# Detecção de Intrusões em Redes: Uma Análise Comparativa entre Aprendizado Supervisionado e Não Supervisionado

**Abdiel A. Gouveia Silva, Andresa Lídia de F. Martins, Pedro Clarindo da S. Neto**

[1] Instituto Federal de Mato Grosso – Campus Cuiabá Cel. Octayde Jorge da Silva
R. Zulmira Canavarros, 95 – Centro – 78005-390 – Cuiabá – MT – Brasil

{abdiel.s, andresa.figueiredo}@estudante.ifmt.edu.br,
pedro.neto@ifmt.edu.br

***Abstract.*** *The growth of cyber attacks has increased the need for intelligent intrusion detection mechanisms. This paper presents a comparative analysis between supervised and unsupervised learning approaches applied to the NSL-KDD dataset. Random Forest and Isolation Forest were evaluated. The results show comparable overall performance. While Random Forest achieved high performance during training, it presented signs of overfitting on the test set. Isolation Forest demonstrated more balanced detection capability, especially for attack patterns.*

***Resumo.*** *O crescimento dos ataques cibernéticos exige mecanismos inteligentes para detecção de intrusões. Este trabalho apresenta uma análise comparativa entre abordagens supervisionada e não supervisionada aplicadas ao conjunto de dados NSL-KDD, utilizando Random Forest e Isolation Forest. Os resultados indicam desempenho global semelhante entre os modelos. O Random Forest apresentou indícios de sobreajuste no teste, enquanto o Isolation Forest demonstrou maior equilíbrio na detecção de ataques.*

## 1. Introdução

O aumento do uso de redes e sistemas conectados tem intensificado os riscos à segurança da informação, com ataques cibernéticos cada vez mais frequentes e complexos, exigindo mecanismos capazes de identificar comportamentos anômalos no tráfego de rede [Cavalcanti 2022].

Nesse cenário, os Sistemas de Detecção de Intrusão (IDS) passaram a incorporar técnicas de Inteligência Artificial e Aprendizado de Máquina, devido à sua capacidade de adaptação e generalização frente a novos padrões de ataque [Oliveira et al. 2018, Casanova and Ponti 2019].

O Aprendizado de Máquina pode ser dividido em abordagens supervisionadas e não supervisionadas, diferenciando-se principalmente pela necessidade de dados rotulados [Duarte and Silva 2021]. Assim, este trabalho propõe uma análise comparativa dessas abordagens aplicadas à detecção de intrusões em redes, utilizando o conjunto de dados NSL-KDD.

## 2. Metodologia

Este trabalho caracteriza-se como uma pesquisa experimental e exploratória, aliada a uma revisão bibliográfica, com o objetivo de analisar o comportamento de algoritmos de aprendizado de máquina aplicados à detecção de intrusões em redes de computadores.

### 2.1. Conjunto de Dados

Para a realização dos experimentos, foi utilizado o conjunto de dados NSL-KDD, que é geralmente utilizado em pesquisas sobre detecção de intrusões em redes de computadores [Ahmed et al. 2016]. Esse conjunto de dados é uma versão aprimorada do KDD Cup 1999, desenvolvida com o objetivo de reduzir redundâncias e melhorar a qualidade dos dados.

O NSL-KDD é composto por atributos numéricos e categóricos que descrevem características das conexões de rede, além de um rótulo que indica se a conexão corresponde a tráfego normal ou a algum tipo de ataque. Neste trabalho, os rótulos foram convertidos para uma classificação binária, considerando apenas as classes normal e ataque.

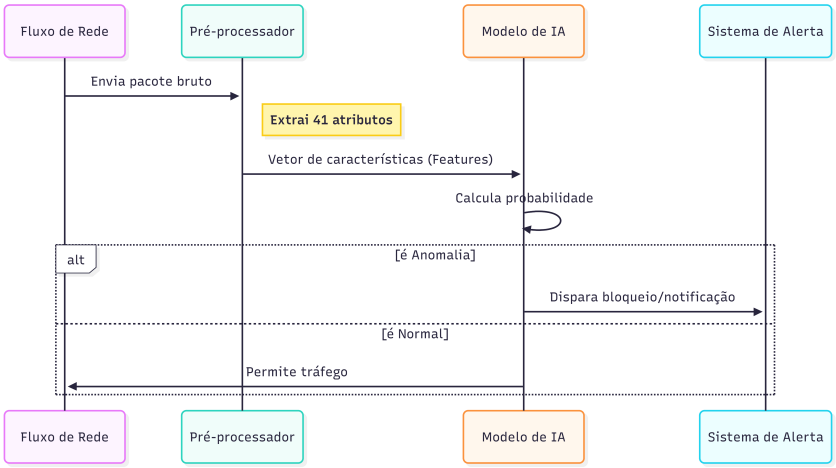


**Figura 1. Estrutura do conjunto de dados NSL-KDD**

### 2.2. Ferramentas Utilizadas

Os experimentos foram desenvolvidos em *Python*, linguagem amplamente utilizada em pesquisas em Inteligência Artificial e Ciência de Dados. As bibliotecas *Pandas* e *Scikit-learn* foram empregadas para o pré-processamento dos dados, implementação dos algoritmos e avaliação dos modelos [LeCun et al. 2015].

A escolha dessas ferramentas deve-se à sua robustez, ampla documentação e uso consolidado em estudos de aprendizado de máquina e cibersegurança [Machado et al. 2019].

### 2.3. Pré-processamento dos Dados

O pré-processamento dos dados envolveu a separação entre atributos e rótulos, a conversão dos rótulos para classificação binária (normal e ataque) e a aplicação de codificação

*one-hot* para atributos categóricos. Além disso, foi realizado o alinhamento das colunas entre os conjuntos de treinamento e teste, garantindo compatibilidade entre os dados fornecidos aos modelos [Duarte and Silva 2021].

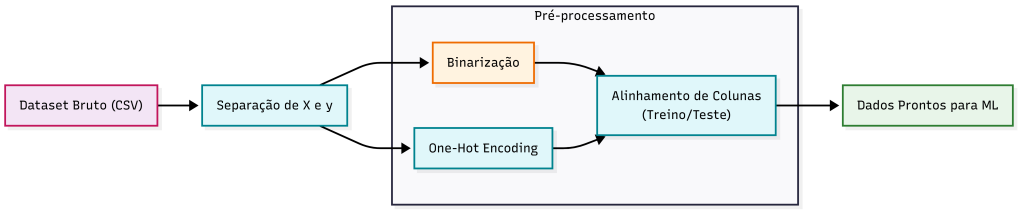


**Figura 2. Fluxo de pré-processamento dos dados**

## 3. Conceitos Básicos

### 3.1. Aprendizado Supervisionado

O aprendizado supervisionado é uma abordagem em que os modelos são treinados a partir de dados previamente rotulados, permitindo aprender a relação entre atributos de entrada e classes de saída [Casanova and Ponti 2019]. Essa abordagem é amplamente utilizada em tarefas de classificação, apresentando alto desempenho quando há disponibilidade de dados confiáveis. O Random Forest é um algoritmo supervisionado baseado em múltiplas árvores de decisão, conhecido por sua robustez e capacidade de reduzir sobreajuste.

### 3.2. Aprendizado Não Supervisionado

O aprendizado não supervisionado busca identificar padrões e estruturas em dados não rotulados, sem conhecimento prévio das classes [Duarte and Silva 2021]. O Isolation Forest é uma técnica voltada à detecção de anomalias, fundamentada no princípio de que comportamentos anômalos são mais facilmente isoláveis do que padrões normais [Ahmed et al. 2016]. Embora apresente maior flexibilidade, essa abordagem tende a obter menor desempenho em métricas tradicionais quando comparada ao aprendizado supervisionado.

## 4. Resultados e Discussões

O Random Forest obteve acurácia de 76,78% no conjunto de teste, enquanto o Isolation Forest alcançou 78,95%. Apesar da alta acurácia média durante a validação cruzada no treinamento (99,87%), o desempenho inferior no teste indica possível sobreajuste, conforme discutido na literatura [Breiman 2001].

A análise da matriz de confusão mostra que o modelo supervisionado apresentou elevado recall para a classe normal (97%), porém recall reduzido para ataques (61%), evidenciando dificuldade na generalização de padrões maliciosos. Em aplicações de segurança, essa limitação pode ser crítica, pois falsos negativos representam ataques não detectados.

O Isolation Forest apresentou desempenho mais equilibrado, com recall de 83% para ataques, demonstrando maior robustez na identificação de anomalias, especialmente em ambientes dinâmicos [Souza and Santana 2020]. Esses resultados reforçam que modelos supervisionados podem alcançar bom desempenho quando bem ajustados, enquanto

abordagens não supervisionadas oferecem maior flexibilidade frente a padrões desconhecidos [Cavalcanti 2022].

**Tabela 1. Resultados e matriz de confusão do modelo supervisionado (Random Forest)**

| **Métricas de Classificação** | | | | |
|---|---|---|---|---|
| Classe | Precisão | Recall | F1-score | Suporte |
| Ataque | 0.97 | 0.61 | 0.75 | 12833 |
| Normal | 0.66 | 0.97 | 0.78 | 9711 |
| Acurácia | | | 0.77 | |

| **Matriz de Confusão** | | |
|---|---|---|
| | Predito: Ataque | Predito: Normal |
| Real: Ataque | 7862 | 4971 |
| Real: Normal | 263 | 9448 |

**Tabela 2. Resultados e matriz de confusão do modelo não supervisionado (Isolation Forest)**

| **Métricas de Classificação** | | | | |
|---|---|---|---|---|
| Classe | Precisão | Recall | F1-score | Suporte |
| Ataque | 0.80 | 0.83 | 0.82 | 12833 |
| Normal | 0.77 | 0.73 | 0.75 | 9711 |
| Acurácia | | | 0.79 | |

| **Matriz de Confusão** | | |
|---|---|---|
| | Predito: Ataque | Predito: Normal |
| Real: Ataque | 10691 | 2142 |
| Real: Normal | 2603 | 7108 |

## 5. Considerações Finais

Este estudo demonstrou a aplicabilidade de técnicas de aprendizado de máquina na detecção de intrusões em rede, comparando abordagens supervisionadas e não supervisionadas. Os resultados indicaram que ambos os modelos apresentaram desempenho satisfatório, com destaque para o Isolation Forest na identificação de padrões anômalos de forma mais equilibrada.

Observou-se que o Random Forest, embora eficiente durante o treinamento, apresentou indícios de sobreajuste quando avaliado em dados não vistos, reforçando a importância da validação adequada e do ajuste de hiperparâmetros. Além disso, a análise da matriz de confusão evidenciou que, em contextos de segurança, métricas como recall para a classe de ataque são tão ou mais relevantes que a acurácia global.

Como trabalhos futuros, sugere-se a ampliação do conjunto de dados, a aplicação de técnicas de balanceamento de classes e a avaliação de modelos híbridos, combinando métodos supervisionados e não supervisionados. Tais estratégias podem contribuir para sistemas de detecção mais robustos e adaptáveis a cenários reais.

## Referências


Ahmed, M., Mahmood, A. N., and Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications*, 60:19–31.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

Casanova, E. and Ponti, M. A. (2019). Aprendizado de máquina: conceitos, aplicações e desafios. *Revista de Informática Teórica e Aplicada (RITA)*, 26(4):11–27. Language: Portuguese.

Cavalcanti, J. C. (2022). O avanço da inteligência artificial e seus impactos sociais. *Revista Brasileira de Ciência, Tecnologia e Sociedade (RBCTS)*, 13(1). Language: Portuguese.

Duarte, D. G. S. and Silva, M. J. S. (2021). Aprendizado supervisionado e não supervisionado: uma revisão sistemática. *Revista Eletrônica de Sistemas de Informação (RESI)*, 20(2). Language: Portuguese.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436–444.

Machado, R. S., Zarpelão, B. B., and Sampaio, L. N. (2019). Detecção de intrusões em redes de computadores utilizando aprendizado de máquina. *Revista Brasileira de Computação Aplicada (RBCA)*, 11(2):1–14. Language: Portuguese.

Oliveira, M. d., Oliveira, A. d., and Silva, L. (2018). Inteligência artificial: conceitos, história e aplicações. *Revista Tecnologia e Sociedade*, 14(31). Language: Portuguese.

Souza, F. R. and Santana, L. C. (2020). Detecção de anomalias em redes de computadores utilizando técnicas de aprendizado não supervisionado. *Revista Eletrônica de Sistemas de Informação*, 19(1). Language: Portuguese.