Prevenção de fraudes com cartões de crédito utilizando o classificador ingênuo de Bayes.

1st Gabriel Pierre Carvalho Coelho
Centro de informática
Universidade Federal de Pernambuco
Recife, Brasil
gpcc@cin.ufpe.br

2nd Klarissa Andrade de Morais
Centro de informática
Universidade Federal de Pernambuco
Recife, Brasil
kam@cin.ufpe.br

3rd Luan Eustáquio Lopes de Farias
Centro de informática
Universidade Federal de Pernambuco
Recife, Brasil
lelf@cin.ufpe.br

***Abstract* — Este trabalho busca através do estudo de uma base de dados, prever se uma transação com cartão de crédito é fraudulenta ou não. A proposta consiste em procurar e utilizar uma base de dados contendo o histórico de transações com cartões de crédito, integrando exemplos positivos e negativos(i.e., fraudulentas e legítimas), e utilizar esses dados para treinar o modelo. O algoritmo Ingênuo de Bayes será utilizado para classificar novas transações com base nas características das transações fraudulentas e legítimas já registrados na base de dados.**

Palavras-chave* — *Detecção de fraudes, transações com cartão de crédito, Classificador Probabilístico, Teorema de Bayes, Python.

I. INTRODUÇÃO

Nos últimos anos, o uso de cartões de crédito em transações financeiras tem apresentado um crescimento significativo, impulsionado pelo aumento da popularidade do comércio eletrônico e das transações sem contato físico. No entanto, esse aumento também tem gerado um grande desafio para as empresas e instituições financeiras: a prevenção de fraudes em transações de cartões de crédito. As fraudes nesse tipo de transação representam um problema crescente, causando prejuízos financeiros significativos para as empresas e consumidores, além de abalar a confiança no uso desse meio de pagamento.

Nesse cenário, o uso de técnicas de aprendizado de máquina tem se apresentado como uma solução eficiente para a prevenção de fraudes em transações de cartões de crédito. Diante disso, este trabalho propõe uma solução baseada em algoritmos de aprendizado de máquina, utilizando o algoritmo Ingênuo de Bayes da biblioteca Scikit-Learn, além de testes com algumas variações do algoritmo da mesma biblioteca, a fim de definir o que se encaixa melhor no modelo. A pretensão é construir um modelo capaz de analisar transações de cartões de crédito e classificá-las como fraudulentas ou legítimas com base em uma análise probabilística.

II. OBJETIVO

O objetivo deste trabalho é propor uma solução para prevenção de fraudes em transações de cartões de crédito com base em aprendizado de máquina, em particular, o algoritmo ingênuo de Bayes. A ideia é construir modelos capazes de analisar transações de cartões de crédito e classificá-las como fraudulentas ou não com base em uma análise probabilística. Para isso, serão utilizadas técnicas de pré-processamento de dados, como normalização e seleção de características, e os algoritmos serão treinados e avaliados em conjuntos de dados públicos de transações de cartões de crédito. Além disso, o trabalho propõe uma comparação entre o desempenho dos algoritmos desenvolvidos, a fim de verificar qual deles apresenta melhores resultados nos cenários avaliados.

III. JUSTIFICATIVA

A fraude em transações com cartões de crédito é um problema crescente que causa prejuízos financeiros significativos para empresas e consumidores, além de abalar a confiança no uso

desse meio de pagamento. A utilização de modelos preditivos para detecção de fraudes em transações com cartão de crédito é uma abordagem promissora para prevenir e reduzir o impacto dessas fraudes. Diversos trabalhos na área de aprendizado de máquina têm sido realizados com o objetivo de detectar operações suspeitas, como o trabalho de Ali et al. [1], que propõe uma nova abordagem utilizando redes neurais convolucionais.

Um exemplo recente que destaca a importância da prevenção de fraudes em transações com cartões de crédito é o caso da Netshoes. A empresa foi condenada a pagar uma multa de R$ 500 mil por um vazamento de dados de mais de 2 milhões de clientes [2]. Esse tipo de violação de dados pode levar a prejuízos financeiros e de reputação para a empresa, além de colocar em risco a privacidade e segurança dos clientes. O desenvolvimento de um modelo preditivo eficaz pode contribuir para a melhoria da segurança dos sistemas de pagamento eletrônico e prevenção de perdas financeiras. Portanto, este projeto justifica-se pela importância da detecção de fraudes em transações com cartão de crédito e pela necessidade de desenvolver modelos preditivos eficazes para prevenir e reduzir o impacto dessas fraudes.

IV. CLASSIFICADOR INGÊNUO DE BAYES

Para prosseguirmos na análise dos dados, será empregado o algoritmo classificador de Bayes, o qual foi desenvolvido pelo matemático inglês Thomas Bayes de mesmo nome.

$$P(y|x) = \frac{P(x|y)P(y)}{P(x)} : \quad \text{Bayes theorem,}$$

Figura 1. Teorema de Bayes.

O método se baseia na independência entre os fatores, por isso é conhecido como “ingênuo” e descreve com boa precisão a classificação dos mesmos , essa precisão, e poder empregar um número menor de dados e mesmo assim produzir um resultado preciso é o motivo a qual ele é empregado em machine learning e por consequência o utilizaremos no projeto.

Na equação da figura 1, P(x) e P(y) representam a probabilidade de um evento “x” e outro “y” ocorrerem, enquanto a notação P(y|x) simboliza a probabilidade do evento y ocorrer após ter ocorrido um evento x.

V. METODOLOGIA

A. Dataset

Coleta de dados: Para este projeto, será utilizado o conjunto de dados *Credit Card Fraud Detection*, disponibilizado publicamente, que contém informações de transações de cartões de crédito realizadas por titulares europeus em setembro de 2013. O conjunto de dados apresenta transações que ocorreram em dois dias, totalizando 284.807 transações, dentre as quais 492 foram rotuladas como fraudulentas e as demais como legítimas. A base de dados é altamente desbalanceada, com a classe positiva (fraudes) representando apenas 0,172% do total de transações. O conjunto de dados contém apenas variáveis numéricas resultantes de uma transformação PCA (Principal Component Analysis). As variáveis V1 a V28 são as componentes principais obtidas com o PCA, enquanto as variáveis *"Time"* e *"Amount"* não foram transformadas. A variável *"Time"* representa o tempo em segundos entre a transação atual e a primeira transação registrada no conjunto de dados, enquanto *"Amount"* representa o valor da transação. A variável *"Class"* é a variável de resposta e assume o valor 1 em caso de fraude e 0 em caso contrário.

Pré-processamento de dados: Para lidar com as particularidades do conjunto de dados *“Credit Card Fraud Detection”*, altamente desbalanceado, serão utilizadas técnicas de amostragem, como subamostragem ou sobreamostragem, para equilibrar as classes e evitar a dominância da classe negativa no conjunto de dados. O objetivo é selecionar as características mais relevantes e tratar dados ausentes ou duplicados de forma adequada, avaliando o impacto das técnicas de pré-processamento no desempenho dos modelos de detecção de fraudes.

Divisão dos dados: Os dados serão divididos em conjuntos de treinamento, validação e teste, em uma proporção de 60%, 20% e 20%, respectivamente.

B. Tecnologias e desenvolvimento

Ferramentas utilizadas: Para este projeto, utilizaremos da linguagem de programação *Python*, que tem se mostrado útil para trabalhos envolvendo computação científica. O passo a passo dos cálculos do classificador, bem como o das probabilidades serão implementados através de recursos da linguagem e de suas bibliotecas, como a *Numpy, Pandas, Scikit Learn e Matplotlib*. Sendo a última a fornecedora de artifícios para a exibição desses dados através de conjuntos de métodos para a geração dos resultados

graficamente. Essas bibliotecas nos fornecerão funções para tratamento de números, manipulação dos dados e aprendizagem de máquina.

Treinamento dos modelos: Serão treinados modelos de detecção, utilizando variações do algoritmo ingênuo de Bayes disponíveis na biblioteca Scikit-Learn.

Ambiente de desenvolvimento: O projeto será desenvolvido no ambiente do Notebook Jupyter, onde realizaremos a escrita do código e análise dos resultados obtidos. O Google Collaboratory também será utilizado como ferramenta adicional na criação e desenvolvimento do projeto.

*Algumas bibliotecas não citadas podem se mostrar necessárias durante o desenvolvimento do projeto e serão adicionadas no relatório final.

C. Avaliação dos modelos e análises dos resultados

Avaliação dos modelos: Com base no conjunto de dados altamente desequilibrado, avaliaremos o desempenho dos modelos de detecção de fraudes utilizando a métrica *AUPRC*, além de outras métricas como *precisão, recall e F1-scor*e. É importante também avaliar a relevância e a importância das características utilizadas no modelo, considerando que as características originais das transações não estão disponíveis devido a questões de confidencialidade.

Comparação de desempenho: Será realizada uma comparação de desempenho entre os modelos desenvolvidos, a fim de verificar qual deles apresenta melhores resultados na detecção de fraudes em transações de cartões de crédito.

Análise dos resultados: O processo de análise será feito utilizando o método k-fold para validação do algoritmo. Com o k-fold, o conjunto de dados será dividido em k-subconjuntos. Em seguida, o algoritmo será treinado e testado k-vezes, utilizando um subconjunto diferente como conjunto de validação em cada rodada e os demais como conjunto de treinamento e teste. Ao final das k-rodadas, os resultados serão combinados para obter as estatísticas de desempenho do modelo. Dessa forma, teremos resultados mais precisos e confiáveis para avaliar a efetividade dos modelos propostos na detecção.

VI. CRONOGRAMA DE ATIVIDADES

Data	**Objetivo**
05/03/23	Reunião com o grupo para escolha do tema da pesquisa.
06/03/23 - 08/03/23	Pesquisa sobre o assunto, coleta de dados e levantamento de informações relevantes.
10/03/23	Finalização da proposta de pesquisa e entrega.
11/03/23 - 15/03/23	Coleta e pré-processamento dos dados, normalização e seleção de características.
16/03/23 - 20/03/23	Análise de distribuição e relações entre as características das transações, criação de gráficos e visualizações.
21/03/23 - 25/03/23	Criação dos modelos de aprendizagem de máquina e ajuste dos modelos com base nas análises dos dados.
26/03/23 - 01/04/23	Treinamento dos modelos com os dados.
02/04/23 - 06/04/23	Verificação do desempenho dos modelos com dados de validação e avaliação das métricas de desempenho.
07/04/23 - 11/04/23	Análise dos resultados obtidos com os modelos treinados e comparação dos desempenhos entre os algoritmos.
12/04/23 - 16/04/23	Elaboração da conclusão e preparação do relatório final.

17/04/23 - 22/04/23	Revisão e finalização do relatório final e criação dos slides para a apresentação.
24/04/23	Apresentação final.

VII. REFERÊNCIAS

[1] ALI, B. et al. Detecção de fraude em cartão de crédito: uma nova abordagem usando redes neurais convolucionais. IEEE Access, v. 6, p. 36647-36656, 2018. Disponível em: https://ieeexplore.ieee.org/abstract/document/8397371

[2] G1. Netshoes terá de pagar R$ 500 mil por vazamento de dados de 2 milhões de clientes. Disponível em: https://g1.globo.com/df/distrito-federal/noticia/2019/02/05/netshoes-tera-de-pagar-r-500-mil-por-vazamento-de-dados-de-2-milhoes-de-clientes.ghtml

[3] ARAGAW, S. et al. Detecção de fraudes em cartões de crédito usando redes neurais artificiais: uma revisão sistemática. IEEE Access, v. 7, p. 10557-10567, 2019. Disponível em: https://ieeexplore.ieee.org/abstract/document/8730573

[4] PEDREGOSA, F. et al. Scikit-learn: machine learning in Python. Journal of Machine Learning Research, v. 12, p. 2825-2830, 2011. Disponível em: https://jmlr.org/papers/v12/pedregosa11a.html

[5] GÉRON, A. Mãos à obra: aprendizado de máquina com Scikit-Learn e TensorFlow: conceitos, ferramentas e técnicas para construir sistemas inteligentes. Rio de Janeiro: Alta Books, 2018. Disponível em: https://www.casadocodigo.com.br/products/livro-machine-learning-python