COMBATE À FRAUDE EM DADOS GOVERNAMENTAIS BRASILEIROS: USO DE MACHINE LEARNING NÃO SUPERVISIONADO PARA DADOS NÃO ROTULADOS

COMBATING FRAUD IN BRAZILIAN GOVERNMENT DATA: USING UNSUPERVISED MACHINE LEARNING FOR UNLABELED DATA

Lucas Gonzalez Martin Junior[1], [i]

RESUMO

No cenário atual, a fraude financeira representa uma preocupação global crescente, impulsionando a busca por métodos de detecção mais eficazes e adaptáveis. Este trabalho se insere nesse contexto, focando na aplicação de técnicas de aprendizado de máquina não supervisionadas para identificar padrões anômalos em transações financeiras governamentais, uma área onde a disponibilidade de dados rotulados de fraude é escassa. Reconhecendo a limitação dos métodos tradicionais e a necessidade de abordagens que operem sem rótulos pré-definidos, o objetivo principal deste estudo é desenvolver e avaliar a eficácia de algoritmos não supervisionados para sinalizar possíveis fraudes em um conjunto de dados transacionais públicos. A metodologia empregada envolve a análise de um conjunto de dados utilizando técnicas como *Z-score*, Intervalo Interquartil (IQR), detecção de transações duplicadas, análise de Lei de Benford e algoritmos de *Machine Learning* não supervisionados, incluindo *Isolation Forest*, *Local Outlier Factor* (LOF), DBSCAN e *One-Class SVM*, para gerar um *score* de risco de fraude. Os resultados preliminares indicam que, em um universo de 127.693 registros, 363 transações foram classificadas como suspeitas (níveis alto e crítico), representando aproximadamente 0,28% do total analisado, demonstrando a capacidade do sistema em identificar anomalias mesmo na ausência de rótulos.

Palavras-chave: detecção de fraude, não supervisionado, aprendizagem de máquina, KDD

ABSTRACT

In the current scenario, financial fraud represents a growing global concern, driving the search for more effective and adaptable detection methods. This work fits into this context, focusing on the application of unsupervised machine learning techniques to identify anomalous patterns in government financial transactions, an area where labeled fraud data is scarce. Recognizing the limitations of traditional methods and the need for approaches that operate without predefined labels, the main objective of this study is to develop and evaluate the effectiveness of unsupervised algorithms to flag potential fraud in a public transactional data set. The methodology employed involves analyzing a dataset using techniques such as Z-score, Interquartile Range (IQR), duplicate transaction detection, Benford's Law analysis, and unsupervised machine learning algorithms, including Isolation Forest, Local Outlier Factor (LOF), DBSCAN,

[1] Instrutor de formação profissional em Desenvolvimento de Sistemas na Escola e Faculdade SENAI Nadir Dias de Figueiredo e mestrando em Engenharia de Software na Universidade Tecnológica Federal do Paraná. E-mail: lcsgonzalez@hotmail.com

and One-Class SVM, to generate a fraud risk score. Preliminary results indicate that, out of a universe of 127,693 records, 363 transactions were classified as suspicious (high and critical levels), representing approximately 0.28% of the total analyzed, demonstrating the system's ability to identify anomalies even in the absence of labels.

Keywords: fraud detection, unsupervised, machine learning, KDD

1 INTRODUÇÃO

A fraude financeira representa uma preocupação significativa no cenário econômico global, ocasionando perdas substanciais para governos, organizações, setores corporativos e indivíduos (MANDAL et al., 2016). A crescente digitalização das transações financeiras, embora traga conveniência e eficiência, também expõe vulnerabilidades e impulsiona o surgimento de novas táticas fraudulentas (KARNAVOU et al., 2025). Nesse contexto, a detecção de fraude tornou-se um pilar fundamental para a segurança e a integridade do sistema financeiro, exigindo mecanismos robustos e adaptáveis.

1.1 Problema de pesquisa

Tradicionalmente, a detecção de fraude se apoia em métodos baseados em regras e abordagens supervisionadas de aprendizado de máquina (ML). No entanto, a evolução constante das táticas de fraude e a inerente dificuldade em obter dados de fraude rotulados em larga escala limitam a eficácia e a aplicabilidade desses métodos em cenários do mundo real (AL-HASHEDI; MAGALINGAM, 2021; MEDURI, 2024). A dependência de rótulos para treinamento de modelos supervisionados torna-os menos flexíveis diante de padrões de fraude emergentes e desconhecidos.

1.2 Objetivo(s)

O objetivo deste trabalho é desenvolver e avaliar a eficácia de técnicas de aprendizado de máquina não supervisionadas para sinalizar possíveis fraudes em um conjunto de dados transacionais do Governo Federal brasileiro, onde as instâncias de fraude não são previamente rotuladas.

1.3 Justificativa

Diante dessa lacuna, o aprendizado de máquina não supervisionado surge como uma alternativa promissora, capaz de identificar anomalias e padrões incomuns em dados sem a necessidade de conhecimento prévio sobre a natureza da fraude (MEDURI, 2024). Métodos baseados em *cluster*, como *K-means*, e técnicas de detecção de *outliers*, como *Isolation Forest* e *One-Class SVM*, já demonstraram eficácia e potencial para aplicação em tempo real em conjuntos de dados não rotulados (BECIROVIC et al., 2020; KARNAVOU et al., 2025). Este trabalho se alinha a essa perspectiva, buscando contribuir para o avanço da detecção de fraude em ambientes onde a rotulagem de dados é inviável.

2 REVISÃO DE LITERATURA

A fraude financeira representa uma preocupação significativa no cenário econômico global, ocasionando perdas substanciais para governos, organizações, setores corporativos e indivíduos. Pode ser definida como um ato de conduta ilícita ou antiética, que culmina em ganho indevido para um indivíduo ou

entidade, obtido por meios desonestos e ilegais (MANDAL et al., 2016). Elas podem ser divididas em quatro tipos: fraude bancária, fraude de seguros, fraude de demonstrações financeiras
e fraude de criptomoedas (AL-HASHEDI; MAGALINGAM, 2021). Uma das maneiras de identificar fraudes é utilizar técnicas de mineração de dados (MANDAL et al., 2016), que podem ser descritas, segundo Albashrawi e Lowell (2016), como uma abordagem usada para extrair dados significativos de um determinado conjunto de dados, usando uma ou mais abordagens, como técnicas estatísticas, de aprendizado de máquina, matemáticas ou de inteligência artificial.
Apesar da prevalência inicial de algoritmos supervisionados para detecção de fraudes, Meduri (2024) fornece uma análise abrangente das ameaças cibernéticas no setor bancário e destaca a importância da detecção de fraude não supervisionada como uma alternativa aos métodos tradicionais baseados em regras, que muitas vezes não conseguem acompanhar a evolução rápida das táticas de fraude.
Dentre os diversos algoritmos de ML existentes, métodos baseados em *cluster*, como *K-means* e detecção de *outliers* baseada em histograma, demonstraram eficácia em conjuntos de dados do mundo real, mostrando o potencial para aplicação em tempo real devido à sua velocidade (BECIROVIC et al., 2020). Complementarmente, o trabalho de Karnavou et al., (2025) explora a detecção de atividades ilícitas em transações bancárias de um banco grego, utilizando aprendizado não supervisionado com *Isolation Forest*, *One-Class SVM* e *autoencoders*, o que se alinha com a abordagem deste estudo de transações financeiras sem perfis fraudulentos rotulados. Ainda analisando a escolha dos algoritmos, Al-Hashedi e Magalingam (2021) oferecem uma revisão detalhada das técnicas de mineração de dados para detecção de fraude financeira, destacando a prevalência de métodos como SVM e *Random Forest*. Em adendo, Ali et al., (2022) realizaram uma revisão sistemática sobre a detecção de fraude financeira utilizando aprendizado de máquina, confirmando a proeminência de SVM e redes neurais. Por essa razão, este trabalho testou o algoritmo *One-Class SVM* com dois *kernel* diferentes.

3 METODOLOGIA

Os dados utilizados neste estudo foram retirados do Portal da Transparência, do Governo Federal Brasileiro, e são referentes aos Recursos Transferidos. Estes são recursos federais aplicados mediante repasse financeiro da União para estados, municípios ou até diretamente para entidades privadas sem fins lucrativos e outras instituições. O conjunto de dados utilizado está disponível publicamente online[2] e contém 127.698 registros, referentes a janeiro de 2025, apresentados em 36 colunas[3].
A Lei de Benford indica que a maioria dos números utilizados começa com o dígito 1 em vez do dígito 9, seguindo uma distribuição logarítmica na frequência de uso entre os dígitos 1 a 9, em números com mais de 4 dígitos (BENFORD, 1938) – caso do conjunto de dados estudado – portanto, a comparação com a Lei de Benford pode sinalizar possíveis desvios que possam indicar a existência de fraudes. De fato, após a análise, é constatado um desvio de 0.9% nos registros. A Figura 1 indica a taxa de

[2] A base pode ser acessada por meio do link: https://portaldatransparencia.gov.br/download-de-dados/transferencias
[3] Dicionário de dados pode ser acessado por meio do link: https://portaldatransparencia.gov.br/dicionario-de-dados/recursos-transferidos

desvio na frequência de cada dígito do conjunto de dados em comparação com o esperado segundo a Lei de Benford.

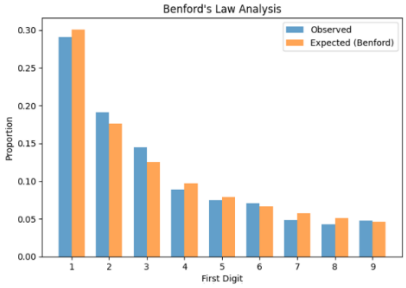

Figura 1. Diferença no conjunto de dados entre frequência esperada e encontrada, em cada dígito, de acordo com a Lei de Benford

Pelo fato de não haver fraudes marcadas para usar como referência, a primeira abordagem é identificar valores discrepantes através de estatística; essas podem indicar possíveis fraudes ou erros de lançamento. O primeiro método de detecção de anomalias utilizado foi *Z-score*, para identificar valores que se distanciam mais de 3 desvios-padrão da média, nesta análise encontrou-se 199 registros. O próximo método aplicado foi o Intervalo Interquartil (IQR), encontrando 20.096 registros discrepantes. Os dados marcados por essas anomalias receberam um peso que foi utilizado no cálculo de risco de fraude.

Dado que fraudes podem ocorrer por meio de lançamentos repetidos ou fracionamento de valores, que dificultam a detecção manual, a detecção de transações duplicadas auxilia nesta etapa. Testou-se tanto a duplicidade exata, de registros idênticos em todos os campos principais (Nome favorecido, Valor transferido, Ano/Mês e UF), encontrando 13.692 registros, quanto as quase duplicatas, transações para o mesmo favorecido e período, com valores próximos (diferença inferior a 1%), não encontrando nenhum registro. Os dados marcados pela anomalia de duplicidade também receberam um peso que foi utilizado no cálculo de risco de fraude.

Após isso, três fatores foram considerados como padrão incomum: (a) transações com números redondos (1.000, 10.000, 100.000, 1.000.000 e 10.000.000); (b) picos de valores muito acima do histórico de transferências; (c) concentração de favorecidos. A quantidade de transações detectadas com valores redondos foi de 9,73%, os dados marcados por esta anomalia receberam um peso que foi utilizado no cálculo de risco de fraude. Não foram encontrados picos de valores e a concentração de transferências recebidas pelos 10% maiores recebedores foi de 77%, esta análise pode ser cruzada com os registros marcados como potenciais fraudes, ao final, para potencializar a suspeita em transferências realizadas para estes recebedores.

Os principais algoritmos de ML para detecção de fraudes são supervisionados, entretanto, neste conjunto de dados não há fraudes marcadas para possibilitar a utilização destes algoritmos. Todavia, há candidatos que podem ser utilizados de forma não supervisionada, destes, foram testados cinco: (i) *Isolation Forest* (IF) (LIU et al., 2009), (ii) *Local Outlier Factor* (LOF)(BREUNIG et al., 2000), (iii) DBSCAN *Clustering* (RAM et al., 2010), (iv) *One-Class Support Vector Machine Radial Basis Function* (*One-Class SVM RBF*) e (v) *One-Class Support Vector Machine Linear* (*One-Class SVM Linear*)(AMER et al., 2013).

A performance dos algoritmos na detecção de possíveis anomalias está detalhada a seguir: os modelos que detectaram a maior taxa de anomalias, e portanto são menos confiáveis, foram *One-class SVM Linear* e IF que detectaram 80,41% e 30,51% dos registros como anomalia, respectivamente. Os modelos LOF e *One-class SVM RBF* encontraram taxas bastante semelhantes, 13,84% e 10,05%. Por fim, DBSCAN teve

a menor taxa, com apenas 0,11% dos registros marcados. Cada registro marcado em cada algoritmo de ML recebeu um peso, cumulativamente, que foi utilizado no cálculo de risco de fraude.
O *dataframe* tratado pelo pré-processamento tem uma coluna adicionada para o *score* de risco das transações, cuja inicia com valor zero para todas as transações. Após as análises estatísticas e de ML já descritas, cada anomalia identificada nos registros foi marcada e considerada para cálculo de risco posterior. Efetuados os cálculos, o *score* final também foi normalizado entre 0 (zero) e 1 (um).O *score* é combinado de forma ponderada para refletir múltiplos sinais de risco, prática recomendada em sistemas modernos de detecção de fraudes e a normalização e os pesos permitem calibrar a sensibilidade do sistema, reduzindo falsos positivos e negativos. Finalmente, as transações são categorizadas em níveis de risco: baixo, médio, alto e crítico.

4 RESULTADOS E DISCUSSÕES

A análise dos 127.693 registros resultou em um total de 363 transações classificadas como suspeitas, níveis alto e crítico, representando aproximadamente 0,28% do total analisado. Como não há rótulos (*ground truth*) indicando quais transações são realmente fraudulentas, a validação direta da assertividade dos métodos não é possível. O baixo percentual de transações suspeitas sugere que o sistema é conservador, priorizando a redução de falsos positivos. Entretanto, sem validação manual ou rótulos, não é possível estimar a taxa de acerto ou o número de fraudes não detectadas.

5 CONCLUSÃO

Este trabalho apresentou um *pipeline* robusto de detecção de fraudes em transações financeiras públicas, integrando métodos estatísticos, regras de negócio e algoritmos de ML não supervisionados. A abordagem permitiu identificar 363 transações suspeitas em um universo de mais de 127 mil registros, mesmo sem a presença de rótulos de fraude. A distribuição dos *scores* de risco e a priorização automática de casos críticos demonstram a utilidade prática do sistema como ferramenta de triagem para auditorias e investigações. Apesar das limitações inerentes à ausência de validação supervisionada, o sistema mostrou-se eficaz para identificar padrões atípicos e potenciais fraudes, alinhando-se às melhores práticas internacionais em análise forense de dados. Futuras evoluções podem ampliar a precisão e a aplicabilidade do método, especialmente com a incorporação de feedback humano, integração de dados contextuais e adoção de técnicas avançadas de aprendizado de máquina, como redes neurais. Concluindo, o trabalho contribui para o fortalecimento dos mecanismos de controle e transparência na gestão de recursos públicos, mas necessita de apoio humano e auditoria das transações identificadas.

REFERÊNCIAS

AL-HASHEDI, K. G.; MAGALINGAM, P. Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. **Computer Science Review**, v. 40, n. 2, p. 100402, maio 2021. ISSN 1574-0137. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1574013721000423>. Acesso em: 1 ago. 2025.

ALBASHRAWI, M. Detecting financial fraud using data mining techniques: A decade review from 2004 to 2015. **Journal of Data Science**, v. 14, n. 3. p. 553–570, jul. 2016.

ALI, A. et al. Financial fraud detection based on machine learning: A systematic literature review. **Applied Sciences**, v. 12, n. 19. p. 9637, set. 2022.

AMER, M.; GOLDSTEIN, M.; ABDENNADHER, S. Enhancing one-class support vector machines for unsupervised anomaly detection. **Proceedings of the ACM SIGKDD Workshop on Outlier Detection and Description - ODD '13***, p. 8–15, ago.* 2013.

ARIYANTO, R.; BONE, H. Fraud awareness in indonesian governmental sector: Multi-agency responses. **Review of Integrative Business and Economics Research**, v. 9, n. Supplementary Issue 2, p. 209–222, jun. 2020.

BECIROVIC, S.; ZUNIC, E.; DONKO, D. A case study of cluster-based and histogram-based multivariate anomaly detection approach in general ledgers. **2020 19th International Symposium INFOTEH-JAHORINA (INFOTEH**), East Sarajevo, Bosnia and Herzegovina, 2020, pp. 1-6, doi: 10.1109/INFOTEH48170.2020.9066333 Mar. 2020. Disponível em: <https://ieeexplore.ieee.org/document/9066333>. Acesso em: 1 ago. 2025.

BENFORD, F. The law of anomalous numbers. Proceedings of the American Philosophical Society, **American Philosophical Society**, v. 78, n. 4, p. 551–572, mar. 1938. ISSN 0003049X. Disponível em: <http://www.jstor.org/stable/984802>. Acesso em: 1 ago. 2025.

BREUNIG, M. et al. Lof: Identifying density-based local outliers. **Proceedings of the 2000 ACM SIGMOD international conference on Management of data - SIGMOD '00***,* v. 29, n.3, p. 93–104, jun. 2000.

KARNAVOU, E. et al. I know you're a fraud: Uncovering illicit activity in a greek bank transactions with unsupervised learning. **Expert Systems with Applications**, v. 288, p. 128148, mai. 2025. ISSN 0957-4174. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0957417425017683>. Acesso em: 1 ago. 2025.

LIU, F. T.; TING, K.; ZHOU, Z.-H. Isolation forest. **2008 Eighth IEEE International Conference on Data Mining***,* p. 413 – 422, jan. 2009.

MANDAL, P. et al. A complete literature review on financial fraud

detection applying
data mining techniques. **International Journal of Trust Management in Computing and Communications**, v. 3, p. 336, jan. 2016.

MEDURI, K. Cybersecurity threats in banking: Unsupervised fraud detection analysis. **International Journal of Science and Research Archive**, v. 11, p. 915–925, abr. 2024.

RAM, A. et al. A density based algorithm for discovering density varied clusters in large spatial databases. **International Journal of Computer Applications**, v. 3, jun. 2010.

AGRADECIMENTOS
Agradeço aos meus pais que sempre colaboraram com a minha educação, e à UTFPR e ao Prof. Dr. Marlon Marcon, que possibilitaram o desenvolvimento desta pesquisa.

SOBRE O(S)AUTOR(ES)

[i] **LUCAS GONZALEZ MARTIN JUNIOR**

Possui graduação em Sistemas de Informação e MBA em Bussines Analytics e Data Science pela Faculdade de Informática e Admininstração Paulista – FIAP, Especialização em Tecnologia Java e cursando atualmente o Mestrado em Engenharia de Software pela Universidade Tecnológica Federal do Paraná - UTFPR. Tem experiência nas áreas de Desenvolvimento de Sistemas, Business Intelligence, Projetos de Redes e Data Science. É Instrutor de Formação Profissional na Escola e Faculdade SENAI “Nadir Dias de Figueiredo” em Osasco e Professor de Graduação na Faculdade de Informática e Admininstração Paulista – FIAP.