Utilização e avaliação de aprendizado de máquina para detecção de ataques DDoS*

Using and evaluating machine learning to detect DDoS attacks

Leonardo Ravaiani da Silva[1]
Felipe Augusto Lima Reis (Orientador)[2]

Resumo

Os Ataques Distribuídos de Negação de Serviço (DDoS - *Distributed Denial of Service*) são uma ameaça persistente e significativa para a segurança cibernética. Este artigo propõe investigar a detecção de ataques DDoS em redes, utilizando o algoritmo Random Forest. A metodologia inclui a análise do tráfego de rede da base de dados *DDoS Evaluation Dataset (CIC-DDoS2019)*, que contém dados de DDoS de diferentes tipos, treinamento do algoritmo Random Forest, e avaliação de seu desempenho, incluindo sua velocidade e tempo de execução. Espera-se que os resultados forneçam *insights* sobre a eficácia e eficiência do algoritmo Random Forest na detecção de ataques DDoS, contribuindo assim para o fortalecimento da segurança cibernética das redes.

Palavras-chave: Ataques de Negação de Seviço, DoS, DDoS, Aprendizado de Máquinas, Inteligencia Artificial, Random Forest.

*Trabalho de conclusão de curso, Sistemas de Informação, Unidade São Gabriel

[1]Programa de Graduação em Sistema da informação da PUC Minas, Brasil– leonardo.rav@hotmail.com
[2]Instituto de Ciências Exatas e de Informática da PUC Minas, Brasil– felipereis@pucminas.br

Abstract

Distributed Denial of Service Attacks (DDoS) are a persistent and significant threat to cybersecurity. This article proposes to investigate the detection of DDoS attacks in networks, using the Random Forest algorithm. The methodology includes analyzing network traffic from the DDoS Evaluation Dataset (CIC-DDoS2019) database, which contains DDoS data of different types, training the Random Forest algorithm, and evaluating its performance, including its speed and execution time. . The results are expected to provide insights into the effectiveness and efficiency of the Random Forest algorithm in detecting DDoS attacks, thus contributing to strengthening the cybersecurity of networks.

Keywords: Denial of Service Attacks, DoS, DDoS, Machine Learning, Artificial Intelligence, Random Forest.

1 INTRODUÇÃO

Nos últimos anos, a crescente dependência da sociedade em redes de computadores tem sido acompanhada por um aumento significativo no número e na sofisticação dos ataques cibernéticos. Entre esses ataques, os chamados Ataques Distribuídos de Negação de Serviço (DDoS) emergiram como uma das formas mais prejudiciais de comprometimento da infraestrutura de rede, representando uma séria ameaça à disponibilidade e integridade dos serviços online. Os ataques DDoS se caracterizam pela sobrecarga de tráfego malicioso de um website ou servidor, resultando em congestionamento ou inacessibilidade do sistema, negando serviços aos usuários e obstruindo a chegada do tráfego legítimo ao seu destino (DAYAL et al., 2016).

Quando ocorre em larga escala, esse tipo de ameaça pode resultar na desaceleração dos servidores das empresas e de seus respectivos serviços, ocasionando perdas financeiras. Além disso, devido às novas ferramentas e tecnologias empregadas por hackers para realizar ataques DDoS, identificar a fonte ou origem do invasor torna-se desafiador, já que estes chegam de múltiplas fontes (DALVI et al., 2021).

Detectar e combater ataques DDoS é crucial para manter a segurança das redes e consequentemente, dos usuários. Com a propagação desses ataques, é necessário usar estratégias inovadoras, como algoritmos de Aprendizado de Máquina (*Machine Learning - ML*), para obter uma resposta rápida e eficaz. Proteger sistemas e a privacidade dos usuários é essencial para um ambiente online seguro.

A elaboração deste artigo se justifica diante do desafio enfrentado na segurança cibernética. Com a constante evolução das táticas dos agressores e o aumento exponencial na frequência e severidade dos ataques (PERAKOVIĆ et al., 2015), torna-se cada vez mais evidente a necessidade de soluções inovadoras e adaptativas para proteger as infraestruturas de rede. A aplicação de modelos de ML oferece uma abordagem promissora para identificar padrões de tráfego malicioso e distinguir entre atividades legítimas e ataques. Além disso, a capacidade dos algoritmos de ML de aprender com dados históricos e se ajustar a novos padrões de ataque aumenta a eficácia e a agilidade das defesas cibernéticas.

Este trabalho tem como objetivo **desenvolver e avaliar modelos de aprendizado de máquina capazes de distinguir entre o tráfego benigno e os ataques DDoS, contribuindo assim para a proteção proativa das redes contra essa forma de ameaça.** Pretende-se analisar como a aplicação de técnicas de ML pode contribuir para a detecção precisa de ataques DDoS, aprimorando a segurança cibernética das redes.

Foram definidos os seguintes objetivos específicos: (i) investigar a eficácia da aplicação de algoritmos de aprendizado de máquina na detecção precisa de ataques DDoS, avaliando sua capacidade de distinguir entre tráfego benigno e malicioso; (ii) treinar o algoritmo Random Forest para análise de tráfego de rede; e, (iii) avaliar o desempenho e a eficiência dos algoritmos, considerando métricas como precisão, revocação (*recall*) e tempo de execução.

Este trabalho está estruturado em 5 seções, para melhor organização. Na Seção 2, é fornecido o referencial teórico, com explicação de conceitos fundamentais para compreensão

do trabalho. A Seção 3, explora trabalhos e sistemas que fundamentaram o desenvolvimento do projeto. Na Seção 4, descreve detalhadamente o método utilizado. Os resultados esperados e parciais são apresentados na Seção 5. A Seção 6 contém o cronograma com as atividades previstas para a confecção deste trabalho.

2 REFERENCIAL TEÓRICO

Esta seção pretende apresentar os principais tópicos acerca de segurança cibernética e inteligência artificial.

2.1 Ataques de Negação de Serviço Distribuído (DDoS)

Um Ataque de Negação de Serviço (DoS) ocorre quando um invasor tenta, maliciosamente, sobrecarregar um serviço ou recursos de rede usando apenas uma única fonte, tornando-os inacessíveis para usuários legítimos. Por outro lado, quando esse ataque é realizado a partir de múltiplas fontes distribuídas, fica conhecido como Ataque Distribuído de Negação de Serviço (DDoS) (HUSSAIN et al., 2020).

Os ataques DDoS, de acordo com Luo e Han (2019), consistem em enviar uma grande quantidade de pacotes de dados específicos para consumir largura de banda ou recursos do sistema da rede-alvo, levando ao bloqueio ou paralisação da mesma. Este tipo de ataque é realizado por múltiplos nós de origem, em vez de apenas um ou alguns, e é baseado em colaboração distribuída. Os atacantes utilizam centenas de nós controlados para lançar ataques cooperativos em larga escala, degradando ou até mesmo paralisando o desempenho do alvo atacado. Esses ataques afetam uma ampla gama de usuários legítimos e são difíceis de rastrear e prevenir, devido à sua implementação fácil e complexidade de detecção.

2.2 Random Forest

O algoritmo Random Forest é uma técnica de classificação que opera sob supervisão, ou seja, requer dados de treinamento rotulados. Ele consiste em múltiplos pontos de decisão e usa valores aleatórios para criar uma abordagem robusta. Este algoritmo é amplamente reconhecido e utilizado no campo de aprendizado de máquina, especialmente em contextos de aprendizado supervisionado. É particularmente útil para lidar com problemas complexos e para melhorar a precisão das previsões. O Random Forest é composto por várias árvores de decisão que empregam um método de votação majoritária para determinar a previsão final, o que geralmente resulta em um desempenho superior (A; DHARMARAJAN, 2022).

2.3 Métricas de Avaliação em Aprendizado de Máquina

Alguns dos critérios mais utilizados para avaliar o desempenho de algoritmos de machine learning incluem a acurácia, que avalia a proporção de classificações corretas feitas pelo modelo, a precisão, que indica a proporção de classificações corretas da classe positiva, e o *recall* (sensibilidade), que representa a eficácia do modelo em prever a classe positiva. O F1-score, uma medida harmoniosa entre precisão e *recall*, é aplicado em conjuntos de dados desbalanceados. A métrica Área Sob a Curva ROC (AUC - *Area Under The Curve*) é utilizada para avaliar a capacidade de classificação do modelo em conjuntos desbalanceados, onde valores próximos a 1 indicam maior precisão. Por fim, a Taxa de Aceitação Falsa (FAR - *False Acceptance Rate*) reflete a taxa de classificações incorretas de conexões normais, sendo valores abaixo de 10% considerados promissores para o modelo (SOUSA; SILVA, 2022).

3 TRABALHOS RELACIONADOS

Dias et al. (2023) apresenta a integração de microsserviços em um ambiente Kubernetes, combinada com técnicas de aprendizado de máquina para otimizar a detecção de ataques DDoS em Redes Definidas por Software (SDN - *Software Defined Networking*). A arquitetura proposta envolve a utilização de modelos de ML, como o Random Forest, para analisar o tráfego de rede e identificar padrões maliciosos que indicam a presença de ataques DDoS. Essa abordagem permite a detecção eficaz dos ataques, enquanto a implementação em microsserviços no Kubernetes proporciona melhor isolamento de recursos e desempenho para cada componente do sistema. Os resultados obtidos indicam a eficácia da abordagem e sugerem a continuação de estudos na área de segurança de redes SDN em nuvem, incluindo a implementação de outros modelos de classificação e a comparação com outras abordagens.

Naing e Thwel (2023) investiga a eficácia de diversos algoritmos de aprendizado de máquina na classificação de ataques de negação de serviço distribuídos (DDoS). Conduzido por pesquisadores do Laboratório de Pesquisa em Segurança Cibernética da Universidade de Estudos de Computação em Yangon, Myanmar, a pesquisa aponta que o classificador de Regressão Logística dentre os estudados, que incluem apenas técnicas básicas, se destaca como o mais eficaz na identificação de ataques DDoS. Utilizando um conjunto de dados abertos de ataques DDoS, os pesquisadores compararam o desempenho de algoritmos como Stochastic Gradient Classifier, Support Vector Machine (SVM), k-Nearest Neighbor (kNN), Naïve Bayes e Regressão Logística. O estudo ressalta a importância de aprimorar métodos de proteção contra ataques DDoS e destaca a necessidade de identificar e classificar esses ataques de forma eficiente para assegurar a integridade e confidencialidade dos sistemas de computadores e redes.

Santos (2023) aborda a crescente ameaça dos ataques de negação de serviço distribuído (DDoS) em redes de computadores, propondo o uso do algoritmo Random Forest para diferenciar entre tráfego benigno e malicioso em ataques do tipo UDP. O trabalho destaca a aplicação

do protocolo NetFlow para coletar e armazenar dados, fornecendo uma base essencial para a análise do comportamento do tráfego de rede. Os resultados revelaram uma precisão de 83% na classificação de dados benignos mesmo em cenários de ataques DDoS UDP. A metodologia adotada foi dividida em três etapas: análise exploratória, treinamento do algoritmo e avaliação da acurácia na base de teste, proporcionando uma estrutura sólida para o estudo. Assim, este trabalho contribui não apenas para a compreensão e melhoria da detecção de anomalias no tráfego de rede, mas também promove avanços significativos na segurança cibernética.

Nitze et al. (2012) analisa e compara o desempenho de diferentes algoritmos de aprendizado de máquina na classificação de tipos de culturas usando dados de sensoriamento remoto. A medida de tempo de execução do algoritmo Random Forest foi realizada durante o treinamento e teste do modelo de classificação. Os tempos de execução foram registrados em segundos por classificação. Para o Random Forest, os tempos de treinamento e teste foram calculados em média para diferentes cenários, como o número de imagens utilizadas no estudo. Por exemplo, o tempo médio de treinamento por classificação para o Random Forest variou de 1.1 a 6.2 segundos, dependendo do número de aquisições de imagens. Já o tempo médio de teste por classificação para o Random Forest foi registrado em média entre 0.011 e 0.175 segundos, novamente dependendo do número de imagens utilizadas no estudo.

Hermawan et al. (2021) apresenta uma análise comparativa de três algoritmos de classificação de árvores de decisão: Algoritmo J48 (implementação do Algoritmo C4.5), Decision Tree e Random Forest, aplicados a dados de recomendação de cupons em compras de veículos. No estudo, o tempo de construção do modelo foi registrado para cada algoritmo e comparado entre eles. Esses tempos de execução foram medidos em segundos e fornecem uma indicação do desempenho computacional de cada algoritmo ao lidar com o conjunto de dados de recomendação de cupons em veículos.

4 MÉTODO

A parte de análise do algoritmo, neste caso o Random Forest, envolverá algumas etapas. Inicialmente, serão coletados e preparados os dados de tráfego de rede, incluindo amostras de tráfego benigno e de ataques DDoS. Em seguida, o algoritmo Random Forest será treinado utilizando os dados preparados. Após o treinamento, será gerado o gráfico de importância das features, que mostrará quais as variáveis do conjunto de dados são mais relevantes para detectar se um fluxo é caracterizado como um ataque DDoS ou não. Por último, o modelo será avaliado. Durante a avaliação, serão registradas métricas como precisão, recall, e F1-score.

A avaliação do desempenho do algoritmo Random Forest também incluirá testes específicos para mensurar sua velocidade e tempo de execução para fazer previsões acerca de novas entradas de dados. Para isso, serão realizados experimentos cronometrados onde o algoritmo será submetido a uma carga de trabalho simulando condições verdadeiras de tráfego de rede. Serão registrados o tempo necessário para que o algoritmo processe e classifique um conjunto

específico de dados de tráfego.

A base de dados utilizada para os primeiros testes foi a *DDoS Evaluation Dataset (CIC-DDoS2019)* que inclui vários dados acerca de diferentes tipos de ataques DDoS, mas a principio foi utilizado apenas os ataques do tipo UDP, caracterizados por uma inundação de pacotes enviados a um servidor, visando sobrecarregar sua capacidade de processamento e resposta. Esse ataque foi executado durante o período de captura de dados, que começou às 12:36 UTC-4 e terminou às 13:04 UTC-4. O texto destaca a importância de construir modelos de detecção para capturar os padrões desses ataques, utilizando algoritmos de aprendizado de máquina comuns, como o random forest e o logistic regression (SHARAFALDIN et al., 2019).

5 RESULTADOS

5.1 Resultados Esperados

Os resultados esperados com esse estudo são centrados na avaliação do modelo de aprendizado de máquina Random Forest para segurança dos usuários, visando distinguir o tráfego malicioso do benigno. O artigo será uma continuação do Trabalho de Conclusão de Curso feito por Santos (2023), ex-aluno do curso de Sistemas de Informação da PUC Minas, e será focado em analisar o algoritmo proposto acerca de seu desempenho, evidenciando sua precisão, recall e tempo de execução. Espera-se também uma análise mais precisa e eficiente do tráfego de rede, proporcionando uma detecção mais rápida e confiável de ataques DDoS. Espera-se que os resultados obtidos forneçam uma compreensão aprofundada do potencial e das limitações desse algoritmo específico para a detecção de ataques DDoS em cenários reais de rede. Por fim, espera-se que os *insights* obtidos contribuam para aprimorar a eficácia das defesas cibernéticas contra ameaças DDoS, fornecendo assim uma base sólida para futuras pesquisas e desenvolvimentos no campo da segurança cibernética.

5.1.1 Planejamento dos experimentos

Para observar o comportamento do algoritmo Random Forest, serão realizados testes variando as características de cada modelo a ser treinado, como o número de árvores ou a profundidade das árvores geradas. É possível também que ao invés de usar o *dataset* completo, seja usado apenas uma parte dele para verificar se há mudanças nos resultados obtidos. Nos 2 primeiros experimentos realizados, o treinamento foi realizado utilizando a biblioteca Scikit-learn, separando a base de dados em 75% para treinamento e 25% para testes, e o algoritmo foi executado no ambiente do Jupyter Notebook. Nos próximos testes será utilizado um ambiente mais controlado a fim de obter maior precisão nos resultados.

Para avaliar o desempenho do algoritmo Random Forest, algumas métricas serão uti-

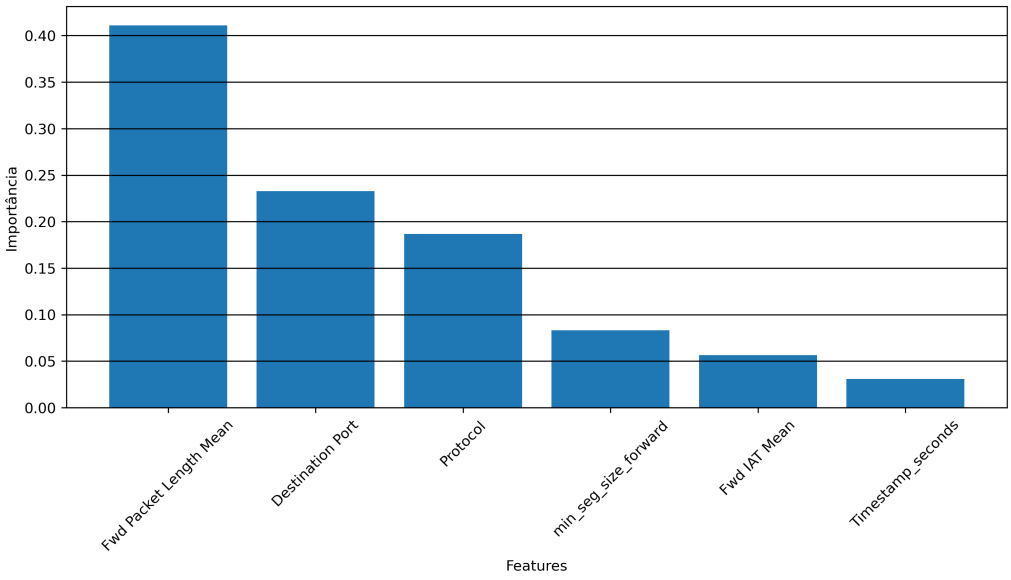
lizadas. Isso inclui precisão, recall, F1-score e acurácia, que fornecem uma visão abrangente da capacidade do modelo de identificar corretamente tanto os ataques DDoS quanto os eventos normais de tráfego de rede. Além disso, será considerado o tempo necessário para o modelo fazer previsões sobre novos dados. Esse aspecto é importante, especialmente em cenários de segurança cibernética, onde a detecção e resposta rápida são essenciais.

5.2 Resultados Parciais

5.2.1 Importância das features

A análise da importância das *features* no modelo Random Forest revelou informações sobre os padrões presentes nos dados de detecção de ataques DDoS. A *feature* mais significativa é o *"Fwd Packet Length Mean"*, com uma importância superior a 0,40. Isso indica que o tamanho médio dos pacotes encaminhados tem uma importância de mais de 40%, e é um forte indicador na diferenciação entre ataques DDoS e tráfego normal. Além disso, a *"Destination Port"*, com uma importância entre 0,20 e 0,25, e o *"Protocol"*, com uma importância entre 0,15 e 0,20, também desempenham papéis importantes na classificação dos dados. *Features* como 'min_seg_size_forward' e *"Fwd IAT Mean"*, com importâncias entre 0,05 e 0,10, contribuem significativamente, mas em menor grau, para o modelo na identificação de padrões associados a ataques DDoS. Por fim, a *feature "Timestamp_seconds"*, embora tenha a menor importância, entre 0,00 e 0,05, fornece informações adicionais sobre o tempo em segundos desde a época Unix, que, embora menos relevante, ainda contribui para o entendimento geral dos padrões de tráfego.

Figura 1 – Gráfico de importância das *features*



Fonte: Elaborado pelo autor.

5.2.2 Experimento 1: Avaliação do algoritmo Random Forest com 100 árvores

O primeiro experimento foi feito treinando o modelo com 100 árvores e utilizando toda a base de dados. Os resultados obtidos foram satisfatórios, com uma precisão de 97% para fluxo benigno, indicando que 97% das instâncias classificadas como benignas realmente são benignas, e 100% para ataques DDoS. O *recall*, que mede a proporção de instâncias positivas que foram corretamente identificadas pelo classificador, indicou que 96% das instâncias de ataques benignos foram corretamente identificadas. Já o F1-score, que é a média harmônica entre precisão e *recall*, evidenciou 96% para o fluxo benigno, indicando um bom equilíbrio.

	Amostras	**Precisão**	**Recall**	**F1-score**
Benign	531	97%	96%	96%
DrDos_UDP	783670	100%	100%	100%

Tabela 1 – Métricas de desempenho do Random Forest com 100 árvores

A acurácia de 99% mostra que o classificador previu corretamente 99% das amostras totais. Quanto ao tempo de execução para o algoritmo fazer previsões acerca de novas amostras, foi cronometrado 2,54 segundos, que corresponde ao tempo total de predição do algoritmo.

Acurácia	**Tempo de execução (s)**
99%	2,54

Tabela 2 – Acurácia e Tempo de Predição do modelo treinado com 100 árvores

5.2.3 Experimento 2: Avaliação do algoritmo Random Forest com 15 árvores

O segundo experimento foi feito treinando o modelo com 15 árvores e utilizando também toda a base de dados. Os resultados mostram que mesmo com apenas 15 árvores, o modelo se comportou bem e se assemelhou com os números das métricas apresentados no primeiro experimento.

	Amostras	**Precisão**	**Recall**	**F1-score**
Benign	531	98%	95%	97%
DrDos_UDP	783670	100%	100%	100%

Tabela 3 – Métricas de desempenho do Random Forest com 15 árvores

Quanto a acurácia e tempo de execução total para predição, o modelo com 15 árvores foi 2,11 segundos mais rápido que o modelo treinado com 100 árvores e manteve a acurácia em 99%, como é mostrado a seguir na Tabela 4.

Acurácia	Tempo de execução (s)
99%	0,43

Tabela 4 – Acurácia e Tempo de Predição do modelo treinado com 15 árvores

Os resultados dos 2 experimentos apontam que o modelo treinado com menos árvores consegue ter métricas de desempenho próximas ou até melhores que o modelo treinado com mais árvores, gastando bem menos tempo para fazer a previsão acerca de novos dados.

6 CRONOGRAMA

Para desenvolvimento do projeto, foi definido o cronograma disponível na Tabela 5. As atividades foram delineadas considerando as etapas necessárias para a construção do projeto, e as datas de entrega foram estabelecidas conforme os requisitos e cronograma de atividades da disciplina de Trabalho de Conclusão de Curso.

Tabela 5 – Cronograma de atividades previstas.

Atividades	Fevereiro	Março	Abril	Maio	Junho
1. Definição do tema	×				
2. Levantamento bibliográfico		×	×		
3. Análise de algoritmos de ML		×	×		
4. Testes e experimentos			×	×	×
5. Melhorias e correções				×	×
6. Preparação para defesa					×
7. Redação do TCC		×	×	×	×

Referências

A, U. N.; DHARMARAJAN, K. Diabetes prediction using random forest classifier with different wrapper methods. In: *2022 International Conference on Edge Computing and Applications (ICECAA)*. [S.l.: s.n.], 2022. p. 1705–1710.

DALVI, Jai et al. Ddos attack detection using artificial neural network. In: **2021 International Conference on Industrial Electronics Research and Applications (ICIERA)**. [S.l.: s.n.], 2021. p. 1–5.

DAYAL, Neelam et al. Research trends in security and ddos in sdn. **Security and Communication Networks**, v. 9, n. 18, p. 6386–6411, 2016. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sec.1759>.

DIAS, Victor et al. Detecção de ataques ddos em redes sdn utilizando aprendizado de máquina: Uma abordagem em microsserviços. In: **Anais Estendidos do XXIII Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais**. Porto Alegre, RS, Brasil: SBC, 2023. p. 141–152. ISSN 0000-0000. Disponível em: <https://sol.sbc.org.br/index.php/sbseg_estendido/article/view/27287>.

HERMAWAN, Dicky Rahma et al. Comparative study of j48 decision tree classification algorithm, random tree, and random forest on in-vehicle couponrecommendation data. In: **2021 International Conference on Artificial Intelligence and Big Data Analytics**. [S.l.: s.n.], 2021. p. 1–6.

HUSSAIN, Faisal et al. Iot dos and ddos attack detection using resnet. In: **2020 IEEE 23rd International Multitopic Conference (INMIC)**. [S.l.: s.n.], 2020. p. 1–6.

LUO, Wenliang; HAN, Wenzhi. Ddos defense strategy in software definition networks. In: **2019 International Conference on Computer Network, Electronic and Automation (ICCNEA)**. [S.l.: s.n.], 2019. p. 186–190.

NAING, Soe Kalayar; THWEL, Tin Thein. A study of ddos attack classification using machine learning classifiers. In: **2023 IEEE Conference on Computer Applications (ICCA)**. [S.l.: s.n.], 2023. p. 108–112.

NITZE, I; SCHULTHESS, U; ASCHE, H. Comparison of machine learning algorithms random forest, artificial neural network and support vector machine to maximum likelihood for supervised crop type classification. **Proceedings of the 4th GEOBIA, Rio de Janeiro, Brazil**, v. 79, p. 3540, 2012.

PERAKOVIĆ, Dragan; PERIŠA, Marko; CVITIĆ, Ivan. Analysis of the iot impact on volume of ddos attacks. **XXXIII Simpozijum o novim tehnologijama u poštanskom i telekomunikacionom saobraćaju–PosTel**, v. 2015, p. 295–304, 2015.

SANTOS, Davi Jorge Leite. **Aplicação do algoritmo Random Forest para classificação de tráfego de rede benigno em meio a ataques DDoS UDP**. 2023. Tese (Trabalho de conclusão de curso) — Pontifícia Universidade Católica de Minas Gerais, PUC Minas São Gabriel, disponível em <http://bib.pucminas.br:8080/pergamumweb/vinculos/0000ba/0000ba97.pdf>.

SHARAFALDIN, Iman et al. Developing realistic distributed denial of service (ddos) attack dataset and taxonomy. In: **2019 International Carnahan Conference on Security Technology (ICCST)**. [S.l.: s.n.], 2019. p. 1–8.

SOUSA, Welton Thiago Martins; SILVA, Carlos Alexandre. Análise de desempenho em algoritmos de aprendizagem de máquina na detecção de intrusão baseada em fluxo de rede usando o conjunto de dados unsw-nb15. **Revista de Sistemas e Computação-RSC**, v. 12, n. 2, 2022.