisão e o "f1-score" para classificar pacotes como "normais" mostrou-se deveras satisfatória. Mas é perceptível que o modelo não conseguiu definir padrões ou lógicas suficientes para estimar pacotes suspeitos com uma precisão ideal (classificados como "attack" neste cenário). Como pode ser

observado na Figura 1 e Figura 2.

Figura 1 – *Data Frame* eixo x do modelo de treinamento com dados aleatórios

```
x.iloc[:10]
```

	source_port	dest_port	packet_size	connection_duration	SourceIPFrequency	DestIPFrequency
0	12345	54321	84	0.0	20	20
1	23456	80	524	5.0	10	15
2	54321	23456	256	5.0	20	15
3	13579	54321	60	5.0	15	20
4	54321	12345	1084	5.0	5	15
5	54321	23456	512	5.0	20	15
6	13579	54321	64	5.0	15	20
7	23456	25	1024	5.0	10	15
8	54321	23456	512	5.0	20	15
9	13579	54321	64	5.0	15	20

Fonte: Autoria Própria, Google Colab Notebook

Figura 2 – Precisão e Relatório de Classificação de modelo para cenário aleatório

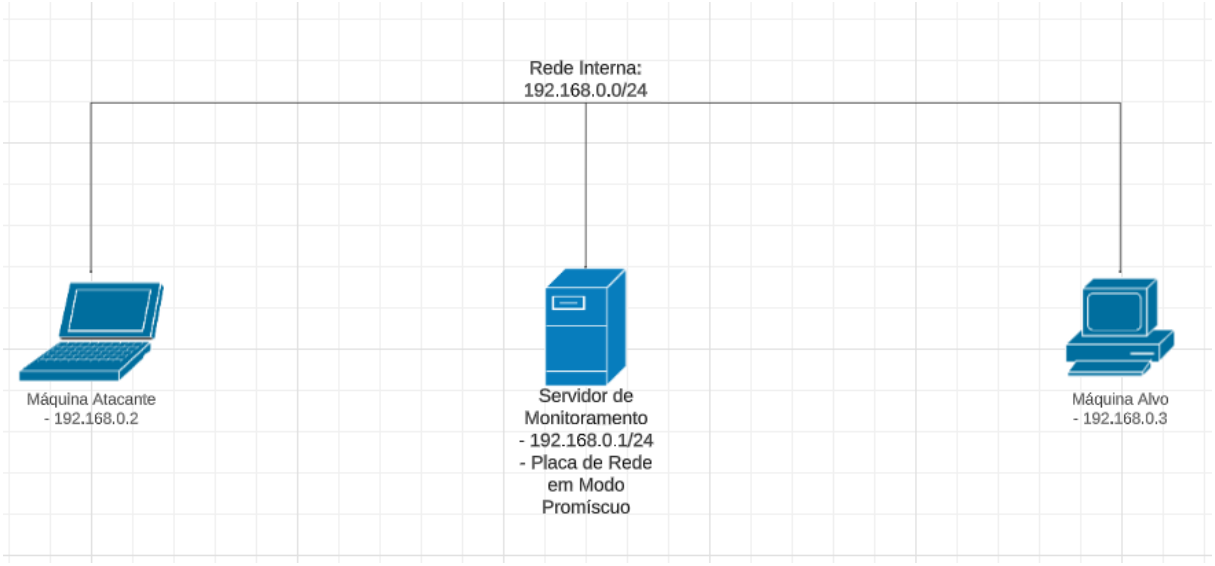
```
Accuracy: 0.8
Classification Report:
              precision    recall  f1-score   support

      Attack       0.50      1.00      0.67         2
      Normal       1.00      0.75      0.86         8

    accuracy                           0.80        10
   macro avg       0.75      0.88      0.76        10
weighted avg       0.90      0.80      0.82        10
```

Fonte: Autoria Própria, Google Colab Notebook

4.3. Eficácia do modelo treinado com pacotes ICMP normais e DOS pelo HPING3.

A partir dos resultados anteriormente obtidos é que foi notada a necessidade de uma massa de dados mais fidedigna a uma situação real enfrentada por uma rede de computadores e que então foi-se realizada a captura de pacotes compartilhados entre máquinas virtuais. A estrutura de rede interna entre essas máquinas foi simplificada apenas às máquinas relevantes e elaborada conforme diagrama apresentado na Figura 3:

Figura 3 – Diagrama de rede das máquinas virtuais utilizadas



Fonte: Autoria Própria

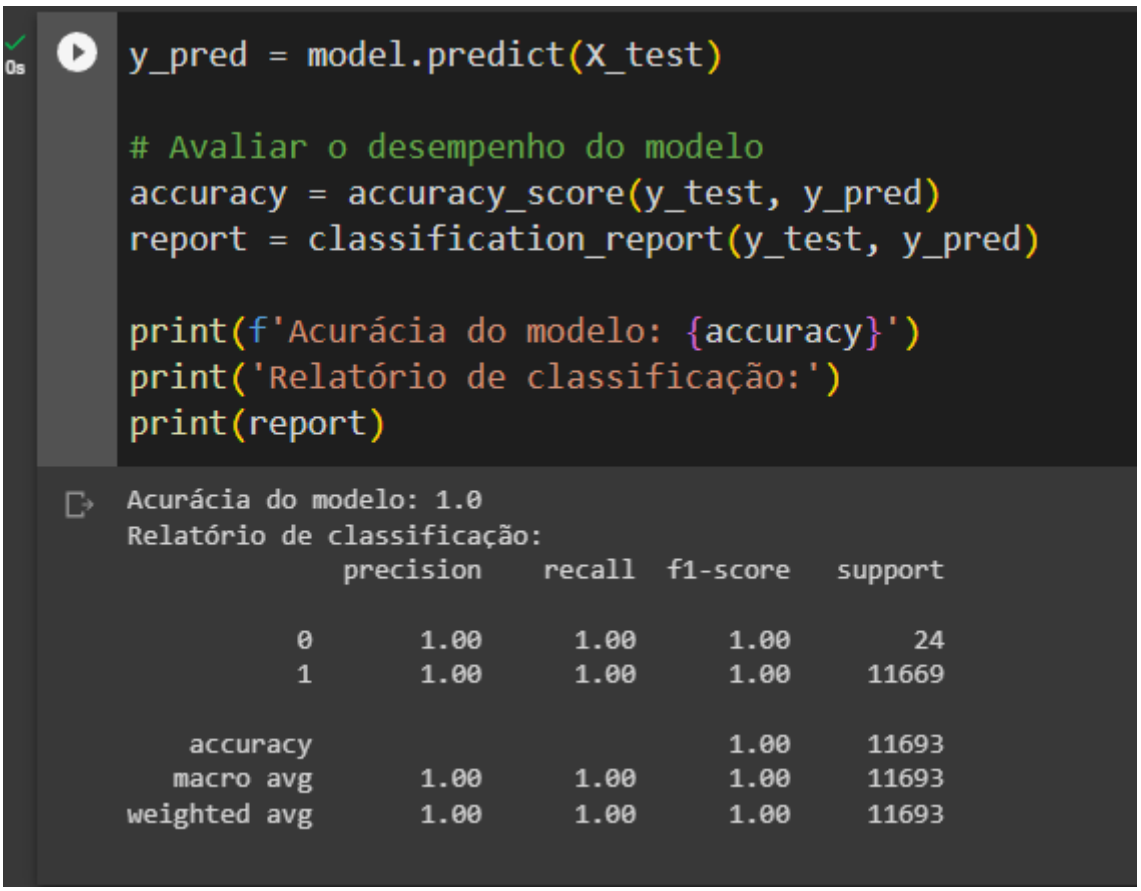
A “Máquina Atacante” realizou o envio dos pacotes ICMP à “Máquina Alvo” tanto comumente quanto de maneira similar a um ataque DoS (*Denial of Service*, Negação de Serviço). Todo tráfego foi capturado pela máquina “Servidor de Monitoramento” e utilizado para treinamento do modelo de ML transformando o rastreio dos pacotes em um *dataframe* (quadro de dados) como no teste anterior e classificando os *pings* comuns como “normal” e os realizados pelo *flood* (inundação de pacotes) do *hping3* como “*threat*” (ameaça).

Durante a execução do *flood* pela máquina atacante, um parâmetro foi utilizado para a utilização de endereços IP de origem dos pacotes serem aleatórios. Tal informação no cabeçalho dos pacotes transmitidos pode ser considerada como sendo de extrema utilidade para o treinamento do modelo uma vez que tal técnica de ofuscação do endereço é comumente utilizada em ataques reais. O problema neste caso é que um modelo treinado com base nas bibliotecas Python disponibilizadas pelo *Scikit-Learn* opera fazendo uso apenas de valores *float* (decimais). Porém, como o endereço IP é formado por 4 conjuntos numéricos separados por um ponto, tal informação é passada ao *Dataframe* como um valor textual.

Devido a este inconveniente, foi necessária a aplicação de uma técnica conhecida como *Hashing de Features*, uma técnica usada na engenharia de características em aprendizado de máquina. Ela envolve a conversão de características (como texto ou categorias) em valores numéricos usando funções de *hash*. Em um contexto de análise de dados, o *hashing* de *features* ajuda a representar dados categóricos, como palavras ou endereços IP, como números que podem ser usados como entradas para algoritmos de *machine learning*.

Após a realização da tratativa dos dados e o do treinamento do modelo, os resultados obtidos apresentaram métricas de desempenho notavelmente altas, indicando uma capacidade de previsão excepcionalmente precisa para as classes de interesse, que são rotuladas como "0" (Normal) e "1" (Ameaça). Conforme mostra o programa desenvolvido no Google Colab Notebook na Figura 4.

Figura 4 – Acurácia e Relatório do Modelo 02



```
y_pred = model.predict(X_test)

# Avaliar o desempenho do modelo
accuracy = accuracy_score(y_test, y_pred)
report = classification_report(y_test, y_pred)

print(f'Acurácia do modelo: {accuracy}')
print('Relatório de classificação:')
print(report)
```

```
Acurácia do modelo: 1.0
Relatório de classificação:
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        24
           1       1.00      1.00      1.00     11669

    accuracy                           1.00     11693
   macro avg       1.00      1.00      1.00     11693
weighted avg       1.00      1.00      1.00     11693
```

Fonte: Autoria Própria (Google Colab Notebook)

A acurácia do modelo é de 1.0, o que significa que o modelo classificou corretamente todas as amostras do conjunto de teste. Isso é um resultado excepcional e sugere que o modelo foi capaz de distinguir com precisão entre os padrões de tráfego de rede classificados como "Normal" e "Ameaça".

A métrica de precisão (precision) é de 1.00 para ambas as classes. Isso indica que o modelo não gerou falsos positivos; todas as previsões de "Ameaça" estavam corretas, bem como todas as previsões de "Normal".

O *recall* (revocação) também é de 1.00 para ambas as classes, indicando que o modelo identificou todos os casos reais de "Normal" e "Ameaça" presentes no conjunto de teste. O F1-Score, que é uma métrica que combina precisão e *recall*, também é de 1.00 para ambas as classes. Isso reflete um equilíbrio notável entre a capacidade do modelo de evitar falsos positivos e identificar todos os casos positivos reais.

Os resultados deste modelo demonstram uma capacidade impressionante de distinguir entre tráfego de rede normal e tráfego associado a ameaças de segurança. A acurácia de 100% indica que o modelo não cometeu erros de classificação no conjunto de teste. Essa alta precisão e *recall* são extremamente desejáveis em cenários de segurança da informação, onde a detecção precisa de ameaças é crucial.

No entanto, é importante mencionar que a análise desses resultados deve ser realizada considerando a possibilidade de desequilíbrio de classe, uma vez que a classe "Normal" tem um número muito menor de amostras em comparação com a classe "Ameaça". É importante entender as implicações do desequilíbrio de classe e considerar estratégias como subamostragem, sobreamostragem ou ajuste de limiar, dependendo das necessidades do cenário de aplicação.

5. Considerações Finais

Este artigo buscou explorar a aplicabilidade do *machine learning* no monitoramento de redes, com foco na segurança da informação. Ao longo do estudo, realizou-se o desenvolvimento e aplicação de *machine learning* para a detecção de ameaças em redes de computadores, usando, a partir da necessidade refletida, dados de tráfego capturados pelo TCPDump em uma rede de máquinas virtuais. A análise dos resultados obtidos permitiu tirar conclusões importantes sobre a viabilidade e eficácia dessa abordagem.

Os resultados do modelo final foram excepcionais, com uma acurácia de 100%. Isso significa que o modelo classificou com precisão todas as amostras do conjunto de teste, demonstrando uma capacidade notável de distinguir entre padrões de tráfego de rede classificados como "Normal" para pacotes ICMP comumente disseminados e "Ameaça" para pacotes utilizados como fonte de um ataque de negação de serviço. As métricas de precisão, *recall* e F1-*Score* também alcançaram 1.00 para ambas as classes, indicando um equilíbrio entre a capacidade do modelo de evitar falsos positivos e identificar todos os casos positivos reais. A partir da aplicabilidade de um modelo similar ao explorado, em um ambiente real, é possível otimizar o tempo de reação à ataques em andamento e auxiliar na tomada de decisão conforme dados obtidos.

Durante as pesquisas de abordagens para a aplicação do *machine learning* para o monitoramento de redes este se provou poder ser executado de diversas formas distintas. O treinamento de um modelo para a classificação de um ataque de negação de serviço é apenas uma das diversas opções existentes de aplicação. Outras opções podem envolver a análise do tráfego geral em uma rede classificando-o conforme protocolos e tamanho de pacotes, a previsão de falhas em dispositivos de rede, como roteadores e *switches*, com base em dados históricos de desempenho, permitindo uma manutenção proativa e a minimização de interrupções na rede, o auxílio na gestão de tráfego e *QoS (Quality of Service,* Qualidade de serviço*)*, etc.

Pormenor, este estudo contribui para o campo da segurança da informação, demonstrando que o uso de *machine learning* no monitoramento de redes pode ser altamente eficaz na detecção precoce de ameaças cibernéticas. As aplicações práticas desta pesquisa incluem aprimorar a segurança de redes empresariais, identificar atividades suspeitas e reduzir os riscos e impactos de ataques cibernéticos.

Referências

BELENTANI, L. C.; MARCELLO, J.; FLORIAN, F. A utilização de ferramentas de monitoramento para otimização do gerenciamento de rede. **Revista Interface Tecnológica**, [S. l.], v. 15, n. 2, p. 99–110, 2018. DOI: 10.31510/infa.v15i2.509. Disponível em:

https://revista.fatectq.edu.br/index.php/interfacetecnologica/article/view/509. Acesso em: 12 fev. 2023.

BLACK, T. **Comparação de Ferramentas de Gerenciamento de Redes**. UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL INSTITUTO DE INFORMÁTICA CURSO DE ESPECIALIZAÇÃO EM TECNOLOGIAS, GERÊNCIA E SEGURANÇA DE REDES DE COMPUTADORES. Disponível em: http://hdl.handle.net/10183/15986. Acesso em: 25 fev .2023

CHUI, M. et al. **Notes from the AI frontier insights from hundreds of use cases** [s.l: s.n.]. Disponível em: <https://www.mckinsey.com/~/media/mckinsey/featured%20insights/artificial%20intelligence/notes%20from%20the%20ai%20frontier%20applications%20and%20value%20of%20deep%20learning/notes-from-the-ai-frontier-insights-from-hundreds-of-use-cases-discussion-paper.ashx>. Acesso em: 01 maio. 2023

LEE, S.; LEVANTI, K.; KIM, H. S. Network monitoring: Present and future. **Computer Networks**, v. 65, p. 84–98, jun. 2014. Disponível em: https://www.sciencedirect.com/science/article/abs/pii/S138912861400111X. Acesso em: 29 maio. 2023

LIMA, J. DOS R. **Monitoramento de redes com Zabbix:** Monitore a saúde dos servidores e equipamentos de redes. [s.l.] Brasport, 2014. Disponível em: https://books.google.com.br/books?hl=pt-BR&lr=&id=G3McAwAAQBAJ&oi=fnd&pg=PA1&dq=LIMA,+J.+DOS+R.+Monitoramento+de+Redes+com+Zabbix:+Monitore+a+sa%C3%BAde+dos+servidores+e+equipamentos+de+redes.+%5Bs.l.%5D+Brasport,+2014.&ots=epvATx0JSO&sig=RVSM1fzt3sJQ0Vn3. Acesso em: 14 jun. 2023

RAMOS, JOHNY DE SOUZA; NASCIMENTO, LUCAS GUILHERME DE SOUZA BORGES; BISCHOFF, ROBERTO ANDREI. **Ferramenta para monitoramento de dispositivos de rede implementada utilizando a linguagem de programação**. 2022. Trabalho de conclusão de curso (Tecnólogo em Sistemas de Telecomunicações) - Universidade Tecnológica Federal do Paraná (UTFPR), [*S. l.*], 2022. Disponível em: http://repositorio.utfpr.edu.br/jspui/handle/1/29729. Acesso em: 14 jun.2023

ZHOU, Z.-H. **Machine learning.** Gateway East, Singapore: Springer, 2021. Disponível em: https://books.google.com.br/books?hl=pt-BR&lr=&id=ctM-EAAAQBAJ&oi=fnd&pg=PR6&dq=ZHOU,+Z.-H.+Machine+learning.+Gateway+East,+Singapore:+Springer,+2021.&ots=oZRlZ8Ry3t&sig=n4vzgUsJINHf81r-URDO8FCGaE0#v=onepage&q&f=false. Acesso em: 29 jul. 2023