**Aplicação de Algoritmos de Aprendizado de Máquina para Detecção de Fraudes Financeiras**

GT: Inteligência Artificial e Análise de Dados

Leonardo de Lima Póss

Acadêmico do Curso de Engenharia de Software do Centro Universitário Cesumar – UNICESUMAR, Curitiba – PR. leonardoposs@hotmail.com

**RESUMO:** A modernização dos serviços financeiros trouxe melhorias importantes, mas também aumentou os riscos de fraudes em operações bancárias. Este estudo, desenvolvido com fins acadêmicos, teve como objetivo aplicar algoritmos de machine learning para identificar possíveis fraudes em um conjunto de dados simulado contendo 50.000 entradas. A biblioteca PyCaret foi escolhida para facilitar a comparação de desempenho entre diferentes algoritmos supervisionados, incluindo Naive Bayes, Regressão Logística, Random Forest e LightGBM. Com a PyCaret, foi possível implementar de forma rápida e eficiente cada uma dessas técnicas, além de contar com uma interface simplificada para a avaliação e comparação dos resultados obtidos. Dentre os algoritmos analisados, o Naive Bayes se destacou, apresentando os melhores desempenhos em métricas como recall e F1-score, o que evidenciou sua eficácia, mesmo sendo uma abordagem simples. Ferramentas adicionais, como YData Profiling e D-Tale, foram empregadas na exploração dos dados, proporcionando uma análise detalhada do perfil dos registros. Estas ferramentas facilitaram a compreensão das distribuições e outliers, além de auxiliarem na preparação adequada dos dados para os algoritmos de aprendizado de máquina. Os resultados indicam que, mesmo em contextos acadêmicos, é possível alcançar bons desempenhos na detecção de fraudes utilizando ferramentas acessíveis e metodologias bem estruturadas, destacando a importância do uso de técnicas analíticas e algoritmos eficientes para a proteção financeira.

**PALAVRAS-CHAVE:** Análise de Padrões, Detecção de Fraudes, Machine Learning.

# **INTRODUÇÃO**

As fraudes bancárias eletrônicas têm se tornado uma crescente preocupação global com o avanço das tecnologias digitais e a expansão dos serviços financeiros online. A natureza dessas fraudes abrange desde subtração dos dados pessoais e segue com invasão em dispositivos, invasão das contas bancárias com movimentações financeiras não autorizadas, causando prejuízos significativos para usuários dos serviços, pessoas físicas, jurídicas, incluindo obviamente as instituições bancárias (Bittencourt, 2024).

É importante registrar que, além das perdas econômicas diretas, esses crimes afetam a confiança de todo ecossistema financeiro impactando na economia de maneira mais ampla, gerando incerteza e desestabilização o sistema financeiro. Alguns aspectos são importantes para entender esse fenômeno, como as novas formas de inclusão de pessoas nos programas de governo, associados a nova tecnologias do sistema bancário a exemplo do arranjo pix, o qual trouxe grandes avanços no que se referente a instantaneidade das transações bancárias. (Bittencourt, 2024).

Nos últimos anos, o avanço da tecnologia tem proporcionado um aumento na complexidade dos golpes financeiros, afetando um número crescente de pessoas. Segundo pesquisa da Confederação Nacional de Dirigentes Lojistas (CNDL) e do Serviço de Proteção ao Crédito (SPC Brasil), cerca de 46% dos internautas brasileiros foram vítimas de golpes financeiros nos últimos 12 meses, totalizando mais de 12 milhões de pessoas (Bacen, 2020).

Esses crimes representam uma preocupação crescente para as instituições financeiras e seus clientes, resultando em danos financeiros significativos, roubo de dados pessoais e outros impactos negativos. Instituições financeiras desempenham um papel fundamental na sociedade ao gerenciar o dinheiro de milhões de pessoas e empresas. No entanto, a vulnerabilidade e complexidade das operações online possibilitam a ocorrência de diversas fraudes (Rezende, 2010, p.77).

Este trabalho tem por objetivo, por meio da aplicação de algoritmos de Machine Learning (LM), desenvolver uma metodologia capaz de identificar com precisão possíveis transações fraudulentas em um grande volume de dados financeiros. Considerando o aumento significativo de fraudes bancárias no cenário digital, o uso dessas técnicas permite automatizar o processo de detecção, proporcionando uma análise mais eficiente e assertiva, além de minimizar a dependência de intervenções manuais, o que torna a prevenção de fraudes mais ágil e eficaz. Ao integrar algoritmos de aprendizado de máquina, como Naive Bayes, Regressão Logística e Random Forest, pretende-se explorar diferentes abordagens e comparar seus desempenhos, visando aprimorar a capacidade de prever comportamentos fraudulentos.

1. **DESENVOLVIMENTO**

## Fundamentação Teórica

Ao desenvolver modelos de aprendizado de máquina para a detecção de fraudes financeiras, a escolha das métricas de avaliação é essencial para garantir que o desempenho do algoritmo represente adequadamente a realidade do problema. Em cenários nos quais há um forte desbalanceamento entre as classes, como no caso das transações legítimas em comparação com as fraudulentas, métricas tradicionais como a acurácia podem levar a interpretações equivocadas. Um modelo que classifica todas as transações como legítimas ainda assim apresentaria uma alta taxa de acerto, ignorando completamente as fraudes reais e comprometendo sua utilidade prática.

Por esse motivo, é importante utilizar métricas que ofereçam uma visão mais precisa da capacidade do modelo em identificar as fraudes, mesmo que elas representem uma minoria dos casos. Dentre essas métricas, o Recall e o F1-Score se destacam por sua relevância em contextos de alto risco e baixa incidência de eventos fraudulentos.

O Recall, também conhecido como sensibilidade, é uma métrica fundamental na avaliação de modelos de classificação, especialmente em cenários de detecção de fraudes, nos quais os falsos negativos representam um risco significativo. Ele mede a proporção de instâncias positivas corretamente identificadas, ou seja, quantas transações fraudulentas reais foram detectadas pelo algoritmo em relação ao total de fraudes existentes no conjunto de dados (RIBEIRO, 2021).

Já o F1-Score é utilizado para encontrar um equilíbrio entre o Recall e a Precisão, sendo calculado por meio da média harmônica dessas duas métricas. Essa medida é particularmente útil quando se trabalha com bases de dados desbalanceadas, como ocorre com frequência no sistema financeiro, onde o número de transações legítimas supera amplamente o número de fraudes. Assim, o F1-Score oferece uma visão mais realista do desempenho do modelo do que apenas a acurácia (RIBEIRO, 2021).

Segundo Chandradeva et al. (2020), os métodos tradicionais de detecção de fraudes são ultrapassados frente ao avanço das transações online, tornando necessária a adoção de técnicas mais dinâmicas, como o machine learning.

Para Ahmadi, Yousefi e Rezazadeh (2021), a categorização e validação de clientes bancários por meio de redes neurais pode apresentar resultados significativamente superior aos métodos convencionais.

Varmedja et al. (2019) destacam a eficácia do algoritmo Random Forest na detecção de fraudes em cartões de crédito, especialmente em bases desbalanceadas tratadas com SMOTE.

Já Khatri, Arora e Agrawal (2020) apontam a Árvore de Decisão como o modelo mais sensível ao tempo de resposta, enfatizando a necessidade de pesquisas que considerem também métodos não supervisionados.

## Metodologia

## Tipo da Pesquisa

Este trabalho caracteriza-se como uma pesquisa experimental e quantitativa, pois utiliza algoritmos de aprendizado de máquina (machine learning) para análise e detecção de fraudes financeiras. A proposta envolve a implementação e comparação de diferentes modelos supervisionados, com base em métricas de desempenho específicas.

## Conjunto de Dados

O conjunto de dados utilizado nesta pesquisa foi extraído de uma fonte pública do Kaggle, desenvolvido para modelagem de detecção de fraudes financeiras. Ele contém informações realistas sobre transações financeiras, incluindo dados numéricos, categóricos e temporais, o que o torna ideal para tarefas de classificação binária. O conjunto abrange 21 características que descrevem diversos aspectos das transações, como o tipo de transação, o valor envolvido, o saldo da conta do usuário antes da transação, a categoria do comerciante, o risco associado à transação, e o histórico de atividades fraudulentas do usuário, entre outros.

A variável alvo, chamada Fraud\_Label, indica se a transação foi fraudulenta (1) ou não (0). Além disso, o conjunto de dados é projetado para promover o desenvolvimento de modelos de aprendizado de máquina precisos, sendo adequado para a construção de sistemas robustos de detecção de fraudes.

## Ferramentas Utilizadas

Neste estudo, diversas ferramentas e ambientes computacionais foram utilizados para executar e desenvolver os experimentos. O projeto foi realizado na linguagem Python, versão 3.10.4, amplamente reconhecida por sua robustez e pelas bibliotecas especializadas em ciência de dados e aprendizado de máquina.

A biblioteca PyCaret foi empregada para facilitar a implementação, o treinamento e a comparação dos algoritmos de classificação. Trata-se de uma ferramenta de baixo código que permite acelerar o desenvolvimento de modelos preditivos por meio de uma interface simplificada e integração com diversas bibliotecas, como scikit-learn, entre outras (PYCARET, 2023).

O conjunto de dados foi obtido na plataforma Kaggle, um repositório popular entre pesquisadores e especialistas em ciência de dados. O dataset selecionado é sintético e foi criado para simular transações financeiras com o propósito de capacitar modelos de detecção de fraudes.

Além disso, foram empregadas duas ferramentas adicionais: D-Tale e YData Profiling. O D-Tale foi utilizado para facilitar a visualização e a exploração interativa dos dados, permitindo uma análise mais dinâmica e intuitiva das variáveis e distribuições do conjunto de dados. Por sua vez, o YData Profiling foi utilizado para criar relatórios automáticos e minuciosos sobre as variáveis do dataset, sendo fundamental para identificar a variável resposta (target) e compreender melhor a estrutura dos dados antes da fase de modelagem.

## Algoritmos Utilizados

* **Naive Bayes**

O algoritmo Naive Bayes é um classificador probabilístico baseado no Teorema de Bayes, que assume independência entre as variáveis preditoras. É amplamente utilizado em tarefas de classificação devido à sua simplicidade e eficiência, especialmente em conjuntos de dados de alta dimensionalidade.

* **Regressão Logística**

A Regressão Logística é um modelo estatístico utilizado para prever a probabilidade de uma variável dependente binária, como a ocorrência de fraude (sim/não). Sua aplicação em detecção de fraudes é comum, pois permite identificar padrões de comportamento associados a transações fraudulentas.

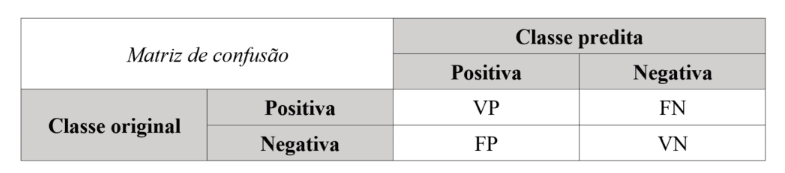
* **Random Forest**

O Random Forest é um método de aprendizado de máquina baseado em múltiplas árvores de decisão, que melhora a precisão do modelo por meio da agregação dos resultados de diversas árvores. Este algoritmo é particularmente eficaz em lidar com conjuntos de dados desbalanceados, como é comum em problemas de detecção de fraudes, onde as transações fraudulentas representam uma pequena fração do total.

* **LightGBM**

O LightGBM (Light Gradient Boosting Machine) é um framework de aprendizado de máquina baseado em boosting, projetado para ser eficiente em termos de tempo e memória. Ele é especialmente eficaz em conjuntos de dados grandes e desbalanceados, como os encontrados em sistemas de detecção de fraudes.

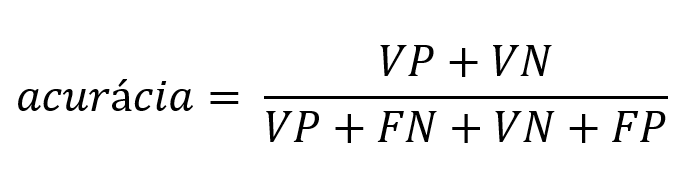
## Métricas de Avaliação

A matriz de confusão é uma ferramenta essencial para avaliar o desempenho de modelos de classificação binária, especialmente em contextos de detecção de fraudes, onde os dados costumam ser desbalanceados. Ela organiza os resultados da predição em quatro categorias: verdadeiros positivos (TP), verdadeiros negativos (TN), falsos positivos (FP) e falsos negativos (FN), permitindo uma análise detalhada dos acertos e erros do modelo (Fawcett, 2006). Ao invés de apenas depender da acurácia, que pode ser enganosa em conjuntos de dados com classes desiguais, a matriz de confusão possibilita a extração de métricas mais informativas, como precisão, recall e F1-score (Saito & Rehmsmeier, 2015).

**Figura 1** – Matriz Confusão

Fonte: COSTA (2020).

A acurácia é uma métrica básica que demonstra a proporção de previsões corretas feitas por um modelo em comparação com o número total de casos. Apesar de ser útil para uma compreensão geral do desempenho, a acurácia pode ser ilusória em situações de dados desbalanceados, onde o modelo pode ter uma vantagem em relação à classe predominante.

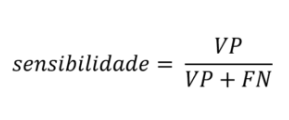


**Figura 2** – Fórmula Acurácia

Fonte: COSTA (2020).

A AUC, ou Área sob a Curva ROC, avalia a habilidade do modelo em diferenciar entre as classes positiva e negativa. A curva ROC, que representa características operacionais do receptor, exibe a taxa de verdadeiros positivos contra a taxa de falsos positivos. Um valor de AUC próximo de 1 denota um modelo extremamente eficiente, enquanto scores baixos indicam que o modelo enfrenta dificuldades em distinguir entre as classes.

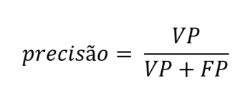
O recall ou também sensibilidade, também conhecido como sensibilidade, mede a proporção de fraudes reais que o modelo conseguiu identificar de maneira correta. Em situações de fraude, onde o impacto de não reconhecer uma transação suspeita é elevado, o recall se torna uma métrica fundamental para evitar que fraudes passem despercebidas.



**Figura 3** – Fórmula Acurácia

Fonte: COSTA (2020).

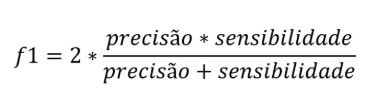
A precisão calcula a fração de transações que foram rotuladas como fraudulentas e que de fato são fraudes. Embora a precisão tenha sua importância, um modelo que exibe uma alta precisão pode ter um recall reduzido, particularmente em cenários de dados desbalanceados, onde a maioria das transações é autêntica.



**Figura 4** – Fórmula Acurácia

Fonte: COSTA (2020).

O F1-Score integra o recall e a precisão em uma métrica única, computando a média harmônica entre os dois. Está métrica é especialmente valiosa em situações onde há uma grande disparidade entre as classes, sendo necessário equilibrar a detecção de fraudes com a minimização de falsos positivos.

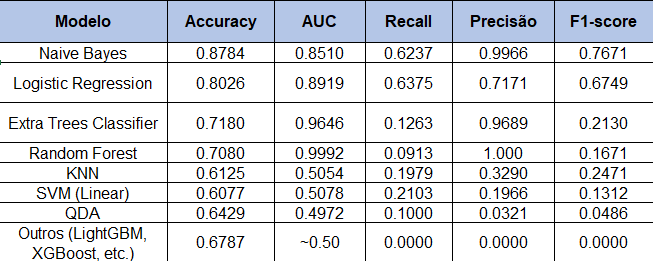


**Figura 5** – Fórmula Acurácia

Fonte: COSTA (2020).

## Resultados e Discuções

Os resultados apresentados na Figura 6 referem-se à execução do código de classificação e comparação de diferentes modelos, conforme destacados na referida figura. O processo de preparação dos dados foi realizado de maneira automatizada, graças à funcionalidade do Pycaret. De forma automática, as variáveis categóricas e numéricas foram separadas, e a variável resposta do algoritmo foi definida como "fraud\_label", que contém o valor binário determinante para indicar se a transação é fraudulenta (1) ou não (0).

**Figura 6 –** Tabela de resultados

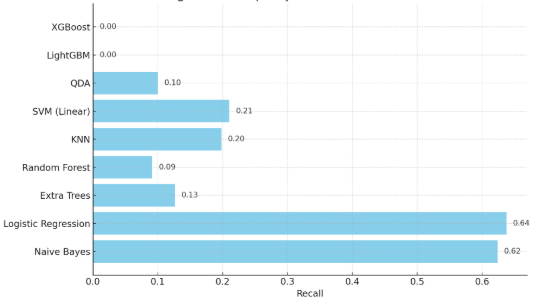
Fonte: O autor

As métricas mais importantes para a comparação dos modelos de classificação foram Acurácia, AUC (Área sob a Curva ROC), Recall, Precisão e F1-Score.

A Acurácia forneceu uma visão geral do desempenho dos modelos, indicando a proporção de acertos em relação ao total de previsões.

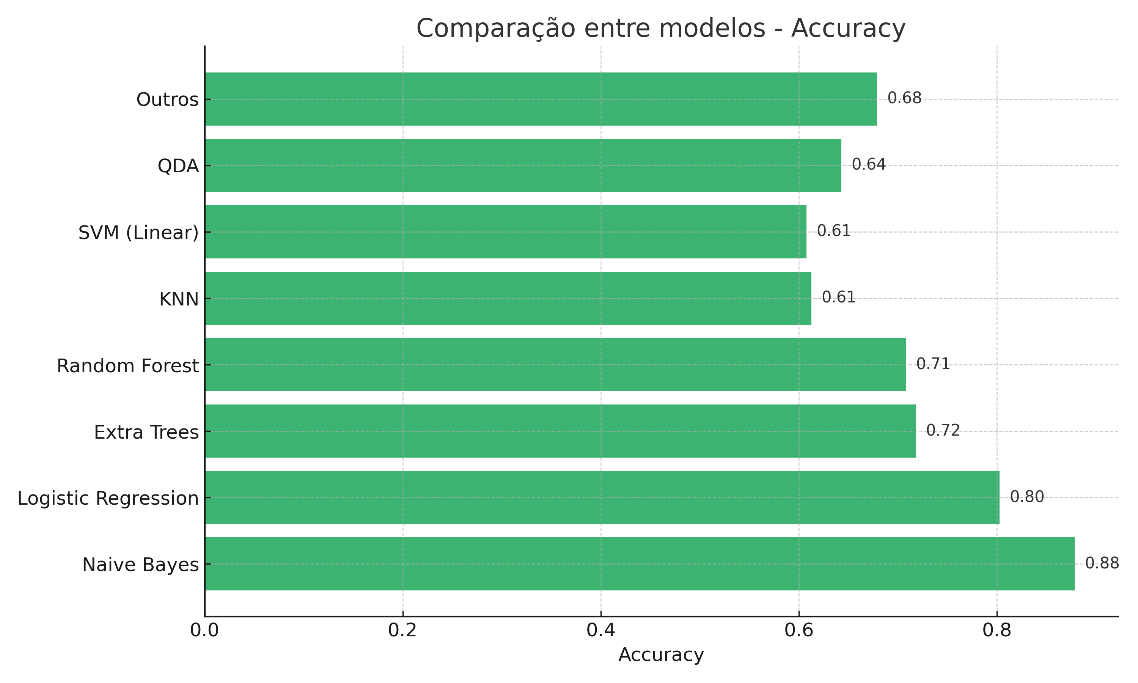
A AUC foi fundamental para avaliar a capacidade de discriminação dos modelos entre as classes de fraude e não fraude, independentemente do limiar de decisão.

O Recall destacou a habilidade de cada modelo em identificar transações fraudulentas, enquanto a Precisão foi importante para verificar a exatidão das fraudes identificadas.

Por fim, o F1-Score foi utilizado para equilibrar os valores de recall e precisão, proporcionando uma visão mais equilibrada sobre a performance do modelo, especialmente em cenários de classes desbalanceadas, como foi o caso encontrado no conjunto de dados utilizado.

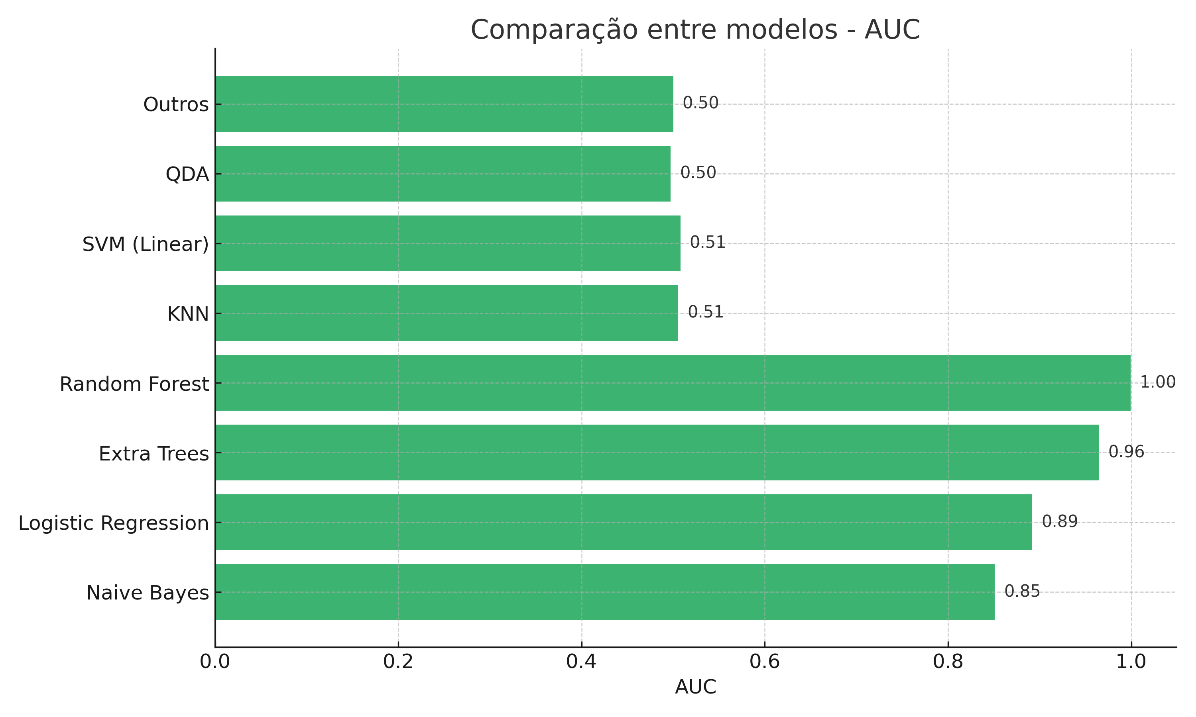
**Figura 7 –** Gráfico de comparação Recall

Fonte: O autor



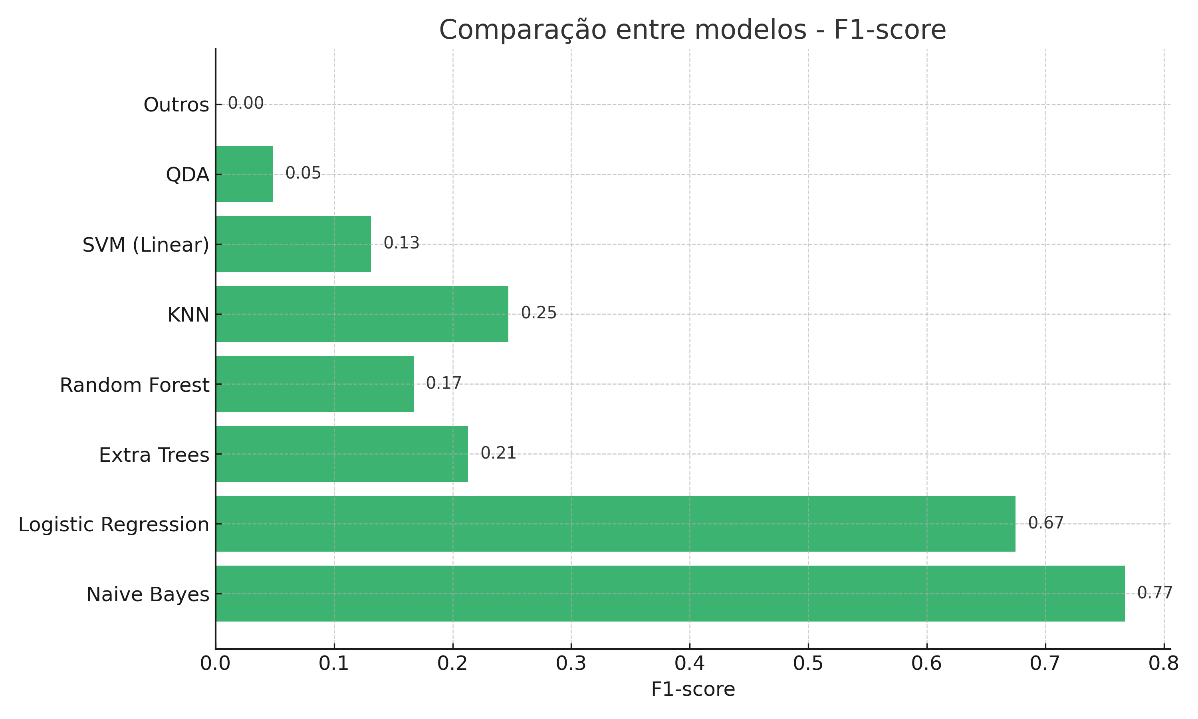
**Figura 8 –** Gráfico de comparação Acurácia

Fonte: O autor



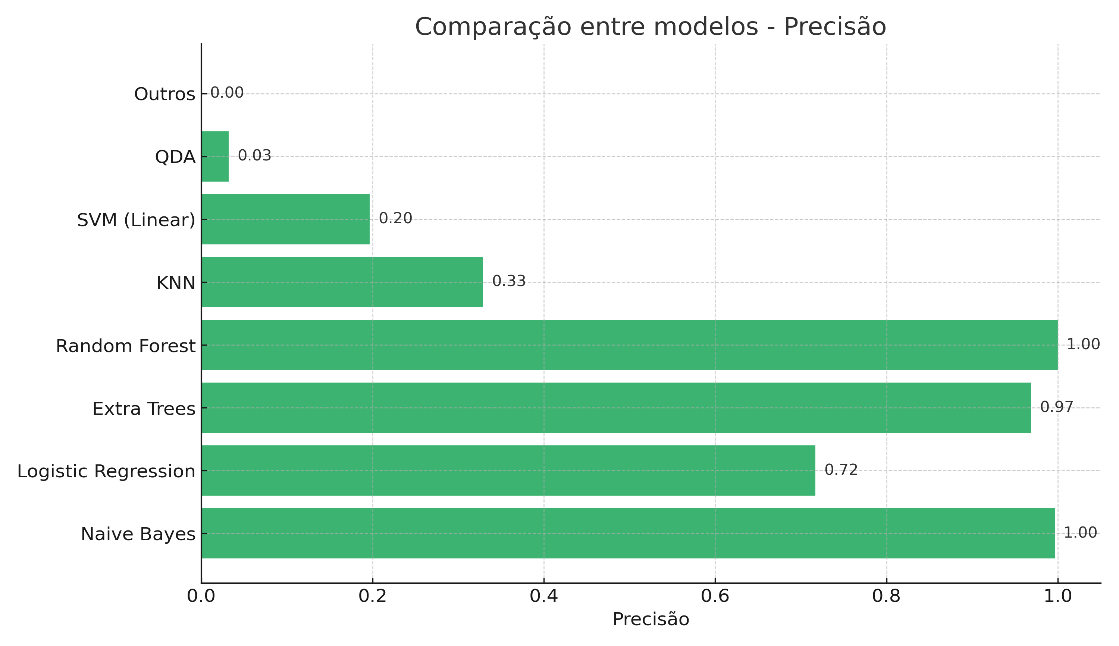
**Figura 9 –** Gráfico de comparação AUC

Fonte: O autor



**Figura 10 –** Gráfico de comparação F1-Score

Fonte: O autor



**Figura 11 –** Gráfico de comparação Precisão

Fonte: O autor

O modelo Naive Bayes destacou-se com o melhor desempenho geral, principalmente em termos de acurácia (87,84%) e precisão (99,66%), mostrando-se extremamente eficaz na detecção de transações legítimas. Embora seu recall (62,37%) não tenha sido o mais elevado, o equilíbrio apresentado por um F1-score de 76,71% reforça sua confiabilidade e força no contexto analisado.

A Regressão Logística também demonstrou um desempenho consistente, alcançando a AUC mais alta (89,19%) entre todos os modelos, o que ressalta sua habilidade em diferenciar entre classes fraudulentas e não fraudulentas. Com um recall de 63,75% e uma precisão razoável, o modelo obteve um F1-score de 67,49%, posicionando-se como uma alternativa viável, mesmo que um pouco abaixo do Naive Bayes.

Por outro lado, os modelos Random Forest e Extra Trees Classifier se destacaram pela alta precisão (100% e 96,89%, respectivamente), mas apresentaram um recall muito baixo (9,13% e 12,63%). Esses resultados indicam que, embora eficazes em minimizar falsos positivos, esses algoritmos falharam em detectar a maioria das fraudes, o que representa um risco significativo em contextos financeiros.

Os modelos KNN, SVM e QDA tiveram um desempenho mais modesto em quase todas as métricas. Essa limitação pode ser atribuída à dificuldade desses algoritmos em lidar com dados que são desbalanceados e possuem alta dimensionalidade, afetando sua utilização neste contexto.

Finalmente, os modelos LightGBM e XGBoost, reconhecidos por suas performances em diversas competições, não alcançaram resultados satisfatórios nesta aplicação. Todas as métricas foram zeradas, possivelmente devido à falta de otimização nos hiperparâmetros ou à sensibilidade desses modelos às distribuições de classe presentes no conjunto de dados.

## CONCLUSÃO

Com base nos resultados obtidos, é possível afirmar que o modelo Naive Bayes se destacou com o melhor desempenho geral na tarefa de identificação de fraudes, combinando alta precisão com um F1-Score satisfatório, o que é crucial em situações com dados desbalanceados. Além disso, a investigação mostrou que o uso de ferramentas como o PyCaret pode acelerar e simplificar a experimentação com diversos algoritmos, tornando o processo mais acessível e reproduzível. Contudo, ainda existem algumas limitações, como a necessidade de ajuste minucioso de hiperparâmetros em modelos mais elaborados e a dificuldade em gerenciar um desbalanceamento severo entre as classes, o que prejudica a eficácia de algoritmos como Random Forest e XGBoost. Para futuros trabalhos, sugere-se a aplicação de técnicas de balanceamento, como o SMOTE ou undersampling, além de realizar análises mais aprofundadas sobre a engenharia de atributos e o efeito de variáveis específicas na detecção de fraudes. Seria também interessante investigar abordagens híbridas ou baseadas em redes neurais, que têm mostrado resultados promissores em cenários semelhantes.

## REFERÊNCIAS

AHMADI, P.; YOUSEFI, A.; REZAZADEH, F. Benefiting machine learning methods to detect fraud in the validation of bank customers' cards. Journal of Bioengineering Research, v. 3, n. 2, p. 12–19, 2021.

BACEN. Notícia sobre economia. Disponível em: https://www.bcb.gov.br/detalhenoticia/412/noticia. Acesso em: 18 abr. 2025.

BACEN. Notícia sobre economia. Disponível em: https://www.bcb.gov.br/detalhenoticia/13011/nota. Acesso em: 18 abr. 2025.

BITTENCOURT, Delmar Araújo. O impacto das fraudes bancárias eletrônicas para a sociedade. 2024. Monografia (Curso Superior de Segurança e Defesa Cibernética) – Escola Superior de Guerra, Rio de Janeiro, 2024. Disponível em: https://repositorio.esg.br/handle/123456789/1961. Acesso em: 19 abr. 2025.

CHANDRADEVA, L. S. et al. Monetary transaction fraud detection system based on machine learning strategies. In: INTERNATIONAL CONGRESS ON INFORMATION AND COMMUNICATION TECHNOLOGY, 4., 2019, London. Proceedings. Singapore: Springer, 2020. v. 1, p. 385–396.

COSTA, L. Métricas de avaliação em Machine Learning: acurácia, sensibilidade, precisão, especificidade e F-score. BioInfo, 2020. Disponível em: https://bioinfo.com.br/metricas-de-avaliacao-em-machine-learning-acuracia-sensibilidade-precisao-especificidade-e-f-score/. Acesso em: 19 abr. 2025.

KAGGLE. Fraud Detection Transactions Dataset. Disponível em: https://www.kaggle.com/datasets/samayashar/fraud-detection-transactions-dataset. Acesso em: 19 abr. 2025.

KHATRI, S.; ARORA, A.; AGRAWAL, A. P. Supervised machine learning algorithms for credit card fraud detection: a comparison. In: INTERNATIONAL CONFERENCE ON CLOUD COMPUTING, DATA SCIENCE & ENGINEERING (CONFLUENCE), 10., 2020, Noida. Proceedings. IEEE, 2020. p. 680–683.

PYCARET. PyCaret – An open-source, low-code machine learning library in Python. Disponível em: https://pycaret.org/. Acesso em: 19 abr. 2025.

PYTHON SOFTWARE FOUNDATION. Python Language Reference, version 3.10.4. Disponível em: https://www.python.org/. Acesso em: 19 abr. 2025.

REZENDE, Frederico Antonio Oliveira. Responsabilidade civil dos bancos em relação às fraudes eletrônicas. FMU Direito – Revista Eletrônica, v. 24, n. 33, 2010. ISSN: 2316-1515.

RIBEIRO, Thayná Abreu. Trabalho de Conclusão de Curso. Universidade Federal Fluminense, 2021. Disponível em: https://app.uff.br/riuff/bitstream/handle/1/14746/Trabalho%20de%20Conclus%C3%A3o%20de%20Curso%20-%20Thayn%C3%A1%20Abreu%20Ribeiro.pdf?isAllowed=y&sequence=1. Acesso em: 19 abr. 2025.

SAITO, T.; REHMSMEIER, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS ONE, v. 10, n. 3, p. e0118432, 2015. Disponível em: https://doi.org/10.1371/journal.pone.0118432. Acesso em: 19 abr. 2025.

VARMEDJA, D. et al. Credit card fraud detection machine learning methods. In: INTERNATIONAL SYMPOSIUM INFOTEH JAHORINA, 18., 2019, Jahorina. Proceedings. IEEE, 2019. p. 1–5.

**Application of Machine Learning Algorithms for Financial Fraud Detection**

**ABSTRACT:** The modernization of financial services has brought significant improvements but also increased the risks of fraud in banking operations. This study applied machine learning algorithms to identify potential fraud in a simulated dataset with 50,000 entries. Using the PyCaret library, several supervised algorithms, including Naive Bayes, Logistic Regression, Random Forest, and LightGBM, were compared. Naive Bayes demonstrated the best performance in recall and F1-score metrics, highlighting its effectiveness despite being a simple approach. Additional tools, such as YData Profiling and D-Tale, aided in data exploration and preparation. The results show that, even in academic contexts, it is possible to achieve good performance in fraud detection using accessible tools and well-structured methodologies, emphasizing the importance of analytical techniques and efficient algorithms for financial protection.

**KeyWords:** Fraud Detection, Machine Learning, Pattern Analysis.