UNIVERSIDADE DE BRASÍLIA
Faculdade do Gama

Aprendizado de Máquina

Mini Trabalho 1

Definição do problema e contextualização

Grupo - 1

Ana Clara Barbosa Borges
André Emanuel Bispo da Silva
Artur Handow Krauspenhar
Gabriel Moura dos Santos
João Artur Leles Ferreira Pinheiro
João Pedro Anacleto Ferreira Machado

Brasília, DF
2025

Problemas reais e como o uso de Machine Learning pode resolver

O conceito de Aprendizado de Máquina (ou Machine Learning, ML) foi apresentado como um subconjunto da Inteligência Artificial e revolucionou a forma como os computadores processam informações, possibilitando que eles aprendessem a partir de dados sem terem sido de fato programados (NEHA, et al., 2024).

As mais diversas técnicas de ML têm se mostrado promissoras em diferentes áreas, como reconhecimento de imagens, processamento de linguagem natural, finanças e automação de sistemas (NEHA, et al., 2024).

O livro "Machine Learning" de Tom Mitchell (Mitchell, T. M., 1997) apresenta algumas tarefas de Aprendizado de Máquina a fim de demonstrar como o conceito pode ser aplicado em diferentes contextos. Dentre eles, seguem alguns:

- Problema de aprendizagem de reconhecimento de caligrafia, em que um dataset de fotos de diferentes caligrafias seria utilizado para o treinamento do modelo.
- Problema de aprendizagem de damas, em que ao jogar diversas formas contra si mesmo, o modelo conseguiria aprimorar o seu comportamento.

Os exemplos citados acima são alguns de uma infinidade de possíveis aplicações de Machine Learning. Após o estudo dos tipos de aprendizado de máquina e de possíveis aplicações, escolhemos aplicar conceitos de ML para o contexto de detecção de fraudes, motivados por pesquisas anteriores que serão melhor abordadas ao longo do documento.

Objetivo do projeto

A fraude em transações financeiras tornou-se um problema global, resultando na perda de bilhões de dólares todos os anos. De acordo com SOROURNEJAD et al. (2016) apenas nos EUA no ano de 2011 as fraudes acarretaram em um prejuízo de 3,4 bilhões de dólares. Detectar essas fraudes é um desafio complexo, pois elas representam uma parcela muito pequena do enorme volume de transações legítimas. Além disso, diversas estratégias de fraude são desenvolvidas constantemente, o que acaba por exigir soluções adaptáveis.

Diante desse cenário, é importante o desenvolvimento de técnicas que permitam a detecção de fraudes. O objetivo do projeto é desenvolver um modelo de detecção de fraudes utilizando técnicas de Machine Learning sobre os dados fornecidos no dataset, identificando assim comportamentos atípicos e suspeitos entre milhares de transações legítimas, de forma a

encontrar padrões que possibilitem identificar fraudes. Visando assim minimizar perdas financeiras e possibilitar a identificação de tentativas de fraude em tempo hábil, algo fundamental para instituições financeiras.

Soluções e métodos existentes

Sistemas de detecção de fraudes baseados em IA são existentes em cerca de 73% dos bancos brasileiros, de acordo com um relatório da Febraban (2024). Esses sistemas geralmente fazem uso de inteligência artificial a fim de se adaptarem às estratégias de fraude, que evoluem com o decorrer do tempo, e identificar padrões de transações ilícitas ou transações fora do padrão, uma técnica chamada de detecção de anomalias (também chamada de detecção de *Outliers*).

Um ponto de foco importante é a falta de algoritmos de aprendizado não supervisionado, que são importantes pois técnicas de fraude e ataques são atualizados e criados constantemente e, assim como outras áreas de cibersegurança, os sistemas devem estar constantemente atualizados. (AHMED; MAHMOOD; ISLAM, 2016).

Soluções existentes de aprendizado de máquina também são muito dependentes de geração de dados sintéticos, por motivos de privacidade e competitividade de dados bancários. Isso apresenta certas dificuldades(AHMED; MAHMOOD; ISLAM, 2016):

- geração de números realmente aleatórios
- dados devem ser atualizados constantemente devido à evolução das técnicas de fraude
- dados sintéticos são específicos a domínio e podem levar a falsos positivos em outros domínios
- Modelos são altamente dependentes da configuração inicial.

Isto também destaca a importância de métodos não supervisionados, que não necessitam de dados sintéticos para atuação. Além disso, há um desbalanceamento de classe, ou seja, existem muitos mais dados sobre transações não fraudulentas do que sobre transações fraudulentas (STOJANOVIĆ et al., 2021).

Além de deteção de de outliers, métodos existentes para classificação de fraude em dados incluem métodos genéricos de *machine learning*, métodos de conjunto (que utilizam uma combinação de modelos a fim de criar um modelo ótimo e eficiente), *deep learning* e métodos baseados em grafos(STOJANOVIĆ et al., 2021). E, de acordo com uma pesquisa de STOJANOVIĆ et al. (2021), técnicas de deteção de anomalias baseadas em

combinação tiveram uma performance mais robusta que outros métodos de detecção de outliers.

Hipóteses do projeto

Considerando o objetivo de desenvolver um modelo de detecção de fraudes utilizando técnicas de Machine Learning e a análise do cenário atual com suas limitações, levantamos um conjunto de hipóteses que guiarão o desenvolvimento e avaliação do projeto. Acreditamos que a combinação de aprendizado supervisionado, técnicas de detecção de anomalias e engenharia de features resultará em um modelo eficaz e robusto na identificação de transações fraudulentas.

As hipóteses centrais incluem a capacidade de modelos supervisionados, como Random Forests ou Redes Neurais, de aprender padrões complexos associados à fraude a partir de dados históricos rotulados.

Ahmed, Mahmood e Islam (2016) demonstraram a eficácia de modelos supervisionados na detecção de fraudes em finanças. Além disso, prevemos que a utilização conjunta de métodos de detecção de anomalias para identificar transações incomuns e algoritmos supervisionados para a classificação final melhora significativamente o desempenho global do sistema de detecção.

Stojanović et al. (2021) enfatizam que técnicas de detecção de anomalias baseadas em combinação apresentam uma performance mais robusta que outros métodos. É também nossa premissa que o tratamento adequado do desbalanceamento de classes nos dados de treinamento e a seleção cuidadosa das features são fatores críticos para o sucesso do modelo.

Por fim, embora a abordagem principal seja o aprendizado supervisionado, exploramos a possibilidade de que modelos de aprendizado não supervisionado possam identificar novos padrões de fraude ainda não representados nos dados rotulados, complementando a capacidade de detecção do sistema

Aprendizado supervisionado e não supervisionado

O aprendizado de máquina pode ser dividido em duas categorias principais: supervisionado e não supervisionado. No aprendizado supervisionado, o algoritmo é treinado com um conjunto de dados rotulados, ou seja, cada entrada é acompanhada de uma saída esperada, permitindo que o modelo aprenda a mapear entradas para saídas corretamente. Já no

aprendizado não supervisionado, os dados não possuem rótulos, e o objetivo é identificar padrões, agrupamentos ou estruturas ocultas nos dados. Segundo Alpaydin(2020), no aprendizado supervisionado, o sistema é apresentado com exemplos de entrada e saída, enquanto no não supervisionado, ele deve encontrar estrutura por conta própria nos dados de entrada. Essa distinção é fundamental para a escolha da abordagem apropriada em cada tipo de problema.

A escolha do aprendizado supervisionado se justifica plenamente no contexto de uma base de dados de transações financeiras rotulada quanto à ocorrência ou não de fraude. Como os dados já indicam quais transações são fraudulentas, o modelo supervisionado pode utilizar essa informação para aprender uma função de decisão precisa. Esse tipo de modelo é ideal quando se tem rótulos confiáveis, pois permite a modelagem direta de problemas bem definidos, como a detecção de fraudes.

Critérios de avaliação

Alguns dos critérios de avaliação que podem ser utilizados para medir o desempenho do modelo incluem: **Precisão**, que indica a proporção de acertos entre todas as previsões realizadas;**Taxa de Falsos Positivos e Falsos Negativos**, que revela os erros na classificação incorreta das categorias; **AUC-ROC**, que avalia a capacidade do modelo de distinguir corretamente entre as classes, além de **Sensibilidade** e **Especificidade**, que medem, respectivamente, a capacidade de identificar corretamente os casos positivos e negativos.

Referências Bibliográficas

FEBRABAN. *Pesquisa FEBRABAN de tecnologia bancária 2024.* São Paulo: FEBRABAN, 2024. Disponível em: https://cmsarquivos.febraban.org.br/Arquivos/documentos/PDF/Pesquisa%20Febraban%20de%20Tecnologia%20Banc%C3%A1ria%202024.pdf. Acesso em: 5 abr. 2025.

AHMED, M.; MAHMOOD, A. N.; ISLAM, M. R. A survey of anomaly detection techniques in financial domain. **Future Generation Computer Systems**, v. 55, p. 278–288, 2016.

ALPAYDIN, Ethem. *Introduction to Machine Learning*. 4. ed. Cambridge: MIT Press, 2020.

STOJANOVIĆ, B. et al. Follow the trail: Machine Learning for fraud detection in Fintech applications. **Sensors (Basel)**, v. 21, p. 1594, fev. 2021.

Mitchell, T. M. (1997). Machine Learning. McGraw-Hill.

NEHA, S. V. S. T.; YADAV, Yogesh; GOYAL, Yashika. Introduction to Machine Learning. *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, mar. 2024.

SOROURNEJAD, Samaneh; ZOJAJI, Zahra; EBRAHIMI ATANI, Reza; MONADJEMI, Amir Hassan. *A Survey of Credit Card Fraud Detection Techniques: Data and Technique Oriented Perspective*. [S.l.], 19 nov. 2016. Disponível em: https://arxiv.org/abs/1611.06439. Acesso em: 6 abr. 2025.