Detecção de transações fraudulentas utilizando diferentes algoritmos de aprendizado de máquina

1st Mabylly Kauany Neres da Silva
UTFPR
Apucarana, Brasil
mabylly@alunos.utfpr.edu.br

2nd Murilo Vital Rondina
UTFPR
Apucarana, Brasil
murilorondina@alunos.utfpr.edu.br

3rd Ruan Mateus Trizotti
UTFPR
Apucarana, Brasil
ruantrizotti@alunos.utfpr.edu.br

***Resumo*—Este estudo aborda a detecção de transações fraudulentas utilizando diferentes algoritmos de aprendizado de máquina, com foco em transações individuais. O aumento das fraudes digitais, especialmente no setor financeiro, tem gerado prejuízos significativos, como os R$ 3,5 bilhões registrados no Brasil em 2023. A pesquisa utiliza o dataset Credit Card Fraud Detection, composto por 284.807 transações, para explorar técnicas como Random Forest, Isolation Forest, XGBoost e Regressão Logística. O objetivo é identificar padrões de fraude e melhorar a precisão na classificação de transações fraudulentas. Os resultados mostram que o XGBoost obteve o melhor desempenho, com um F1-score de 0.84, indicando um bom equilíbrio entre precisão e recall. O estudo conclui que a combinação de diferentes algoritmos pode ser eficaz para a detecção de fraudes, especialmente em cenários com dados altamente desbalanceados.**

***Palavras-chaves*—Detecção de fraudes, Aprendizado de máquina, Transações financeiras, XGBoost, Regressão Logística, Random Forest, Desbalanceamento de classes**

I. Introdução

Fraudes em transações financeiras são um dos maiores problemas enfrentados por instituições financeiras e financeiras atualmente. Nos últimos anos, a fraude digital se tornou uma das principais ameaças aos mercados globais, com grande impacto no Brasil. Em 2023, foram registrados mais de 277 milhões de pedidos online, dos quais cerca de 3,7 milhões foram tentativas de fraude, o que causou um prejuízo financeiro estimado em R$ 3,5 bilhões [2]. Este cenário reflete o aumento dos ataques virtuais e a evolução das ferramentas usadas por criminosos para obter lucros ilegais, afetando setores críticos como o comércio eletrônico e o setor financeiro [8].

Com o crescimento exponencial das transações virtuais, impulsionado pela expansão do comércio eletrônico e pelo uso de plataformas bancárias online, o desenvolvimento de sistemas eficientes para identificar e prevenir fraudes se tornou uma prioridade.O progresso da Inteligência Artificial (IA) tem revolucionado a segurança digital, fornecendo ferramentas capazes de realizar análises preditivas e comportamentais com alta precisão e agilidade na identificação de padrões suspeitos [3]. Segundo Norvig e Russell [7], a IA tem se consolidado como uma tecnologia essencial para resolver problemas complexos, como a detecção de fraudes, graças à sua capacidade de aprender padrões em grandes volumes de dados.. Essas tecnologias têm se consolidado como um elemento fundamental na redução dos riscos em transações financeiras e no aprimoramento da experiência dos clientes.

Entretanto, a detecção de fraudes permanece um desafio devido à disparidade de classes nos dados, transações fraudulentas geralmente representam uma fração mínima do total e à necessidade de identificar padrões sutis e dinâmicos em grandes quantidades de dados. Os bancos, como o Nubank, têm adotado soluções baseadas em aprendizado de máquina para lidar com esses desafios. Ao lidar com transações sequenciais, muitas dessas organizações empregam estruturas específicas, tais como Redes Neurais Recorrentes (RNNs) ou *Memory Long Short-Term Memory*(LSTMs), que são adequadas para capturar relações temporais entre eventos [6]. No entanto, essas abordagens não são adequadas para situações em que as transações são analisadas isoladamente, sem levar em conta uma sequência temporal.

Este estudo explora métodos alternativos de aprendizado de máquina para detectar fraudes financeiras, focando em algoritmos que operam diretamente sobre transações individuais. Foram examinadas técnicas como *Random Forest, Gradient Boosting Machines* (GBM) e redes neurais profundas em diferentes situações. Obtemos informações detalhadas de transações reais, incluindo fraudes previamente identificadas [4].

O propósito deste estudo é colaborar para a criação de soluções mais robustas e abrangentes, alinhadas às demandas do mercado financeiro atual. Dessa forma, este estudo também examina os efeitos das fraudes digitais no mercado brasileiro, as tendências recentes e as soluções tecnológicas que têm mudado a abordagem desse problema. Ademais, espera-se que as informações apresentadas sirvam de guia para futuras pesquisas e aplicações em ambientes reais.

II. Trabalhos Relacionados

A detecção de fraudes financeiras é um dos desafios mais significativos no setor bancário e de pagamentos. Diversos trabalhos têm explorado o uso de algoritmos de aprendizado de máquina para identificar comportamentos anômalos em grandes volumes de transações, contribuindo para melhorar a segurança e reduzir prejuízos.

Giovane Piola Místico apresenta um estudo abrangente sobre o uso de aprendizado de máquina na detecção de

fraudes, destacando a aplicabilidade de algoritmos supervisionados como *Random Forests* e Redes Neurais. O trabalho enfatiza a importância de técnicas de engenharia de atributos para a construção de modelos robustos, abordando também os desafios relacionados ao desbalanceamento de classes no conjunto de dados. Os autores sugerem o uso de estratégias como *oversampling* para mitigar esse problema e melhorar a sensibilidade do modelo [5].

Em outro estudo, Andrei Camilo dos Santos explora diferentes abordagens de aprendizado supervisionado e não supervisionado, incluindo o uso de algoritmos como *Support Vector Machines (SVM) e Autoencoders* . O artigo destaca a aplicação de técnicas como detecção de *outliers* para identificar padrões não usuais em conjuntos de dados transacionais. A eficiência dos modelos é avaliada em termos de métricas como precisão e taxa de falsos positivos, evidenciando a eficácia de arquiteturas híbridas [1].

Adicionalmente, um artigo publicado na plataforma *Towards Data Science* apresenta uma visão geral das tendências e práticas recentes no uso de aprendizado de máquina para detecção de fraudes. O texto destaca a combinação de métodos baseados em árvores de decisão com técnicas de aprendizado profundo, ressaltando que tais combinações oferecem maior capacidade de generalização. Além disso, é discutida a aplicação de algoritmos como *Gradient Boosting* e redes neurais recorrentes (RNNs) em cenários de transações sequenciais, abordando casos em que a ordem temporal das transações fornece pistas críticas para a detecção de fraudes [9].

Esses estudos reforçam a relevância e a adaptabilidade dos algoritmos de aprendizado de máquina na detecção de transações fraudulentas, além de fornecerem diretrizes práticas para o desenvolvimento de sistemas mais eficientes e confiáveis.

III. Metodologia

Neste trabalho, foi desenvolvido um estudo para detecção de transações fraudulentas utilizando diferentes algoritmos de aprendizado de máquina. O estudo foi realizado com o dataset *Credit Card Fraud Detection*, amplamente utilizado em pesquisas acadêmicas, composto por 284.807 transações realizadas com cartões de crédito na Europa ao longo de dois dias. O dataset contém 31 colunas (*features*), onde as variáveis se resumem a: tempo da transação, valor (*amount*), classe (que indica se a transação é fraudulenta ou não) e 28 variáveis anonimizadas por meio de Análise de Componentes Principais *(PCA)* para proteger a privacidade dos dados. O objetivo é identificar padrões em dados financeiros e melhorar a precisão na classificação de fraudes. A linguagem escolhida para implementação foi *Python*, devido à ampla disponibilidade de bibliotecas e *frameworks* voltados para aprendizado de máquina e análise de dados, além de ser amplamente utilizada na comunidade científica e tecnológica.

A metodologia adotada neste trabalho segue um fluxo bem definido, conforme ilustrado no Fluxograma 1. Inicialmente, os dados são coletados do dataset *Credit Card Fraud Detection*, seguindo-se as etapas de Analise exploratória*(EDA)*, pré-processamento, seleção de algoritmos, treinamento dos modelos e avaliação dos resultados.

Fig. 1. Fluxograma da metodologia de pesquisa.

A. Pré-processamento

Para o pré-processamento dos dados, foram utilizadas as bibliotecas *pandas, numpy, seaborn* e *matplotlib*. Os dados foram explorados inicialmente, analisando sua distribuição, possíveis *outliers* e correlações entre variáveis. A detecção de desbalanceamento foi realizada e tratada nos modelos supervisados, ajustando os pesos das classes proporcionalmente à sua frequência. Além disso, foi utilizada a técnica de escalonamento das variáveis para adequá-las aos requisitos do algoritmo de Regressão Logística.

B. Seleção de Algoritmos

O estudo utilizou os algoritmos *Random Forest, Isolation Forest, XGBoost* e Regressão Logística. Cada um desses métodos foi selecionado considerando suas características e capacidades para lidar com o problema de detecção de fraudes, que apresenta dados altamente desbalanceados. O *Random Forest* é amplamente utilizado por sua robustez e capacidade de capturar interações não lineares entre variáveis, sendo combinado com *Isolation Forest*, que é um método não supervisionado eficaz para a detecção de anomalias em grandes conjuntos de dados. O *XGBoost* foi escolhido por ser um modelo de *gradiente boosting* de alto desempenho, eficiente em termos computacionais e capaz de lidar com dados desbalanceados. Já a Regressão Logística foi empregada como uma *baseline*, devido à sua simplicidade e eficiência em problemas lineares, mas exigiu o escalonamento das variáveis *Time* e *Amount*, a fim de garantir uma melhor performance do modelo.

C. Treinamento e Avaliação

Os modelos foram treinados utilizando uma divisão de 70% dos dados para treino e 30% para teste. A avaliação foi realizada com base em métricas como acurácia, precisão, recall e F1-score. O melhor resultado para cada modelo foi obtido utilizando os valores padrões dos hiperparâmetros dos algoritmos, com o acréscimo do parâmetro *class_weight='balanced'* quando aplicável. Essa abordagem permitiu lidar de forma eficiente com o desbalanceamento das classes, melhorando a sensibilidade do modelo na detecção de transações fraudulentas.

IV. Resultados

Nesta seção, são apresentados os resultados obtidos com os modelos de detecção de fraudes. Para avaliar o desempenho

dos algoritmos, foram utilizadas métricas como acurácia, precisão, recall e F1-score. Além disso, as matrizes de confusão de cada modelo são apresentadas para uma análise mais detalhada dos erros de classificação.

A. Métricas de Desempenho

A Tabela I resume as métricas de desempenho dos modelos testados, considerando apenas a classe fraudulenta (classe 1), uma vez que a classe não fraudulenta (classe 0) apresentou métricas próximas a 1 em todos os modelos, exceto na Regressão Logística, que obteve *recall* de 0.97 e *F1-score* de 0.99. O melhor resultado foi obtido com o *XGBoost*, que apresentou o maior *F1-score* (0.84), indicando um bom equilíbrio entre *precision* e *recall*.

TABLE I
MÉTRICAS DE DESEMPENHO DOS MODELOS.

Modelo	Acurácia	Precisão	Recall	F1-score
Random Forest	1.00	0.96	0.71	0.82
Random Forest + Isolation Forest	1.00	0.96	0.73	0.83
XGBoost	1.00	0.96	0.75	0.84
Regressão Logística	0.97	0.05	0.89	0.10

B. Matrizes de Confusão

As matrizes de confusão dos modelos são apresentadas nas Figuras 2, 3, 4 e 5. Essas matrizes permitem visualizar a quantidade de falsos positivos e falsos negativos, fornecendo insights sobre o desempenho de cada modelo na detecção de transações fraudulentas.

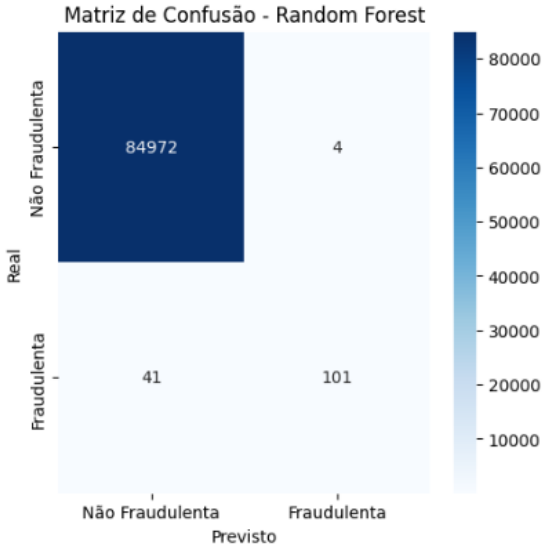

Fig. 2. Matriz de confusão do modelo Random Forest.

C. Análise dos Resultados

A seguir, são apresentadas as análises dos resultados de cada modelo, com base nas métricas e nas matrizes de confusão.

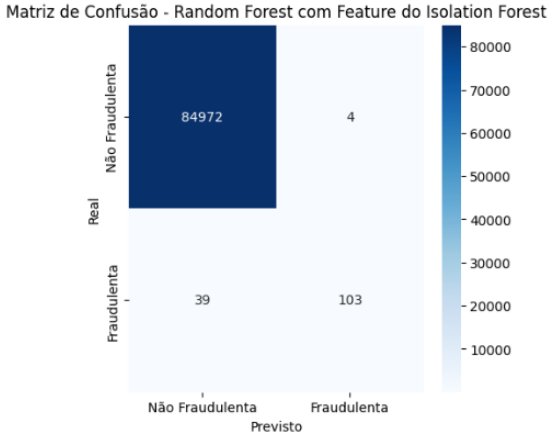

Fig. 3. Matriz de confusão do modelo Random Forest + Isolation Forest.

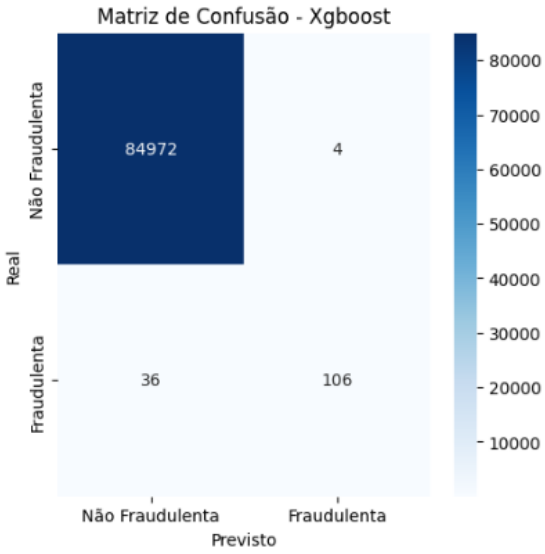

Fig. 4. Matriz de confusão do modelo XGBoost.

1) Random Forest: O modelo *Random Forest* apresentou excelente precisão (0.96) e uma acurácia de 1.00, indicando que ele é altamente eficaz na classificação geral das transações. No entanto, sua sensibilidade *(recall)* foi de 0.71, o que significa que ele deixou de identificar aproximadamente 29% das transações fraudulentas. Esse resultado sugere que, embora o modelo seja muito bom em evitar falsos positivos, ele pode não ser suficientemente agressivo na detecção de fraudes. A matriz de confusão mostra que o modelo tem um número significativo de falsos negativos, o que pode ser problemático em cenários onde a identificação de fraudes deve ser prioritária. O *F1-score* de 0.82 indica um equilíbrio razoável entre precisão e *recall*.

2) Random Forest + Isolation Forest: A combinação de *Random Forest* com *Isolation Forest* mostrou um leve aumento no *recall* (0.73 em comparação com 0.71 no *Random Forest* puro), indicando que o modelo conseguiu identificar um pouco mais de transações fraudulentas. Esse pequeno ganho na sensibilidade veio sem comprometer a precisão, que se

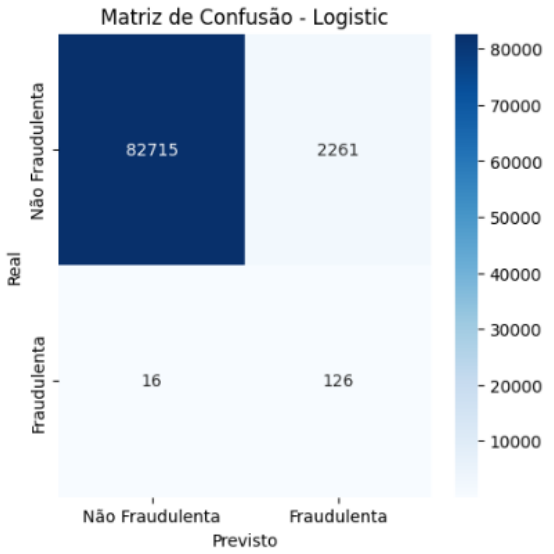

Fig. 5. Matriz de confusão do modelo Regressão Logística.

manteve em 0.96. A acurácia permaneceu em 1.00, o que sugere que o modelo ainda classifica muito bem a maioria das transações. O *F1-score* subiu ligeiramente para 0.83, o que indica que a adição do *Isolation Forest* trouxe uma melhoria no desempenho geral, reduzindo o número de falsos negativos sem aumentar significativamente os falsos positivos. A matriz de confusão reforça essa análise, mostrando uma redução marginal nos erros de classificação de fraudes.

3) XGBoost: A acurácia permanecendo em 1.00, sugere que o modelo classifica muito bem a maioria das transações ainda sim conseguindo também aumentar mais o *F1-score* subindo em 0.01 em relação ao *Isolation* o modelo conseguiu identificar um número maior de transações fraudulentas, reduzindo a quantidade de falsos negativos. Essa melhora na detecção de fraudes ocorreu sem comprometer significativamente a precisão do modelo, representando um equilíbrio ainda melhor entre precisão e *recall* em comparação com os modelos anteriores. A matriz de confusão reforça essa análise, mostrando uma redução nos erros de classificação de transações fraudulentas, o que pode ser um diferencial importante em um cenário onde a detecção de fraudes é uma prioridade. A acurácia permanecendo em 1.00, sugere que o modelo classifica muito bem a maioria das transações ainda sim conseguindo também aumentar mais o *F1-score* subindo em 0.01 em relação ao *Isolation Forest* indo para 0.84, sendo assim o modelo com maior taxa de confiabilidade entre todos os apresentados.

4) Regressão Logística:

V. Conclusão

References

[1] Andrei, Camilo dos Santos. (2023). Aplicações de aprendizado de máquina em cenários de detecção de fraudes financeiras. Disponível em: https://www.monografias.ufop.br/bitstream/35400000/6454/7/MONOGRAFIA-AprendizadoM%C3%A1quinaAplicado.pdf. Acesso em: 04 dez. 2024.

[2] CLEARSALE. Relatório de Identidade Digital e Fraudes 2023. Disponível em: https://br.clear.sale/mapa-da-fraude. Acesso em: 04 dez. 2024.

[3] ESTADO DE MINAS. Contra fraudes, bancos utilizam análise comportamental com IA. Disponível em: https://www.em.com.br/economia/2024/04/6842400-contra-fraudes-bancos-utilizam-analise-comportamental-com-ia.html. Acesso em: 04 dez. 2024.

[4] Kaggle. Credit Card Fraud Detection Dataset. Disponível em: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud Acesso em: 04 dez. 2024.

[5] Místico, Giovane Piola. (2023). Detecção de fraudes em transações financeiras com aprendizado de máquina. Disponível em: https://repositorio.usp.br/directbitstream/f41a5ae8-4c56-4db9-b274-d794576d72cc/Mistico-Giovane-tcc.pdf. Acesso em: 04 dez. 2024.

[6] Nubank. Behavior modeling with sequential architectures: Exploring LSTMs in machine learning. Publicado no blog Building Nubank, 25 ago. 2021. Disponível em: https://building.nubank.com.br/behavior-modeling-with-sequential-architectures-exploring-lstms-in-machine-learning/. Acesso em: 04 dez. 2024.

[7] RUSSELL, Stuart J.; NORVIG, Peter. Inteligência Artificial: Uma Abordagem Moderna. 4. ed. Rio de Janeiro: GEN LTC, 2022. E-book. p.57. ISBN 9788595159495. Disponível em: https://app.minhabiblioteca.com.br/reader/books/9788595159495/. Acesso em: 14 fev. 2025.

[8] TI INSIDE. Da detecção à prevenção: a revolução da IA no combate às fraudes do setor financeiro. Disponível em: https://tiinside.com.br/20/09/2024/da-deteccao-a-prevencao-a-revolucao-da-ia-no-combate-as-fraudes-do-setor-financeiro/. Acesso em: 04 dez. 2024.

[9] Towards Data Science (2023). Machine Learning in Fraud Detection: A Primer. Disponível em: https://towardsdatascience.com/machine-learning-in-fraud-detection-a-primer-8005b8c88cde?gi=1fbd6cc21063. Acesso em: 04 dez. 2024.