Análise Comparativa de Random Forest e Multi-Layer Perceptron para Detecção de Intrusões em Redes com Redução de Dimensionalidade no CICIDS2017

Pedro H.D. Silva[1], **Paulo R.S.L. Coelho**[1], **Rafael Pasquini**[1], **Márcia A. Fernandes**[1], **Luís F. Faina**[1]

[1]Faculdade de Computação – Universidade Federal de Uberlândia (UFU)
Caixa Postal 592 – 38.400-902 – Uberlândia – MG – Brazil

{pedrodomeneghi,paulocoelho,rafael.pasquini,marcia,luisfaina}@ufu.br

Abstract. *The increasing sophistication of cyberattacks has driven the use of Machine Learning for network intrusion detection. This work presents a comparative analysis of Random Forest (RF) and Multi-Layer Perceptron (MLP) using the CICIDS2017 dataset, considering dimensionality reduction techniques. Mean Decrease in Impurity (MDI) and Principal Component Analysis (PCA) were applied along with preprocessing and Stratified K-Fold Cross-Validation. Both models achieved F1-scores above 0.99, with Random Forest showing better computational efficiency. The results highlight the potential of Machine Learning for intrusion detection in real-world scenarios.*

Resumo. *A crescente sofisticação dos ataques cibernéticos tem impulsionado o uso de Aprendizado de Máquina para detecção de intrusões em redes. Este trabalho apresenta uma análise comparativa entre Random Forest (RF) e Multi-Layer Perceptron (MLP) utilizando o conjunto de dados CICIDS2017, considerando técnicas de redução de dimensionalidade. Foram aplicadas Mean Decrease in Impurity (MDI) e Principal Component Analysis (PCA), aliadas a pré-processamento e validação por Stratified K-Fold Cross-Validation. Ambos os modelos alcançaram F1-Score superior a 0,99, sendo o Random Forest mais eficiente computacionalmente. Os resultados reforçam o potencial do uso de ML na detecção de intrusões em cenários reais.*

1. Introdução

A crescente complexidade das infraestruturas de rede e o aumento expressivo de ataques cibernéticos têm intensificado a necessidade de mecanismos eficazes de segurança. Nesse contexto, os Sistemas de Detecção de Intrusão em Redes (NIDS) são fundamentais na identificação de atividades maliciosas.

Tradicionalmente, os NIDS são classificados em abordagens baseadas em assinaturas e em anomalias. Enquanto sistemas baseados em assinaturas são eficazes na detecção de ataques conhecidos, abordagens baseadas em anomalias permitem identificar comportamentos desconhecidos ao modelar padrões normais de tráfego [Labonne 2020].

Com o avanço do Aprendizado de Máquina, tornou-se possível identificar padrões complexos em grandes volumes de dados. Estudos recentes demonstram que a aplicação

de modelos supervisionados pode alcançar alta precisão na detecção de ataques, especialmente quando combinados com conjuntos de dados representativos, como o CICIDS2017 [Neto and Gomes 2019, Sharafaldin et al. 2018].

Neste trabalho, adota-se uma abordagem supervisionada de classificação, na qual modelos de Aprendizado de Máquina são treinados para identificar diferentes classes de tráfego, incluindo tipos específicos de ataques e tráfego benigno, com base em dados previamente rotulados. São investigados os modelos Random Forest (RF) e Multi-Layer Perceptron (MLP), bem como técnicas de redução de dimensionalidade, com o objetivo de avaliar o impacto dessas abordagens no desempenho dos modelos [ref. omitida].

O restante do artigo está organizado da seguinte forma. A Seção 2 descreve a metodologia, a Seção 3 apresenta os resultados e a Seção 4 traz as conclusões.

2. Metodologia

Esta seção descreve a metodologia adotada para a análise comparativa dos modelos de Aprendizado de Máquina aplicados à detecção de intrusões em redes. A metodologia foi estruturada em etapas que incluem seleção do conjunto de dados, pré-processamento, redução de dimensionalidade, treinamento dos modelos e avaliação de desempenho.

2.1. Conjunto de Dados

O conjunto de dados utilizado foi o CICIDS2017 [Sharafaldin et al. 2018], amplamente empregado na literatura para avaliação de Sistemas de Detecção de Intrusão em Redes. Esse *dataset* foi escolhido por sua representatividade de cenários reais, contendo tráfego benigno e diversas categorias de ataques, como DoS (*Deny-of-Service*), DDoS (*Distributed Deny-of-Service*), *PortScan* e ataques Web.

Neste trabalho, o problema é formulado como uma tarefa de classificação supervisionada multiclasse. As 78 características extraídas dos fluxos de rede são utilizadas como variáveis de entrada, enquanto o rótulo (*label*) do *dataset* é utilizado como variável alvo, representando as diferentes classes de tráfego, incluindo tipos específicos de ataques e tráfego benigno.

2.2. Pré-processamento dos Dados

O pré-processamento contempla as etapas: **limpeza dos dados** - remoção de valores ausentes, infinitos e duplicados; **remoção de atributos irrelevantes** - exclusão de características com baixa variabilidade ou valores constantes; **normalização** - uso da técnica *StandardScaler*, que transforma os dados para média zero e desvio padrão unitário.

Essas etapas contribuem para reduzir ruídos e garantir que todas as características tenham influência equilibrada no processo de aprendizado.

2.3. Redução de Dimensionalidade

Para reduzir a complexidade do modelo e melhorar a eficiência computacional, foram aplicadas duas técnicas de redução de dimensionalidade: **Mean Decrease in Impurity (MDI)** - técnica baseada em Random Forest que avalia a importância das características com base na redução de impureza [Breiman 2001]; **Principal Component Analysis (PCA)** - técnica estatística que projeta os dados em um espaço de menor dimensão, preservando a maior parte da variância [Pearson 2010].

A PCA foi adotada como abordagem principal, pois proporcionou maior redução de dimensionalidade sem dependência direta de um modelo específico.

2.4. Modelos de Aprendizado de Máquina

Foram avaliados dois modelos supervisionados amplamente utilizados na literatura: **Random Forest - (RF)** - modelo baseado em *ensemble learning*, composto por múltiplas árvores de decisão, conhecido por sua robustez e capacidade de lidar com dados desbalanceados [Breiman 2001]; **Multi-Layer Perceptron (MLP)** - rede neural artificial capaz de modelar relações não lineares e capturar padrões complexos nos dados [Rosay et al. 2020].

Os modelos foram configurados com parâmetros ajustados empiricamente, buscando equilíbrio entre desempenho e custo computacional.

2.5. Validação dos Modelos

Para avaliação robusta dos modelos, foi utilizada a técnica de validação cruzada *Stratified K-Fold Cross-Validation* [Kohavi et al. 1995], que preserva a proporção das classes em cada partição do *dataset*.

Essa abordagem é especialmente importante em cenários com classes desbalanceadas, garantindo maior confiabilidade nos resultados obtidos.

2.6. Métricas de Avaliação

O desempenho dos modelos foi avaliado utilizando métricas amplamente consolidadas: Precisão (*Precision*); Revocação (*Recall*); e F1-Score.

Foram utilizadas versões ponderadas dessas métricas, considerando o desbalanceamento das classes, o que permite uma avaliação mais representativa do desempenho dos modelos no contexto de detecção de intrusões.

3. Resultados

Esta seção apresenta os resultados obtidos a partir da aplicação dos modelos de Aprendizado de Máquina no conjunto de dados CICIDS2017, considerando diferentes configurações e técnicas de redução de dimensionalidade.

Inicialmente, avaliou-se o impacto da variação do parâmetro *n_estimators*, que define a quantidade de árvores utilizadas Random Forest (RF). A Tabela 1 apresenta os resultados obtidos para diferentes valores deste parâmetro.

Tabela 1. Resultados do Random Forest com diferentes valores de *n_estimators*

n_estimators	F1-Score	Tempo (min)
10	0,9978	5,2
25	0,9981	8,7
50	0,9981	14,3
100	0,9982	26,5

Observa-se que o aumento no número de árvores proporciona ganhos marginais de desempenho, mas crescimento no tempo de treinamento. Dessa forma, o valor de 25 árvores apresentou o melhor equilíbrio entre desempenho e eficiência computacional.

Para o modelo Multi-Layer Perceptron (MLP), foram avaliadas diferentes configurações de treinamento. A Tabela 2 apresenta os resultados obtidos. Para o modelo MLP, foram avaliadas diferentes configurações de hiperparâmetros, variando principalmente a arquitetura da rede e parâmetros de treinamento.

Tabela 2. Resultados do MLP com diferentes configurações

Configuração	F1-Score	Tempo (min)
Configuração 1	0,9948	10,5
Configuração 2	0,9959	13,1
Configuração 3	0,9955	15,8

As configurações consideradas são descritas a seguir: i) **Configuração 1** - arquitetura com menor número de neurônios na camada oculta e menor número de épocas de treinamento; ii) **Configuração 2** - arquitetura intermediária, com maior número de neurônios e ajuste no número de épocas, buscando melhor capacidade de generalização; e iii) **Configuração 3** - arquitetura mais complexa, com maior número de neurônios e maior tempo de treinamento.

Os resultados indicam que o MLP também apresenta alto desempenho, com pequenas variações entre as configurações testadas. No entanto, observa-se um custo computacional mais elevado em comparação ao Random Forest. Por fim, a Tabela 3 apresenta a comparação consolidada entre os modelos avaliados.

Tabela 3. Comparação entre Random Forest e Multi-Layer Perceptron

Modelo	Precisão	Recall	F1-Score	Tempo (min)
Random Forest	0,9981	0,9981	0,9981	8,7
MLP	0,9959	0,9959	0,9959	13,1

Ambos os modelos apresentaram desempenho superior a 0,99, indicando alta eficácia na classificação das classes previamente rotuladas do dataset. Ressalta-se que os resultados refletem a identificação de padrões conhecidos.

O Random Forest apresentou melhor relação entre desempenho e custo computacional, com menor tempo de treinamento face ao uso de múltiplas árvores de decisão independente. Por outro lado, o MLP apresentou desempenho semelhante, porém com maior custo computacional devido ao treinamento iterativo.

A redução de dimensionalidade com PCA contribuiu para diminuir a complexidade dos dados sem comprometer o desempenho.

Além disso, destaca-se que o conjunto de dados utilizado apresenta forte desbalanceamento entre classes, o que reforça a importância do uso de métricas ponderadas e validação estratificada para garantir uma avaliação mais confiável dos modelos.

De forma geral, os resultados indicam que o Random Forest é mais adequado para cenários que exigem eficiência computacional, enquanto o MLP pode ser explorado em contextos onde há maior disponibilidade de recursos computacionais e necessidade de modelagem mais complexa.

4. Conclusão

Este trabalho apresentou uma análise comparativa de técnicas de Aprendizado de Máquina aplicadas à detecção de intrusões em redes, utilizando o conjunto de dados CICIDS2017. Foram avaliados os modelos Random Forest e Multi-Layer Perceptron, combinados com técnicas de redução de dimensionalidade. Os resultados obtidos estão diretamente associados à capacidade dos modelos em discriminar classes conhecidas, não sendo avaliada a detecção de novos tipos de ataques.

Os resultados demonstraram que ambos os modelos são altamente eficazes na identificação de padrões de tráfego malicioso, alcançando elevados valores de F1-Score. No entanto, o modelo Random Forest apresentou melhor desempenho geral, especialmente em termos de eficiência computacional, tornando-se mais adequado para aplicações práticas que exigem processamento em tempo reduzido.

Além disso, verificou-se que a redução de dimensionalidade desempenha papel fundamental na melhoria do desempenho dos modelos, reduzindo a complexidade dos dados e o tempo de treinamento sem comprometer significativamente a precisão.

Como limitações, destaca-se a dependência de dados rotulados e o uso de um único conjunto de dados, o que pode impactar a generalização dos resultados. Dessa forma, como trabalhos futuros, sugere-se a avaliação de técnicas não supervisionadas, o uso de outros *datasets* e a aplicação de métodos mais avançados de otimização.

Por fim, os resultados obtidos reforçam o potencial das técnicas de Aprendizado de Máquina como ferramentas eficazes no suporte à segurança cibernética, contribuindo para o desenvolvimento de sistemas de detecção de intrusão mais robustos e eficientes.

Referências

Breiman, L. (2001). Random Forests. *Machine learning*, 45:5–32.

Kohavi, R. et al. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In *Ijcai*, volume 14, pages 1137–1145. Montreal, Canada.

Labonne, M. (2020). *Anomaly-Based Network Intrusion Detection Using Machine Learning*. Theses, Institut Polytechnique de Paris.

Neto, M. S. and Gomes, D. G. (2019). Network Intrusion Detection Systems Design: A Machine Learning Approach. In *Anais do XXXVII SBRC*, pages 932–945, Porto Alegre, RS, Brasil. SBC.

Pearson, K. (2010). On Lines and Planes of Closest Fit to Systems of Points in Space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572.

Rosay, A., Carlier, F., and Leroux, P. (2020). MLP4NIDS: An Efficient MLP-Based Network Intrusion Detection for CICIDS2017 Dataset. In *Machine Learning for Networking*, pages 240–254, Cham. Springer International Publishing.

Sharafaldin, I., Lashkari, A. H., and Ghorbani, A. A. (2018). Toward Generating a New Intrusion Detection Dataset and Intrusion Traffic Characterization. *ICISSP*, 1(2018):108–116.