Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Biológicas
Departamento de Estatística
Bacharelado em Estatística

Aprendizado de Máquina Aplicado na Detecção de Fraudes em Cartão de Crédito

Andrei Camilo dos Santos

Ouro Preto-MG
Setembro 2023

Andrei Camilo dos Santos

Aprendizagem de Máquina Aplicado na Detecção de Fraudes em Cartão de Crédito

Monografia de Graduação apresentada ao Departamento de Estatística do Instituto de Ciências Exatas e Biológicas da Universidade Federal de Ouro Preto como requisito parcial para a obtenção do grau de bacharel em Estatística.

Orientador(a)

Dr. Tiago Martins Pereira

Universidade Federal de Ouro Preto – UFOP
Departamento de Estatística – DEEST

Ouro Preto-MG

1 Setembro de 2023

MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
INSTITUTO DE CIENCIAS EXATAS E BIOLOGICAS
COLEGIADO DO CURSO DE ESTATISTICA

UFOP

FOLHA DE APROVAÇÃO

Andrei Camilo dos Santos

Aprendizagem de máquina aplicado na detecção de fraudes em cartão de crédito

Monografia apresentada ao Curso de Estatística da Universidade Federal
de Ouro Preto como requisito parcial para obtenção do título de Bacharel em Estatística

Aprovada em 1º de setembro de 2023

Membros da banca

Dr. Tiago Martins Pereira - Orientador (Universidade Federal de Ouro Preto)
Dr. Fernando Luiz Pereira (Universidade Federal de Ouro Preto)
Dra. Diana Campos de Oliveira (Universidade Federal de Ouro Preto)

Prof. Dr. Tiago Martins Pereira, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 01/09/2023

Documento assinado eletronicamente por **Tiago Martins Pereira**, **PROFESSOR DE MAGISTERIO SUPERIOR**, em 10/02/2024, às 19:10, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.

A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0 , informando o código verificador **0585868** e o código CRC **90CCFE16**.

R. Diogo de Vasconcelos, 122, - Bairro Pilar Ouro Preto/MG, CEP 35402-163
Telefone: (31)3559-1312 - www.ufop.br

Agradecimentos

Dedico esta monografia à minha família e amigos, cujo amor, apoio e encorajamento me guiaram ao longo dessa jornada acadêmica. Vocês são a minha força, e estou além de agradecido por ter vocês ao meu lado. O sucesso deste trabalho é, em grande parte, devido à minha dedicação e apoio contínuo de todos vocês.

A minha irmã, agradeço por estar ao meu lado, compartilhando risos, ajudando com suas palavras encorajadoras e lembrando-me da importância de equilibrar o trabalho, com momentos de descontração. Seu apoio me motivou a continuar perseverando, mesmo quando as coisas ficaram difíceis.

Aos amigos que Ouro Preto me deu, sou grato por terem me apoiado durante todas as etapas desta jornada acadêmica. Vocês foram essenciais, ouvindo minhas preocupações, oferecendo conselhos valiosos e permanecendo ao meu lado nos momentos de dúvida.

Além disso, gostaria de expressar minha gratidão aos meus professores e orientadores, cujo conhecimento e orientação foram cruciais para o sucesso deste trabalho. Suas contribuições valiosas, paciência e incentivo foram fundamentais para minha compreensão aprofundada do tema e para o desenvolvimento das habilidades necessárias para concluir esta monografia.

Muito obrigado!

Aprendizagem de Máquina Aplicado na Detecção de Fraudes em Cartão de Crédito

Autor: Andrei Camilo dos Santos
Orientador(a): Dr. Tiago Martins Pereira

Resumo

O presente trabalho descreve uma proposta de aplicar diferentes técnicas de aprendizagem de máquina, estatísticas de teste e estudo de caso de transações fraudulentas por meio de cartões de credito. Hoje no mundo sendo a mais popular forma de pagamento, tanto em compras online ou mesmo offline o número crescente de fraudes impacta negativamente a receita de várias instituições financeiras além das pessoas que são vítimas de golpes. Desta maneira uma detecção categórica e eficiente se torna de grande valia no mantimento da confiança no sistema de pagamento. Desse modo metodologias que utilizam de aprendizagem de máquina e inteligência artificial veem sendo desenvolvidas com intuito de melhorar a segurança das instituições. Nesse contexto existem dois principais tipos de algoritmos, um que utiliza de aprendizagem supervisionada, tais como Naive Bayes, regressão logística, arvores de decisão, KNN e redes neurais, e outros conhecidos como aprendizagem não supervisionada como por exemplo, K – means, Clustering e hierárquico. Embora muitas técnicas propostas tenham atingido resultados satisfatórios, hoje ainda é um grande desafio detectar com alta precisão transações consideradas fraudulentas (falso positivo) em relação a transações normais (falso negativo), visto que os algoritmos usuais aprendem com os dados e estes geralmente estão desbalanceados devido a grande maioria das transações serem legitimas.

Palavras-chave: Aprendizagem de Máquina, Aprendizagem Supervisionada, Inteligência Artificial.

Machine Learning Applied to the Detection of Credit Card Fraud)

Author: Andrei Camilo dos Santoso
Advisor: Dr. Tiago Martins Pereira

Abstract

The present work describes a proposal to apply different machine learning techniques, test statistics and a case study of fraudulent transactions through credit cards. Today in the world being the most popular form of payment, both in online and offline purchases, the growing number of fraud negatively impacts the revenue of several financial institutions in addition to the people who are victims of scams. In this way, a categorical and efficient detection becomes of great value in maintaining trust in the payment system. Thus, methodologies that use machine learning and artificial intelligence are being developed in order to improve the security of institutions. In this context, there are two main types of algorithms, one that uses supervised learning, such as Naive Bayes (reference), logistic regression (reference), decision trees (reference), Knn and neural networks (reference), and others known as learning. unsupervised such as K – means (reference), Clustering and hierarchical. Although techniques have satisfactory results, even today they are satisfactory results, they are still great results obtained with high precision since they were fraudulent in relation to normal transactions (false positive results) seen with fraudulent transactions (false) because the vast majority of transactions are legitimate.

Keywords: Machine Learning, Supervised Learning, Artificial Intelligence.

Lista de figuras

Lista de tabelas

Sumário

1 Introdução

Os pagamentos com cartões de crédito cresceram cerca de 24,6% no ano de 2022 em comparação com o período de janeiro à dezembro de 2021, sendo assim foram movimentados cerca de R$ 2,1 trilhões em pagamentos por meio de cartões de credito segundo um balanço divulgado pela Associação Brasileira das Empresas de Cartões de credito e serviços (ABECS, 2022). Com o crescimento do *e-commerce* (comercio eletrônico) e o tempo de confinamento devido grande parte ao controle da crise de saúde COVID-19, as compras pela internet tiveram um grande salto. Contudo, vale destacar que com um cenário econômico adverso de pandemia mundial a capacidade das famílias de manterem suas contas em dia vem sendo prejudicada, o que pode estar diretamente ligado ao crescimento descontrolado da inflação, o aumento do desemprego e estes consequentemente podem gerar um aumento da inadimplência.

A tecnologia tem constante impacto, em especial em nossas relações interpessoais, molda nossos comportamentos e nossas expectativas em relação as ferramentas que estão a nosso dispor. O volume de transações realizadas através da internet é o maior já visto nesses últimos anos, porém, ao passo em que o avanço tecnológico pode ser usado para o bem, pode ser utilizado para o mal também, isso de alguma forma facilita a ação de grupos criminosos em variados crimes cibernéticos. Esses novos tipos de crime não possuem limites territoriais, desta maneira com agravante da não necessidade de deslocamento do criminoso para a prática de tais condutas, em resumo a pessoa pode praticar inúmeras condutas lesivas sem sair de sua residência, contudo prejudicando pessoas residentes em todo planeta.

Muitas praticas desses crimes utilizam dados pessoais e começaram a ficar cada vez mais constantes, o que fortaleceu o debate quanto a necessidade de regulamentação em práticas envolvendo o uso de dados pessoais. No Brasil a discussão ganhou espaço, e após anos de debate em 14 de agosto de 2018 foi sancionada a Lei Geral de Proteção de Dados Pessoais do Brasil (LGPD), lei 13.709/2018 entrando assim em vigor em 18 de setembro de 2020. Esta lei muda a forma de funcionamento e operação das organizações

ao estabelecer regras claras sobre a coleta, armazenamento, tratamento e principalmente compartilhamento de dados pessoais, impondo um padrão mais elevados de proteção e penalidades para o não cumprimento das normas.

Por outro lado, houve grandes avanços em metodologias que utilizam de aprendizagem de máquina e inteligência artificial, que por meio de algoritmos torna -se possível que sistemas consigam prever padrões gerados pelas informações coletadas identificando possíveis vulnerabilidades em sistemas de pagamentos, e implementando soluções para impedir que ameaças ou mesmo fraudes prejudiquem as instituições financeiras e a sua reputação quando o assunto é segurança. Assim o *machine learning* é um importante aliado para evitar fraudes em sistemas de pagamento, seja quando ocorre uma mudança de comportamento que não está de acordo com o padrão de uso de determinado cliente, ou mesmo tentativas erradas de uso do cartão, podendo evitar assim não apenas um prejuízo para a organização, como também para o próprio usuário.

Estas técnicas possibilitam o processamento e a análise de conjuntos volumosos de dados, permitindo a identificação de padrões e características subjacentes. Dessa forma, a análise de fenômenos complexos relacionados a crimes cibernéticos se torna possível, considerando a interconexão de diversas variáveis influenciadoras. (HASTIE T.; TIBSHIRANI, 2008). Assim este trabalho possui uma abordagem que ressalta a capacidade das técnicas de Aprendizado de Máquina em lidar com desafios de alta dimensionalidade e interdependência entre variáveis.

Na busca pelo combate de fraudes é necessário recorrer a duas boas práticas que fazem grande diferença: prevenção e detecção de fraude. Basicamente as duas estão muito interligadas, a prevenção descreve medidas para impedir a ocorrência de fraudes, como números de identificação para cartões de banco, sistemas de segurança na internet para transações de credito, cartão com *Subscriber Identity Module* (SIM), medidas e estratégias de construção de regras, entre outros. Enquanto a detecção entra em ação justamente quando a prevenção falha, envolve a identificação da fraude o mais breve possível, uma vez que tenha sido cometida. As mesmas devem estar em constante melhoria e utilização com objetivo de barrar a ocorrência de novas fraudes com o mesmo comportamento por meio de ações e ferramentas estatísticas, visto que uma vez o método de detectar estiver em vigor e os criminosos entenderem como funciona irão adaptar as suas estratégias e tentarão de outras maneiras violar os sistemas (BOLTON; HAND., 2002).

Contudo existem duas principais formas práticas iniciais pela qual se desenvolvem esses algoritmos: aprendizado de máquina não supervisionado e supervisionado. O primeiro

ocorre quando o modelo aprende a partir de dados sem rótulos, por meio de técnicas estatísticas conseguimos extrair uma estrutura dos dados ou mesmo a relação entre os seus atributos. Enquanto o segundo o modelo aprende a partir de resultados pré definidos, utilizando os valores observados atribuídos a variável resposta para aprender e gerar previsões. Estes mesmos valores servirão de supervisão de tais resultados, que nos permite ajusta-los baseados nos erros. Entretanto podemos estender o objetivo geral da análise estatística com um retorno de pontuação de suspeita de fraude, ou seja, através de pontos classificamos que quanto maior a pontuação mais incomum é a observação da transação, ou mesmo quanto mais semelhantes as transações consideradas anteriormente como fraudulentas.

Desta maneira podemos perceber que há diferentes maneiras pelas quais uma fraude pode ser interpretada, pois muitos são os tipos de crimes e com o passar do tempo existe mudanças em suas práticas, assim com cenários diversos existem maneiras diferentes de calcular e medir operações suspeitas (BOLTON; HAND., 2002).

Essas transações quando podem ser então classificadas e ordenadas em grupos menores facilitam de certa maneira a investigação, e esses *outliers* recebem a devida atenção. Uma grande dificuldade enfrentada até então é que normalmente existem milhões de registros simultâneos legítimos para milhares de fraudes, trazendo assim um problema de desbalanceamento dos dados, bem como uma certa dificuldade da aprendizagem de máquina em dizer qual realmente é ou não considerado crime, visto que a mesma por meio de regras estatística e de forma empírica entende o mundo a partir do que sua leitura é capaz de compreender.

Neste trabalho objetiva - se discutir as vantagens de conceitos estatísticos e dos modelos de aprendizagem de máquina supervisionados com algoritmos conhecidos como classificadores, tendo como resposta binária 1 = sim é fraude, 0 = não é fraude, visto o desenvolvimento de processos contra fraudes no setor financeiro na qual foi desenvolvida uma pesquisa com dados bancários. As atividades realizadas foram estruturadas a partir de investigações estatísticas e os resultados trazem à tona um grande potencial de cunho investigativo com intuito de redução de danos causados pelos criminosos, bem como a importância de abstração e desenvolvimento de conceitos matemáticos.

1.1 Cartão de Crédito e Fraudes

Cartões de crédito são ferramentas financeiras que permitem aos usuários realizar compras e pagar por elas em uma data posterior. Quando uma pessoa usa um cartão de crédito para fazer uma compra, ela está efetivamente fazendo um empréstimo de dinheiro do emissor do cartão, geralmente um banco ou uma empresa financeira para pagar pelo produto ou serviço desejado. O usuário é então obrigado a pagar de volta o dinheiro emprestado, geralmente com juros e outros encargos financeiros, em uma data de vencimento predeterminado. Os cartões também podem ter limites de crédito, que são os valores máximos que os usuários podem gastar usando o mesmo. Hoje em dia possuem uma variedade de benefícios, como programas de recompensas, proteção contra fraudes e outros serviços adicionais, no entanto o uso exige responsabilidade para evitar dividas excessivas e altas taxas de juros.

Com avanço acelerado do *e-commerce* e a criação de bancos digitais que trazem grandes facilidades de abertura de novas contas, a disponibilidade de cartões de crédito teve um aumento considerado nos últimos anos, além de algumas vantagens que várias empresas oferecem para o seu uso no dia a dia como o *cash back*, que é o acúmulo de pontos que incentivam os clientes a pagar tudo por meio dos cartões, com a promessa de um retorno financeiro seja ele em créditos de viagens ou mesmo em produtos. Desta maneira um sistema de detecção de fraudes baseada em análise de dados de compras existentes do titular da conta é uma maneira promissora de reduzir os crimes, juntamente com evolução da tecnologia da informação e as leis que abrangem o uso e proteção dos dados.

As fraudes com cartões de crédito têm se expandido, e sua detecção não é uma tarefa fácil quando se utiliza processos normais, devido a isto o desenvolvimento de modelos estatísticos na detecção de crimes tornou – se importante, tanto em organizações acadêmicas ou mesmo empresarias. A fraude é definida quando um indivíduo usa o cartão de credito de outra pessoa para uso pessoal, enquanto o proprietário do mesmo, bem como o emissor do cartão não estão cientes de que está sendo utilizado. Desta maneira com o intuito de impedir tais crimes tem - se reunido grandes quantidades de dados em transações de contas, sendo vários sistemas, bem como modelos matemáticos e medidas preventivas que ajudam instituições como bancos e operadoras de cartões a reduzir os riscos financeiros e a aumentar confiabilidade desse serviço.

Hoje existem variadas maneiras de crimes cometidos aos clientes, sendo que podemos dividir em 2 tipos: fraudes offline ou online, sendo a primeira por meio do roubo do cartão

físico, interceptação de correspondências ou mesmo falsificação e a segunda via internet, telefone, compras web ou aplicativos na ausência do titular do cartão (BOLTON; HAND., 2002). Tais atividades fraudulentas afeta a todos que estão envolvidos na operação dos negócios feitos, ou seja, o titular da conta, o comerciante, o adquirente e também o emissor do cartão. Contudo analisando o impacto que os crimes causam nota - se que os titulares de contas é a parte menos afetada, visto que a responsabilidade do consumidor limita - se apenas pela transação, sendo tanto para cartões ausentes ou presentes.

Em grande maioria de bancos e financiadoras possuem políticas de proteção ao titular que cobre as perdas por uso indevido dos cartões, ficando o por conta do usuário apenas relatar as cobranças suspeitas ao banco emissor, que irá investigar o problema com o adquirente e o comerciante, fazendo assim o processo de estorno do valor contestado. Em contrapartida os comerciantes são os mais atingidos pelos golpes praticados, pois devem assumir a total responsabilidades pelas perdas, visto que os bancos dos emissores do cartão enviam um estorno ao lojista e revertem o crédito da transação, caso o comerciante não tenha nenhuma comprovação física como circuito de segurança de câmeras ou mesmo assinatura de entrega disponível para uma tentativa de contestação ficará muito difícil de reaver o valor.

Além da perda dessa transação a mercadoria ou serviço envolvidos na venda no ato da fraude são também considerados no impacto de prejuízo, pois não serão recuperados, sendo pior nos casos quando as margens de lucro do produto são baixas e outros gastos que podem existir, como o frete dos produtos quando compras são realizadas online (BHATLA; PRABHU; DUA, 2003). Dentre as várias formas de crimes cometidos, os que geram impacto significativo tanto para empresas como para clientes destacam -se:

Fraude de Identidade: ocorre com os criminosos roubam informações pessoais dos titulares dos cartões, como nome completo, endereço, data de nascimento, e número do cartão. Assim utilizam para solicitar empréstimos ou realizar transações financeiras em nome das vitimas

Fraude por *Skimming*: quando fraudadores instalam dispositivos em um determinado terminal de pagamento com o intuito de capturar informações dos clientes ao inserir os cartões.

Fraude por *Phishing*: por meio da trapaça os fraudadores tentam enganar os usuários por meio de e-mails ou mensagens falsas, fazendo com que os mesmos revelem informações confidenciais, como senhas e informações do cartão de credito.

Fraude interna: quando o funcionário de empresa que processa os pagamentos dos cartões ou mesmo que detém tais dados usa dos mesmos para praticar crimes.

Fraude por *Malware*: o *malware* é considerado um software malicioso, que pode ser instalado em dispositivos como celulares ou computadores com objetivo de obter informações confidencias dos usuários.

Tendo em vista os problemas apresentados acima, percebe – se que os crimes geram um impacto negativo significativo emocional tanto na vida das vítimas, causando estresse, preocupação e ansiedade em relação a segurança financeira como um grande prejuízo à reputação das empresas financeiras e bancos de forma geral. Visto isso, se torna de grande importância o investimento e desenvolvimento por meio de medidas através da segurança da lei de proteção dos dados pessoais e a otimização de processos de detecção e prevenção de tais práticas ilícitas, reduzir o custo de operação e assim minimizar os danos causados.

1.2 Problemas de Negócio

A fraude em transações financeiras representa um desafio significativo, cujo custo total abrange os prejuízos reais decorrentes dos crimes e os gastos com revisão, que estão diretamente relacionados ao volume de transações sob investigação. A segmentação de transações, produtos e titulares de cartão torna-se fundamental para identificar perfis potencialmente criminosos, permitindo que o sistema financeiro concentre seus esforços em um grupo reduzido de transações que necessitam de revisão.

O crescimento das transações comerciais com cartões também tem impulsionado o aumento das fraudes, especialmente com a expansão da internet, que abre novos canais para negócios, mas também aumenta a probabilidade de atividades fraudulentas em transações com cartão de crédito. Identificar corretamente o segmento de transações que requer revisão é crucial, uma vez que nem todas as transações apresentam o mesmo nível de risco associado. Ao adotar estratégias de segmentação eficientes, é possível mitigar os custos operacionais e focar nas transações mais suscetíveis a fraudes. (BHATLA; PRABHU; DUA, 2003)

A diversidade de atividades fraudulentas, manifestada através de várias formas de fraude, torna a detecção de comportamento criminoso uma tarefa complexa. Após a emissão de um cartão, as empresas geralmente confiam nos dados gerados periodicamente pelo usuário para avaliar o comportamento da conta e determinar qualquer suspeita de fraude. As regras de detecção de fraude são desenvolvidas com base em análises históricas

de comportamentos inadequados anteriores nas carteiras de clientes. No entanto, muitas instituições financeiras ainda se limitam ao uso de análises estatísticas básicas para criar essas regras, o que é exacerbado pela base de dados desbalanceada, com a grande maioria das transações sendo legítimas em comparação com as fraudulentas. Geralmente, as regras e condições adotadas são baseadas em variáveis de conta simples. No entanto, a maioria das das bases de dados geralmente é considerada desbalanceada, devido a grande maioria das transações sendo legais em contrapartida as transações fraudulentas.

O estudo atual foi projetado para simular a eficácia de um sistema de redução de crimes, para ser implementado como uma etapa pós - processamento em sistemas de autorizações de crédito em instituições financeiras. Quando a transação é enviada para autorização a mesma carrega informações que identificam o titular e as característica da transação. Não surpreendentemente, o uso de inteligência artificial e estatísticas mais robustas para detecção de fraude podem levar a melhores resultados. Desta maneira busca - se viabilizar um algoritmo baseado em uma amostra de contas legais e criminosas, seguido pela execução de testes a partir de modelos estatísticos treinados em um conjunto de transações.

1.3 Hipóteses da Pesquisa

Nesse contexto, o uso de algoritmos de Aprendizagem de Máquina (AM) tem se destacado como uma abordagem promissora para aprimorar a detecção de fraudes em cartões de crédito. A capacidade desses algoritmos em aprender padrões complexos e não lineares a partir de dados é um dos principais motivos para sua crescente aplicação nesse campo. Contudo, para garantir a eficiência e a confiabilidade desses sistemas, vamos investigar duas hipóteses importantes:

Hipótese 1: O custo agregado em verificar transações falsos positivos é significativamente maior do que o prejuízo causado por fraudes não detectadas.

Para avaliar essa hipótese, é fundamental entender que o processo de detecção de fraudes pode gerar dois tipos de erros: falsos positivos e falsos negativos. Falsos positivos são transações legítimas erroneamente classificadas como fraudes, o que pode levar a inconveniências e custos adicionais para os clientes e as instituições financeiras, pois requerem uma revisão manual para confirmar sua legitimidade. Por outro lado, falsos negativos são transações fraudulentas não detectadas pelo sistema, resultando em prejuízos financeiros diretos para a instituição e seus clientes.

É necessário, portanto, analisar o impacto econômico desses dois tipos de erros. Um elevado número de falsos positivos pode levar à insatisfação dos clientes e afetar negativamente a reputação da instituição. Por outro lado, uma alta taxa de falsos negativos resultará em perdas financeiras substanciais devido às fraudes não detectadas. Assim, é essencial encontrar um equilíbrio entre a redução de falsos positivos e falsos negativos para garantir a eficácia e a eficiência do sistema de detecção de fraudes.

Hipótese 2: A combinação de diferentes tipos de algoritmos e técnicas de Aprendizagem de Máquina potencializa a detecção de fraudes em cartão de crédito.

Essa hipótese parte do pressuposto de que a utilização de uma abordagem diversificada, combinando diferentes algoritmos e técnicas de AM, pode aumentar a capacidade de detecção de fraudes. Cada algoritmo possui suas próprias vantagens e limitações, e a combinação deles pode explorar suas características complementares, permitindo uma análise mais abrangente e precisa dos padrões de comportamento suspeitos.

A combinação desses diferentes algoritmos pode potencializar os resultados na detecção de fraudes, uma vez que cada um deles pode contribuir de forma única para a análise dos dados. Além disso, a adaptação contínua dos modelos, à medida que os fraudadores desenvolvem novas táticas, é fundamental para garantir que o sistema permaneça atualizado e eficaz na detecção de fraudes em constante evolução.

1.4 Objetivo

O objetivo desta pesquisa é desenvolver e implementar modelos estatísticos e algoritmos de aprendizado de máquina para a detecção e prevenção de fraudes em transações com cartões de crédito. A pesquisa se concentrará em utilizar técnicas de análise descritiva e preditiva em dados bancários de transações existentes para identificar padrões e comportamentos suspeitos que possam indicar fraudes. Além disso, será explorada a utilização de técnicas de aprendizado de máquina supervisionado, como classificadores, para criar modelos capazes de distinguir entre transações legítimas e fraudulentas. Desta maneira pretende - se:

1. Utilizar técnicas de análise descritiva: Realizar uma análise exploratória dos dados bancários de transações com cartões de crédito, identificando padrões, tendências e comportamentos associados aos clientes. Através dessa análise, busca-se compreender as características das transações fraudulentas em comparação com as legítimas.

2. Aplicar técnicas de análise preditiva: Desenvolver modelos estatísticos e de aprendizado de máquina que permitam a classificação das transações em duas categorias: legítimas e fraudulentas. Serão exploradas técnicas de aprendizado de máquina supervisionado, como classificadores, para criar modelos que sejam capazes de fazer essa distinção com alta precisão.

3. Avaliar a eficácia dos modelos: Realizar uma extensa avaliação dos modelos desenvolvidos, utilizando conjuntos de dados de teste independentes, para verificar sua capacidade de detecção e prevenção de fraudes em tempo real. Serão consideradas métricas como precisão, *recall* e F1-score para medir o desempenho dos modelos.

4. Propor melhorias nos processos de prevenção de fraudes: Com base nos resultados obtidos, a pesquisa buscará propor melhorias e recomendações para aprimorar os processos de detecção e prevenção de fraudes em transações com cartões de crédito, tanto para instituições financeiras quanto para comerciantes.

2 Métodos Propostos e Discussão

2.1 Modelos Supervisionados e Não Supervisionados

A modelagem de dados é uma tarefa essencial na ciência de dados e envolve a construção de modelos estatísticos ou algoritmos de aprendizado de máquina para extrair informações e realizar previsões. Caracterizam o aprendizado de máquina como sendo um subárea da inteligência artificial que busca a utilização de técnicas que possibilitem que computadores "aprendam" com a "experiência", baseadas na análise de dados passados (de teste), de modo indutivo, a partir da observação de um conjunto de dados, identificando padrões que determinem, com certa acurácia, os resultados pretendidos (AL, 2019). Dois principais tipos de modelagem são amplamente utilizados: modelagem supervisionada e modelagem não supervisionada.

Na modelagem supervisionada, o objetivo é treinar um modelo usando um conjunto de dados rotulados, onde cada exemplo de dados possui uma variável de resposta conhecida. O modelo é treinado para aprender a relação entre os atributos de entrada (variáveis independentes) e a variável de resposta (variável dependente) com base nos rótulos fornecidos. O processo de modelagem supervisionada envolve várias etapas, como pré-processamento de dados, seleção de atributos, treinamento do modelo e avaliação do desempenho. Alguns dos algoritmos de modelagem supervisionada incluem regressão linear, regressão logística, árvores de decisão, *random forest*, *support vector machines (SVM)* e redes neurais.

Os modelos supervisionados são úteis para prever ou classificar valores em uma variável de resposta (fraude = 1 e não fraude = 0) com base em informações disponíveis nos atributos de entrada. Eles são amplamente utilizados em problemas de classificação (como detecção de spam, diagnóstico médico) e regressão (como previsão de vendas, estimativa de preços imobiliários).

Na modelagem não supervisionada, não há variável de resposta ou rótulos conhecidos. O objetivo é explorar a estrutura e os padrões subjacentes nos dados e agrupá-los

ou encontrar representações latentes sem ter informações prévias sobre as categorias ou classes. Esses são usados para tarefas como análise de *cluster* (agrupamento), redução de dimensionalidade, descoberta de padrões ou associações. Alguns dos algoritmos comuns incluem *K-means*, DBSCAN, análise de componentes principais (PCA) e algoritmos de associação, como Priori. A modelagem não supervisionada é útil para explorar e descobrir informações ocultas nos dados, identificar grupos semelhantes de exemplos ou reduzir a dimensionalidade dos dados para visualização ou pré-processamento.

É importante mencionar que existe também a modelagem semi-supervisionada, que combina elementos de modelagem supervisionada e não supervisionada, onde parte dos dados possui rótulos e parte não possui. Isso permite aproveitar tanto os dados rotulados quanto os não rotulados para melhorar o desempenho do modelo.

O aprendizado desempenha um papel de significativa importância, especialmente no campo da mineração de dados, onde se almeja extrair padrões e informações valiosas de conjuntos extensos de dados, bem como na esfera da inteligência artificial, onde o desenvolvimento de algoritmos e modelos que possam simular processos cognitivos humanos tornou-se um desafio inovador (HASTIE T.; TIBSHIRANI, 2008). O estudo em questão se concentrará exclusivamente na utilização da modelagem supervisionada, cujo objetivo é prever ou classificar uma variável de resposta já conhecida. Nesse contexto, a modelagem supervisionada é uma abordagem muito poderosa para tarefas de previsão e classificação, uma vez que permite que o modelo seja treinado com dados em que já se conhece o resultado desejado. Isso possibilita que o algoritmo aprenda a mapear as informações presentes nos dados para fazer previsões precisas em novos casos.

Assim, ao se concentrar na modelagem supervisionada, o estudo visa criar um sistema de previsão ou classificação robusto, capaz de trazer *insights* valiosos e contribuir para o avanço do conhecimento em relação à variável de interesse. Através da aplicação de técnicas de Aprendizagem de Máquina, espera-se que o modelo alcance um alto desempenho na tarefa proposta, o que terá importantes aplicações práticas em diversas áreas, como medicina, finanças, marketing e muitas outras.

Antes de apresentar os algoritmos, é importante ressaltar que utilizaremos métricas na avaliação do desempenho de modelos de aprendizagem de máquina. Elas fornecem uma maneira objetiva de medir quão bem o modelo está realizando suas previsões ou classificações em relação aos dados de teste. As três métricas que serão abordadas a seguir são amplamente utilizadas na avaliação de modelos de classificação.

- Precisão:

  Avalia a proporção de casos positivos corretamente identificados em relação ao total de casos identificados como positivos pelo modelo. É especialmente útil quando o custo de falsos positivos é alto, ou seja, quando classificar erroneamente um caso negativo como positivo pode ter consequências graves.

- *Recall* (Taxa de verdadeiros positivos):

  Também conhecido como sensibilidade, mede a proporção de casos positivos corretamente identificados pelo modelo em relação ao total de casos positivos presentes nos dados. É relevante quando o custo de falsos negativos é alto, ou seja, quando classificar erroneamente um caso positivo como negativo pode levar a perdas significativas.

- F1-score:

O F1-score é uma métrica que combina a precisão e o *recall* em uma única medida. É calculado como a média harmônica entre precisão e *recall*, o que proporciona uma balança entre essas duas métricas. É útil quando se deseja encontrar um equilíbrio entre evitar falsos positivos e falsos negativos. Essas métricas são cruciais para avaliar a eficácia e o desempenho do modelo de classificação em diferentes cenários, permitindo que os desenvolvedores e pesquisadores selecionem o modelo mais adequado para a tarefa em questão.

2.1.1 Máquina de Vetores de Suporte (Vetor Machine)

A Máquina de Vetores de Suporte é um algoritmo de aprendizado de máquina supervisionado que é amplamente utilizado para tarefas de classificação e regressão. O objetivo principal do SVM é encontrar o hiperplano ótimo que divide o conjunto de dados em duas classes distintas. O hiperplano é uma fronteira de decisão que maximiza a margem entre as duas classes. A margem é definida como a distância entre o hiperplano e os pontos de dados mais próximos de cada classe, conhecidos como vetores de suporte. Os vetores de suporte são os pontos críticos para a definição do hiperplano, pois determinam sua posição e orientação.

A principal ideia por trás do SVM é transformar o problema de classificação em um problema de otimização. O algoritmo tenta encontrar o hiperplano ótimo minimizando uma função objetivo, que leva em consideração tanto a maximização da margem quanto

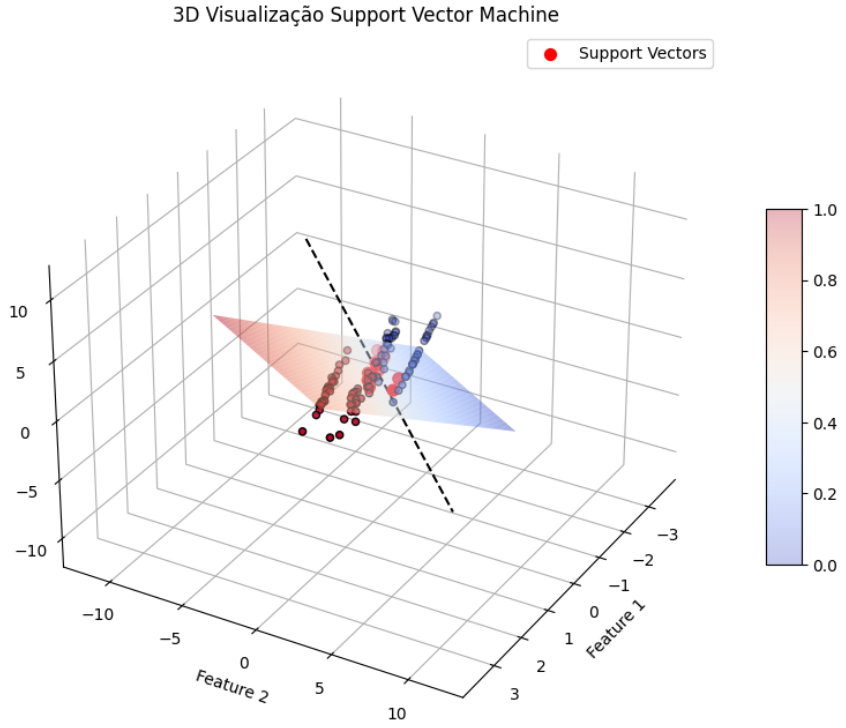

Figura 1: Visualização - Support Vector Machine 3d

a minimização do erro de classificação. O SVM é capaz de lidar com conjuntos de dados linearmente separáveis e também com conjuntos de dados que não podem ser separados linearmente diretamente.

Apesar disso, o treinamento de SVM em conjuntos de dados muito grandes pode ser computacionalmente caro. O presente estudo analisa o desempenho do modelo de aprendizado de máquina aplicado a um problema de classificação. O objetivo é avaliar a eficácia do modelo em termos de precisão, *recall* e F1-score.

O modelo apresentou um desempenho notável, com uma precisão de 93%, *recall* de 80% e F1-score de 86% conforme mostrado na tabela 2. Essas métricas destacam a capacidade do modelo em classificar corretamente as amostras fraudulentas. O recall do modelo foi de 80%, esse valor mostra que o modelo é capaz de encontrar a maioria das instâncias positivas.

Tabela 1: Resultados Máquina de Vetores de Suporte

Alvo	Previsão	Recall	F1 - Score
0	0.93	0.80	0.86
1	0.92	0.80	0.86
Acurácia			0.85

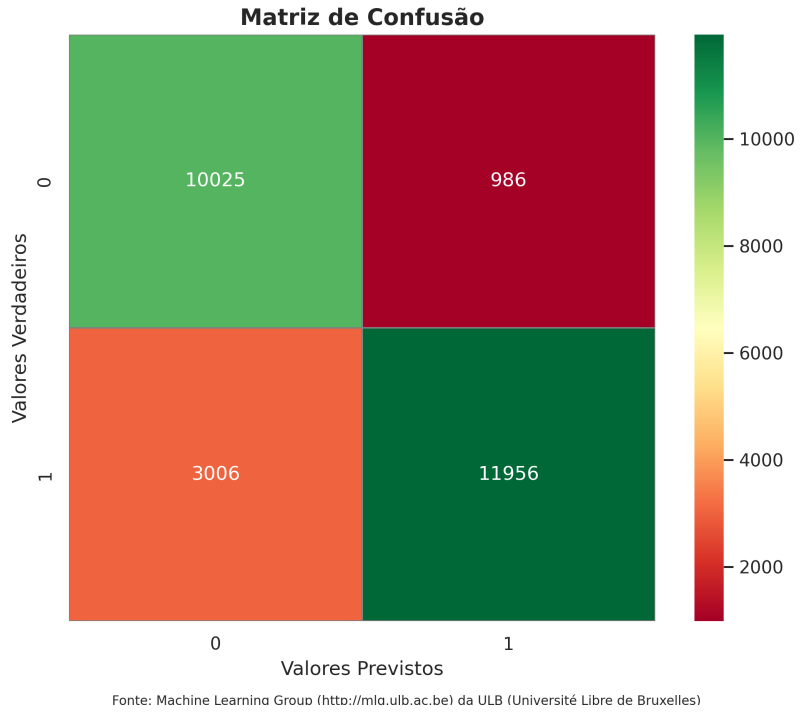

Figura 2: Gráfico - Matriz de Confusão *Vetor Machine*

O F1-score do modelo foi de 86%, essa métrica que combina a precisão e o *recall* em uma única medida balanceada, fornecendo uma medida geral do desempenho do modelo. Desta maneira esse valor indica que o modelo apresenta um bom equilíbrio entre a precisão e o *recall*, e sugerindo um desempenho promissor na tarefa de classificação. A alta precisão indica que o modelo minimiza os falsos positivos.

2.1.2 Regressão Logística

Técnica de aprendizado de máquina supervisionado amplamente utilizada para tarefas de classificação binária e multiclasse. É uma abordagem estatística que estima a probabilidade de um evento ocorrer com base em um conjunto de variáveis independentes. Ela estima a probabilidade de um evento ocorrer com base em variáveis independentes e utiliza uma função logística para transformar os valores de entrada em uma escala de probabilidade. É popular devido à sua simplicidade, interpretabilidade e eficiência computacional, sendo aplicada em várias áreas da ciência de dados.

O objetivo é modelar a relação entre as variáveis independentes e a variável dependente, que é uma variável categórica. Ao contrário da regressão linear, que prevê valores contínuos, a regressão logística prevê a probabilidade de uma observação pertencer a uma

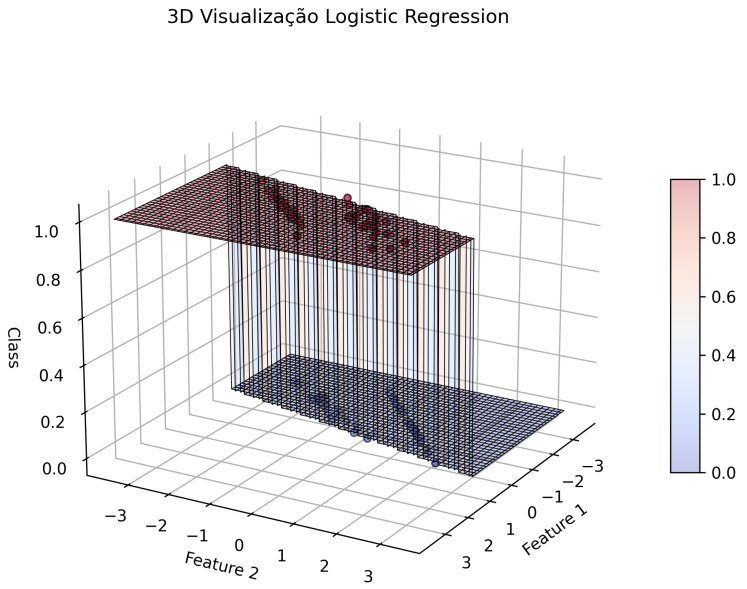

Figura 3: Visualização - Logistc Regression 3d

determinada classe.

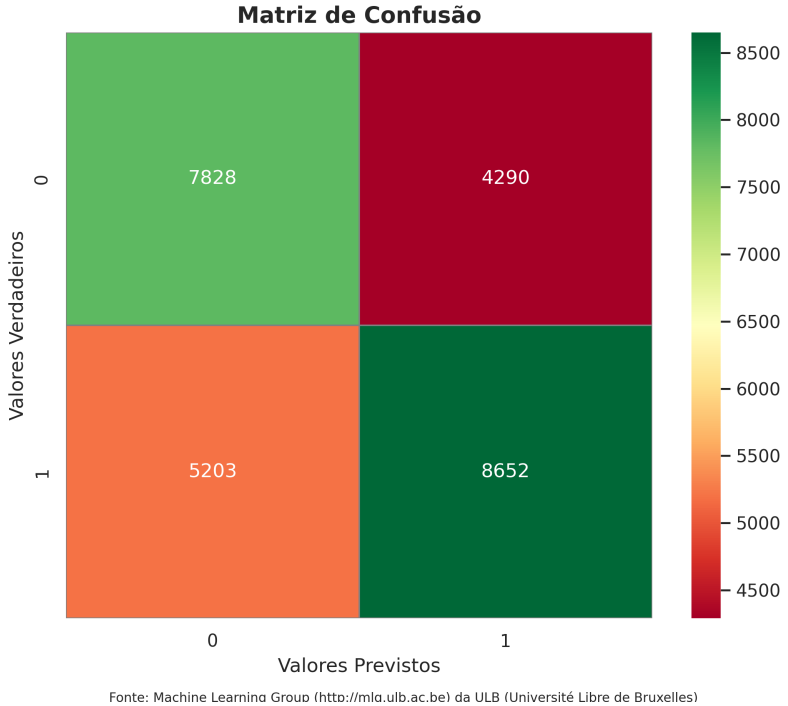

Figura 4: Gráfico - Matriz de Confusão Regressão Logística

Tabela 2: Resultados Regressão Logística

Alvo	Previsão	Recall	F1 - Score
0	0.59	0.65	0.62
1	0.67	0.62	0.64
Acurácia			0.63

O modelo de regressão demonstrou um desempenho moderado na tarefa de classificação, com acurácia de 63%, *recall* de 64% e precisão de 60% tabela 3. Embora esses resultados indiquem um progresso na classificação correta das transações consideradas fraude e na capacidade de encontrar casos positivos, é importante considerar a utilização da regressão logística como uma técnica inicial para a modelagem de problemas de classificação, mas também evidenciar a importância de aprimoramento do desempenho do modelo, visando melhorar a acurácia, o *recall* e a precisão.

2.1.3 Árvore de Decisão

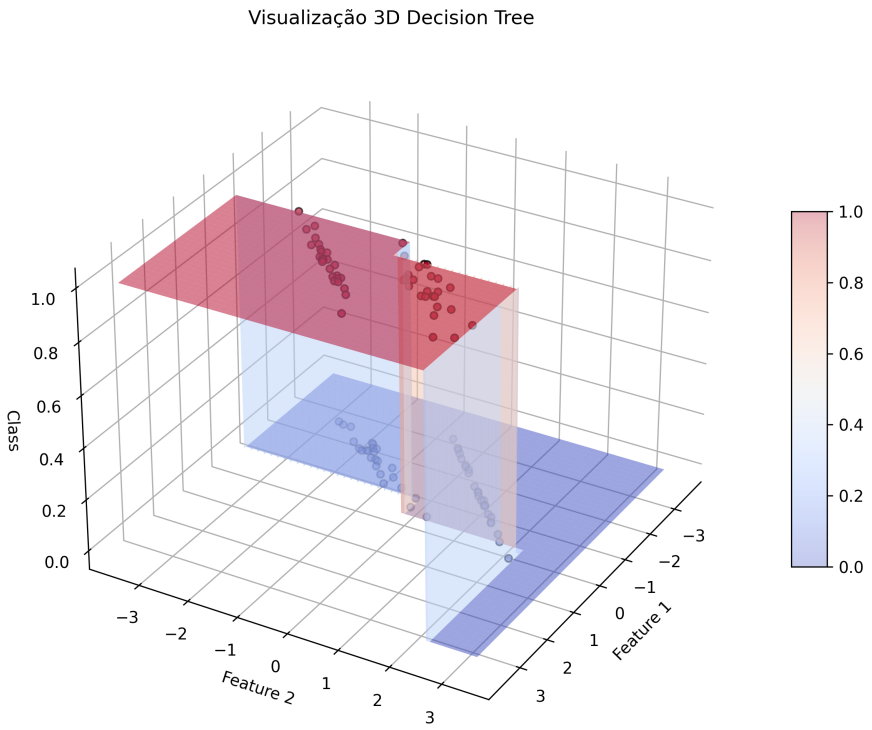

Figura 5: Visualização - Decision Tree 3d

A árvore de decisão é um algoritmo de aprendizado de máquina supervisionado usado para tarefas de classificação e regressão. É uma técnica poderosa e amplamente utilizada, especialmente por sua interpretabilidade e facilidade de uso. A ideia fundamental por trás da árvore de decisão é criar um modelo preditivo em forma de árvore, na qual cada nó interno representa uma decisão com base em um atributo específico, e cada folha representa uma classe ou valor de saída.

O algoritmo busca dividir o conjunto de dados de forma recursiva com base nos atributos disponíveis. O objetivo é encontrar as divisões que maximizam a homogeneidade dos dados dentro de cada subconjunto resultante. Uma vez que a árvore de decisão tenha sido construída, ela pode ser usada para fazer previsões em novos exemplos. A árvore é percorrida de cima para baixo, seguindo os ramos correspondentes às decisões baseadas nos valores dos atributos. No final, a previsão é feita com base na classe ou valor associado à folha alcançada. Essa técnica possui algumas vantagens, incluindo a capacidade de lidar com dados numéricos e categóricos, além disso, ela é relativamente rápida de construir e pode lidar com grandes volumes de dados.

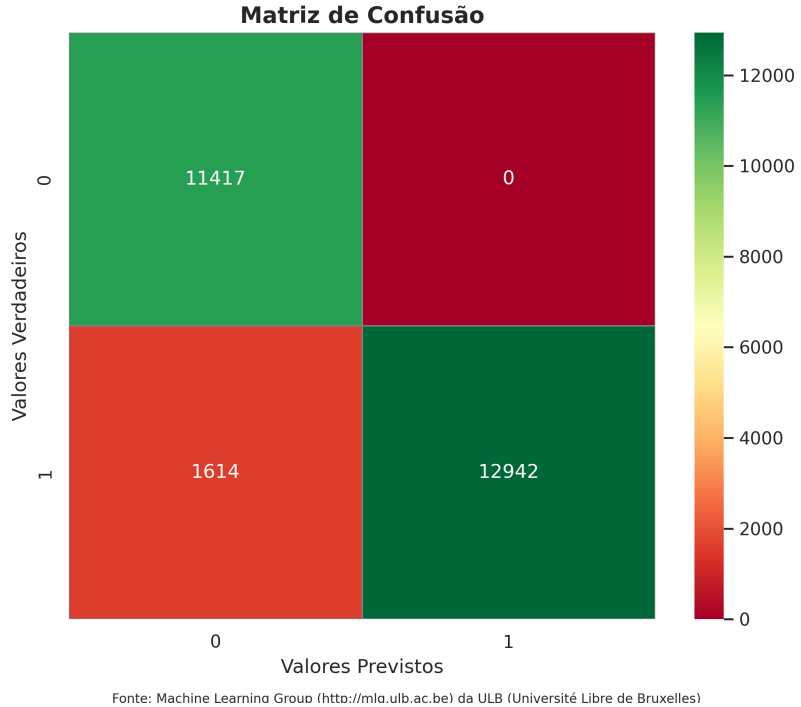

Figura 6: Gráfico - Matriz de Confusão *Decision tree learning*

Os resultados obtidos revelaram um excelente desempenho do modelo de Árvore de Decisão, alcançando uma precisão de 1.00. Esse valor significa que o modelo classificou corretamente todas as amostras positivas, o que demonstra uma alta taxa de acerto. Além

Tabela 3: Resultados Árvore de Decisão

Alvo	Previsão	Recall	F1 - Score
0	0.88	1.00	0.93
1	1.00	0.89	0.94
Acurácia			0.94

disso, o *recall* de 0.89 indica que o modelo conseguiu capturar uma parte significativa das instâncias positivas, evidenciando sua capacidade de identificar corretamente as amostras positivas. Com um F1-score de 0.94, o modelo demonstra um bom equilíbrio entre a precisão e o *recall*, sendo uma escolha promissora para a tarefa de classificação em questão.

Em resumo, o modelo de Árvore de Decisão alcançou resultados altamente satisfatórios, mostrando-se uma escolha promissora para a tarefa de classificação das amostras, com um alto grau de precisão, boa capacidade de identificação das fraudes e equilíbrio entre precisão e *recall*.

2.1.4 KNN - K-ésimo Vizinho mais Próximo (k-nearest neighbors algorithm)

O KNN *(K-Nearest Neighbors)* é baseado no princípio de que exemplos semelhantes tendem a estar próximos uns dos outros no espaço de características. É um algoritmo não paramétrico, o que significa que não faz suposições explícitas sobre a distribuição dos dados. Em vez disso, ele armazena todo o conjunto de dados de treinamento no momento da fase de treinamento, para que possa ser usado posteriormente durante a fase de teste para fazer previsões.

O funcionamento do KNN é relativamente simples. Quando um novo exemplo é apresentado ao algoritmo, ele calcula a distância entre o exemplo de teste e todos os exemplos de treinamento. A distância mais comumente usada é a distância euclidiana, mas outras medidas de distância, como a distância de Manhattan, também podem ser usadas. Em seguida, os K exemplos mais próximos (onde K é um parâmetro definido pelo usuário) são selecionados com base na distância. Para tarefas de classificação, a classe mais frequente entre os K vizinhos é atribuída ao exemplo de teste. Para tarefas de regressão, uma média ou uma combinação ponderada dos valores alvo dos K vizinhos pode ser usada para fazer a previsão.

Uma das principais vantagens do KNN é poder lidar com dados não lineares sendo também considerado robusto a ruídos nos dados. Outra vantagem é a capacidade de

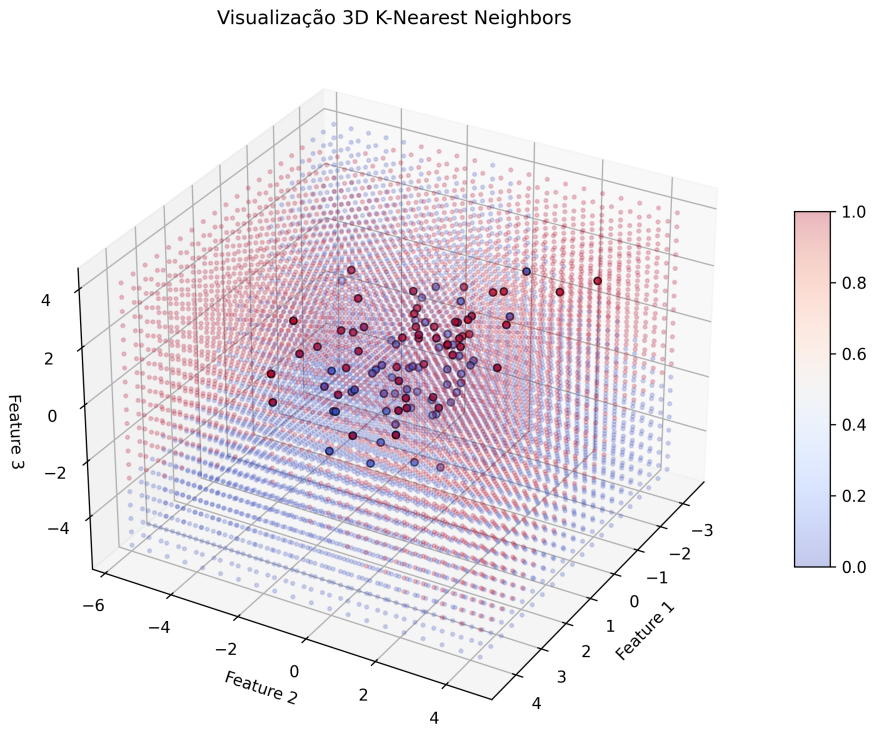

Figura 7: Visualização - K - Nearest Neighbors 3d

adaptação a novos dados, já que o modelo é atualizado a cada novo exemplo adicionado. No entanto, também possui algumas limitações. Uma delas é a necessidade de armazenar todo o conjunto de dados de treinamento, o que pode levar a uma alta demanda de memória, especialmente para conjuntos de dados grandes. Além disso, o cálculo das distâncias entre todos os exemplos pode ser computacionalmente caro, especialmente em conjuntos de dados com muitas características.

Os resultados revelam que o modelo KNN apresentou um desempenho promissor na classificação das amostras. Com uma acurácia de 86%, o modelo demonstrou uma boa taxa geral de acerto. Além disso, o *recall* de 78% indica que o modelo foi capaz de encontrar a maioria das instâncias positivas, identificando corretamente a maior parte dos casos que eram fraudes.

Impressionantemente, a precisão atingiu 99%, o que significa uma alta taxa de acerto na classificação das transações criminosas e uma minimização significativa dos falsos positivos. O F1-score, que combina precisão e *recall*, ficou em 87%, indicando que o modelo conseguiu alcançar um bom equilíbrio entre identificar corretamente as instâncias positi-

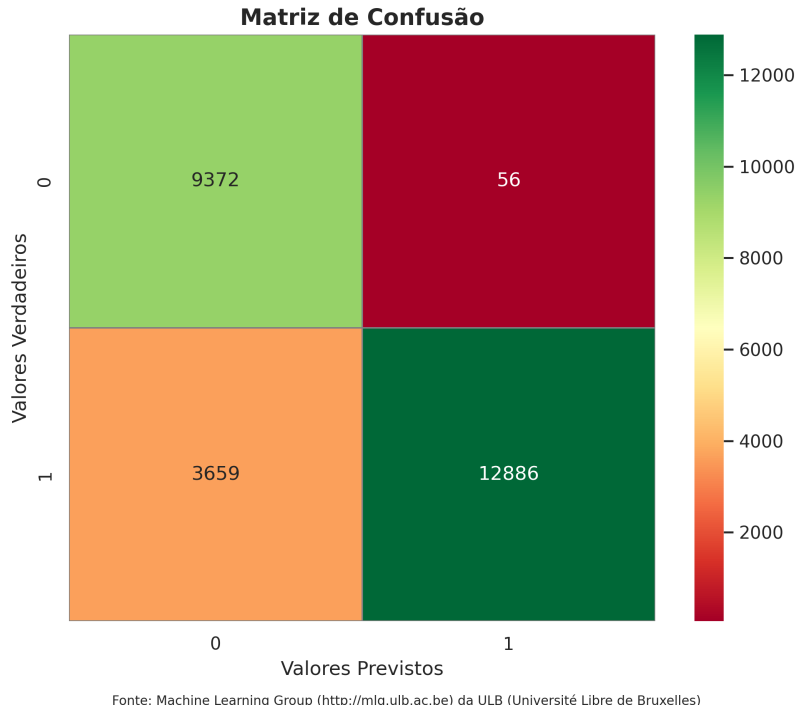

Figura 8: Gráfico - Matriz de Confusão *k-nearest neighbors algorithm*

Tabela 4: Resultados KNN - K-ésimo Vizinho mais Próximo

Alvo	Previsão	Recall	F1 - Score
0	0.72	0.99	0.84
1	1.00	0.78	0.88
Acurácia			0.86

vas e evitar tanto os falsos positivos quanto os falsos negativos.

Em resumo, o modelo KNN apresentou resultados encorajadores, com alta precisão na classificação das fraudes e uma capacidade razoável de encontrar a maioria delas. O F1-score demonstra um equilíbrio satisfatório entre essas duas métricas cruciais, tornando o modelo uma escolha promissora para a tarefa de classificação em questão.

2.1.5 Perceptron multicamadas (Multilayer perceptron)

O *MLP Classifier*, ou Perceptron Multicamadas, é uma rede neural artificial composta por várias camadas de neurônios, incluindo uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída. Durante a fase de treinamento, o MLP ajusta os pesos sinápticos entre os neurônios para minimizar uma função de perda, que mede a diferença entre as previsões do modelo e os rótulos corretos dos exemplos de treina-

mento. O algoritmo de otimização, como o gradiente descendente, é usado para atualizar iterativamente os pesos de forma a minimizar a função de perda.

Uma das principais vantagens do *MLP Classifier* é sua capacidade de aprender representações complexas dos dados. Com camadas ocultas e funções de ativação não lineares, o MLP pode capturar padrões e interações complexas entre as variáveis de entrada. Além disso, o MLP é capaz de lidar com problemas de classificação não lineares, onde as fronteiras de decisão não podem ser separadas linearmente.

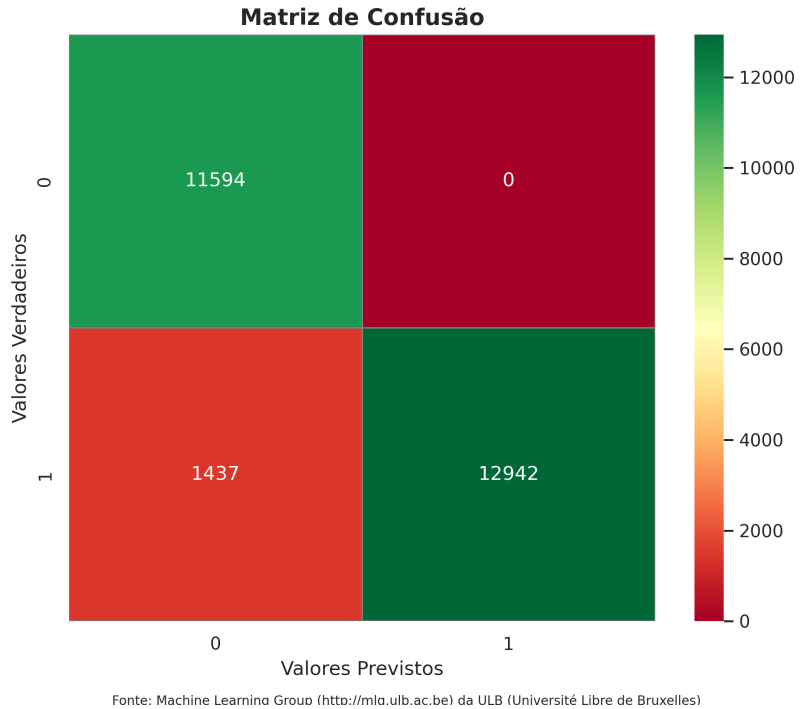

Figura 9: Gráfico - Matriz de Confusão *Multilayer perceptron*

Tabela 5: Resultados MLP Classifier - Perceptron multicamadas

Alvo	Previsão	Recall	F1 - Score
0	0.89	1.00	0.94
1	1.00	0.90	0.95
Acurácia			0.94

Os resultados obtidos revelaram um excelente desempenho do modelo MLP Classifier, alcançando uma acurácia de 95%. Esse valor indica um alto grau de precisão na classificação geral das amostras, demonstrando que o modelo foi capaz de classificar corretamente a grande maioria das transações.

Além disso, o recall do modelo foi de 90%, o que sugere que o modelo conseguiu

encontrar a maioria das transações fraudulentas. Esse resultado é bastante significativo, pois mostra a capacidade do modelo de identificar corretamente as transações que são realmente fraudulentas.

Outro destaque importante é a precisão do modelo, que atingiu 100%. Esse valor representa a proporção de instâncias classificadas como positivas que são realmente positivas, indicando que o modelo não teve nenhum falso positivo na classificação das transações fraudulentas. Isso é extremamente relevante, pois significa que todas as vezes que o modelo classificou uma transação como fraudulenta, ela realmente era fraudulenta.

Em resumo, os resultados obtidos pelo modelo MLP Classifier são altamente satisfatórios, com uma alta acurácia, recall significativo e uma precisão perfeita na classificação das transações fraudulentas.

2.1.6 Floresta Aleatória (Random Forest)

A Floresta Aleatória *(Random Forest)* é um algoritmo que combina várias árvores de decisão individuais para realizar tarefas de classificação e regressão. Ela é conhecida por sua capacidade de lidar com problemas complexos e produzir resultados precisos.

O algoritmo da Floresta Aleatória funciona construindo múltiplas árvores de decisão independentes entre si. Cada árvore é treinada em uma amostra aleatória do conjunto de dados de treinamento, chamada de amostragem de *bootstrap*, na qual exemplos são selecionados com substituição.

Durante a fase de treinamento, cada árvore é expandida até sua capacidade máxima ou até que um critério de parada seja atingido. Em problemas de classificação, a árvore toma decisões com base na maioria das classes dos exemplos de treinamento nas folhas. Para fazer uma previsão com uma Floresta Aleatória treinada, cada árvore individual na floresta é percorrida e sua previsão é computada. No caso de classificação, a classe mais frequente entre as árvores é escolhida como a previsão final.

A Floresta Aleatória é altamente precisa e robusta, lida bem com conjuntos de dados grandes e de alta dimensionalidade, e é menos suscetível a overfitting em comparação com uma única árvore de decisão. Além dissso, é eficiente em termos computacionais, pois as árvores podem ser construídas em paralelo.

Os resultados obtidos revelaram que o modelo de Floresta Aleatória alcançou uma precisão de 1.00, o que indica um alto grau de precisão na classificação correta das amostras

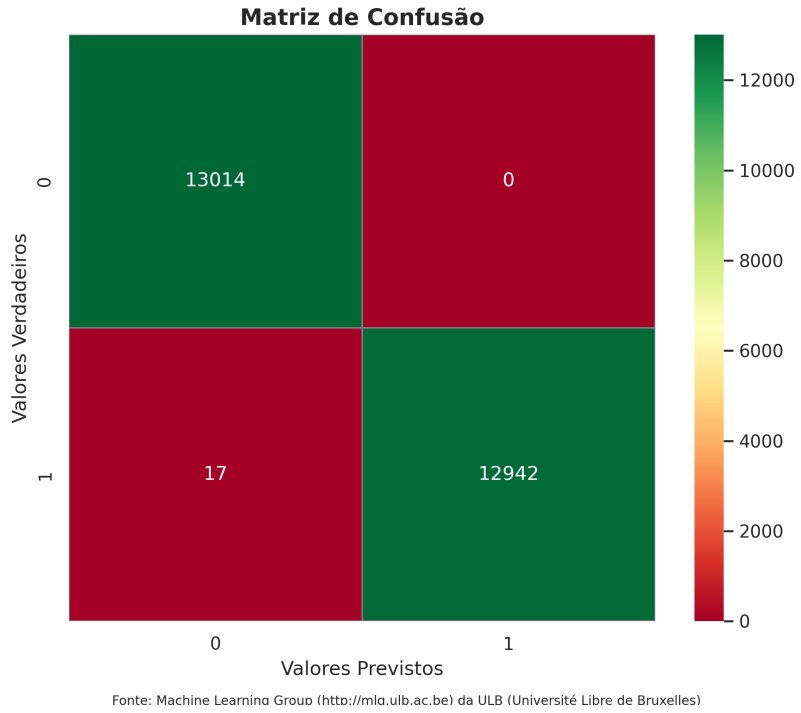

Figura 10: Gráfico - Matriz de Confusão *Random forest*

Tabela 6: Resultados Floresta Aleatória

Alvo	Previsão	Recall	F1 - Score
0	1.00	1.00	1.00
1	1.00	1.00	1.00
Acurácia			0.99

positivas. Além disso, o recall do modelo também foi de 1.00, demonstrando que o modelo foi capaz de capturar todas as instâncias positivas, não deixando nenhum caso positivo sem identificação. O F1-score do modelo também foi de 1.00, o que mostra que o modelo apresenta um perfeito equilíbrio entre a precisão e o recall, sendo capaz de identificar corretamente todas as instâncias positivas sem cometer erros. Esses resultados destacam o excelente desempenho do modelo de Floresta Aleatória na tarefa de detecção de fraudes em cartão de crédito.

2.1.7 Classificador de aumento de gradiente - Gradient Boosting Classifier

Algoritmo de pertencente à família de métodos de *boosting*. Ele é usado principalmente para tarefas de classificação, mas também pode ser aplicado a problemas de regressão.

O mesmo constrói um modelo preditivo forte combinando vários modelos de aprendizado fracos.

A ideia central é construir os modelos de forma iterativa, onde cada novo modelo é treinado para corrigir os erros cometidos pelos modelos anteriores. A construção é feita através de um processo chamado *boosting*, que combina as previsões dos modelos individuais para obter uma previsão final.

O algoritmo inicia criando um modelo fraco, como uma árvore de decisão rasa, que tem um desempenho apenas um pouco melhor do que uma classificação aleatória. Em seguida, ele analisa os erros cometidos pelo modelo e ajusta os pesos dos exemplos de treinamento para se concentrar mais nos exemplos que foram classificados incorretamente.

Na iteração seguinte, outro modelo fraco é construído, que se concentra nos exemplos que foram mal classificados pelo modelo anterior. Esse processo é repetido várias vezes, com cada modelo sucessivo corrigindo os erros do modelo anterior.

Para combinar as previsões dos modelos individuais e obter uma previsão final, o *Gradient Boosting Classifier* utiliza uma abordagem chamada de agregação por soma ponderada. Cada modelo contribui com uma ponderação para a previsão final, levando em consideração seu desempenho durante o treinamento.

Uma das vantagens é que ao corrigir os erros dos modelos anteriores, ele melhora continuamente seu desempenho ao longo das iterações. Além disso, o algoritmo pode lidar com dados heterogêneos e capturar relações não lineares entre as características.

No entanto, o *Gradient Boosting Classifier* também apresenta algumas considerações importantes. Ele tende a ser mais suscetível ao *overfitting* em comparação com outros algoritmos, especialmente se o número de iterações for muito alto. Portanto, é necessário ajustar adequadamente os hiperparâmetros, como a taxa de aprendizado e o número de iterações, para obter um bom equilíbrio entre o desempenho e a generalização.

Tabela 7: Resultados Classificador de aumento de gradiente

Alvo	Previsão	Recall	F1 - Score
0	0.62	0.66	0.64
1	0.69	0.64	0.66
Acurácia			0.65

Os resultados obtidos revelaram que o modelo de *Gradient Boosting Classifier* obteve uma precisão de 0.69, indicando uma taxa razoável de acerto na classificação correta das

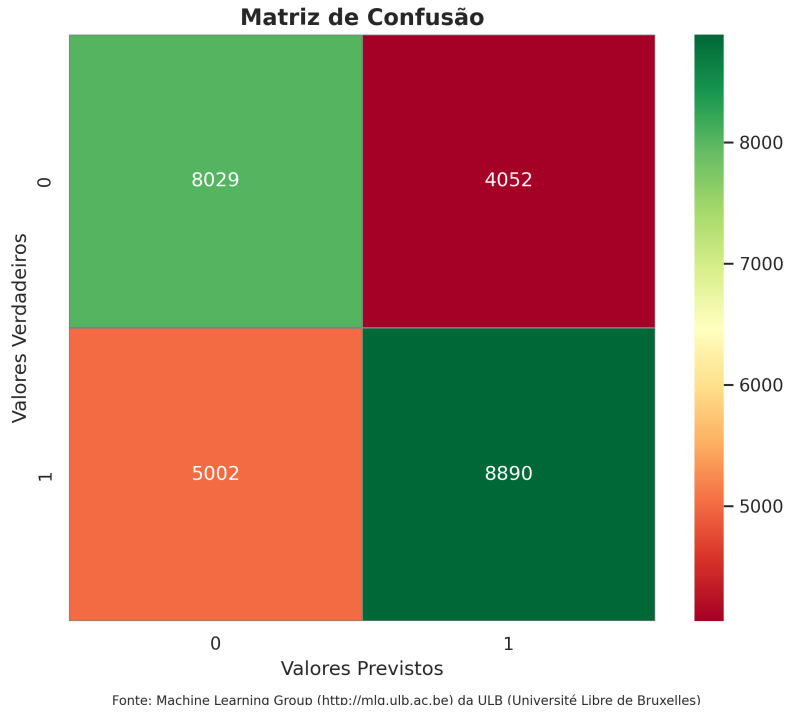

Figura 11: Gráfico - Matriz de Confusão *Gradient Boosting Classifier*

amostras positivas. No entanto, o *recall* do modelo foi de 0.64, mostrando que o modelo não foi tão eficiente em capturar todas as instâncias positivas presentes nos dados.

O F1-score do modelo foi de 0.66, refletindo um equilíbrio razoável entre a precisão e o *recall*, mas ainda deixando espaço para melhorias na capacidade do modelo em identificar corretamente todas as amostras positivas.

Esses resultados sugerem que o modelo de *Gradient Boosting Classifier* apresenta um desempenho moderado na tarefa de detecção de fraudes em cartão de crédito. É importante considerar aprimoramentos, a fim de buscar um modelo com melhor desempenho e maior capacidade de identificar transações fraudulentas com maior precisão e *recall*.

2.1.8 Aumento de Gradiente Extremo - XGBoost (Extreme Gradient Boosting)

Também pertencente à família de métodos de *boosting*, é uma versão otimizada e aprimorada do algoritmo de *Gradient Boosting*.

Uma das principais melhorias do XGBoost em relação ao *Gradient Boosting* tradicional é o tratamento eficiente de recursos como a regularização. Ele utiliza uma técnica chamada regularização L1 e L2 para controlar a complexidade do modelo e evitar o *over-*

fitting. Além disso, o XGBoost possui uma função de perda customizável, permitindo otimizar diferentes objetivos, como erro quadrático médio, entre outros.

Outra melhoria significativa do XGBoost é o uso de um algoritmo de otimização eficiente para encontrar os melhores valores para os hiperparâmetros do modelo. Isso inclui a otimização da taxa de aprendizado, o número máximo de árvores, a profundidade máxima das árvores e outros parâmetros relacionados à estrutura do modelo. A otimização dos hiperparâmetros permite ajustar o modelo de forma mais precisa e obter um desempenho melhor.

O XGBoost também introduz um mecanismo de tratamento de dados ausentes *(missing values)* de forma nativa. Ele automaticamente lida com dados ausentes durante o treinamento, evitando a necessidade de pré-processamento adicional. Além disso, XGBoost é conhecido por sua eficiência computacional e escalabilidade. Ele foi projetado para lidar com grandes conjuntos de dados e pode ser executado em sistemas distribuídos para acelerar o treinamento.

Devido a todas essas melhorias e otimizações, o XGBoost se tornou uma escolha popular em competições de ciência de dados e é amplamente utilizado em problemas do mundo real. Ele tem um desempenho impressionante em termos de precisão e é considerado um dos algoritmos mais poderosos disponíveis.

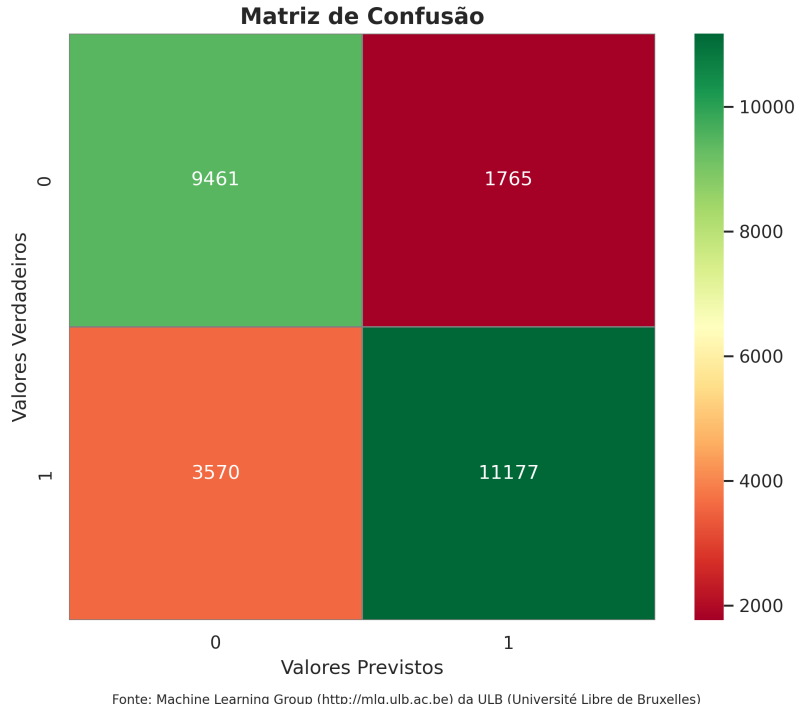

Figura 12: Gráfico - Matriz de Confusão *Xgboost*

Tabela 8: Resultados Xgboost

Alvo	Previsão	Recall	F1 - Score
0	0.73	0.84	0.78
1	0.86	0.76	0.81
Acurácia			0.79

Os resultados obtidos demonstraram um ótimo desempenho do modelo XGBoost na detecção de fraudes em cartão de crédito. Com uma precisão de 0.86, o modelo apresentou uma alta taxa de acerto na classificação correta das amostras positivas. Isso significa que a maioria das transações identificadas como fraudulentas pelo modelo eram realmente fraudulentas.

O *recall* do modelo, que foi de 0.76, indica que o modelo conseguiu capturar uma parte significativa das instâncias positivas, identificando corretamente a maioria das transações fraudulentas presentes nos dados.

O F1-score do modelo, que foi de 0.81, isso sugere que o modelo conseguiu alcançar uma combinação satisfatória de acurácia na classificação das transações positivas e na capacidade de encontrar a maioria das fraudes.

Esses resultados indicam que o modelo XGBoost é uma escolha promissora para a detecção de fraudes em cartão de crédito, apresentando um bom equilíbrio entre precisão e *recall*. No entanto, é sempre importante considerar a possibilidade de ajustes e melhorias no modelo, bem como a utilização de mais dados de treinamento, para aprimorar ainda mais seu desempenho e garantir uma detecção eficaz e confiável de transações fraudulentas.

2.1.9 Light Gradient Boosting Model (LGBM)

Algoritmo baseado em *boosting* que se destaca pela sua eficiência e desempenho em problemas de classificação e regressão. Ele foi desenvolvido pela *Microsoft* e é conhecido por ser rápido, preciso e escalável.

O LGBM é uma implementação otimizada do algoritmo *Gradient Boosting*, projetada para lidar com grandes volumes de dados e conjuntos de características de alta dimensionalidade. Ele utiliza uma estrutura de árvore de decisão em conjunto com uma abordagem de crescimento por nível, o que torna o treinamento e a previsão mais rápidos em comparação com outros algoritmos de *boosting*.

Uma das principais características do LGBM é o uso de uma estratégia de split ba-

seada em histogramas para encontrar as melhores divisões nas árvores de decisão. Essa abordagem agrupa os valores dos atributos em intervalos discretos e constrói histogramas para acelerar o processo de seleção dos *splits* mais informativos. Essa técnica permite uma eficiente busca pelos melhores pontos de divisão, resultando em tempos de treinamento mais curtos.

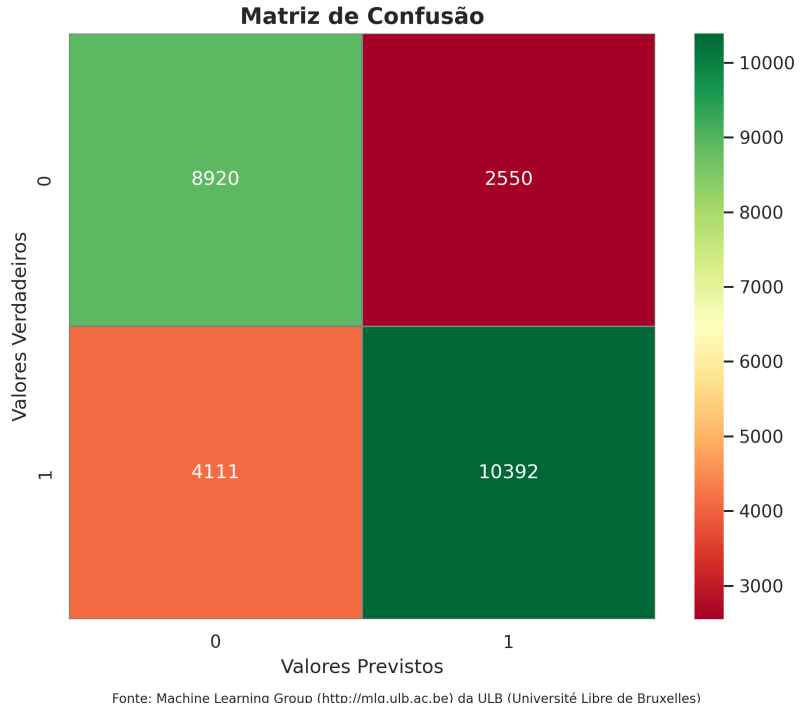

Figura 13: Gráfico - Matriz de Confusão *Light Gradient Boosting Model*

Tabela 9: Resultados Modelo de aumento de gradiente

Alvo	Previsão	Recall	F1 - Score
0	0.68	0.78	0.73
1	0.80	0.72	0.76
Acurácia			0.74

Outro aspecto importante do LGBM é o uso de uma técnica chamada *Leaf-wise* (crescimento por folha), na qual a árvore é construída de forma a maximizar a redução do erro em cada divisão, selecionando a folha com maior ganho em vez de seguir a abordagem tradicional por nível. Isso resulta em uma árvore mais profunda e potencialmente mais precisa.

Além disso, o LGBM utiliza a técnica de exclusão de características por amostragem aleatória *(feature sub-sampling)* durante a construção de cada árvore. Essa técnica me-

lhora a capacidade de generalização do modelo, reduzindo a correlação entre as árvores e evitando *overfitting*.

O LGBM também permite ajustar vários hiperparâmetros, como a taxa de aprendizado, a profundidade máxima das árvores, o número de árvores na floresta e o tamanho mínimo das folhas. A otimização desses hiperparâmetros pode ser realizada usando técnicas como busca em grade *(grid search)* ou otimização bayesiana.

Os resultados obtidos demonstraram um bom desempenho do modelo LGBM na detecção de fraudes em cartão de crédito, com uma precisão de 0.80, o modelo apresentou uma alta taxa de acerto na classificação correta das amostras positivas. Isso significa que a maioria das transações identificadas como fraudulentas pelo modelo eram realmente fraudulentas.

O *recall* do modelo, que foi de 0.73, indica que o modelo conseguiu capturar uma parte significativa das instâncias positivas, identificando corretamente a maioria das transações fraudulentas presentes nos dados.

O F1-score do modelo, que foi de 0.76, isso sugere que o modelo conseguiu alcançar uma combinação satisfatória de acurácia na classificação das transações positivas e na capacidade de encontrar a maioria das fraudes.

Esses resultados indicam que o modelo LGBM é uma escolha promissora para a detecção de fraudes em cartão de crédito, apresentando um bom equilíbrio entre precisão e *recall*.

3 Conclusões finais

Neste trabalho, exploramos o desempenho de diferentes algoritmos classificadores em relação à detecção de fraudes em um conjunto de dados de transações de cartão de crédito. Primeiramente, constatamos que a maioria dos algoritmos apresentaram bons resultados em termos de precisão, *recall* e F1-score, indicando que são capazes de identificar corretamente a ocorrência de fraudes, ao mesmo tempo que minimizam os falsos positivos e falsos negativos. Após a execução dos algoritmos classificadores e a avaliação de suas métricas de desempenho, podemos fazer algumas conclusões sobre nossas hipóteses.

Tabela 10: Resultados Medidas de Desempenho

Classificador	Alvo	Previsão	Recall	F1 - Score	Acurácia
Vetor Machine	0	0.93	0.80	0.86	
	1	0.92	0.80	0.86	0.85
Regressão Logística	0	0.59	0.65	0.62	
	1	0.67	0.62	0.64	0.63
Árvore de Decisão	0	0.88	1.00	0.93	
	1	1.00	0.89	0.94	0.94
Knn	0	0.72	0.99	0.84	
	1	1.00	0.78	0.88	0.86
Perceptron Multicamadas	0	0.89	1.00	0.94	
	1	1.00	0.90	0.95	0.94
Floresta Aleatória	0	1.00	1.00	1.00	
	1	1.00	1.00	1.00	0.99
Gradient Boosting Classifier	0	0.62	0.66	0.64	
	1	0.69	0.64	0.66	0.65
XGBoost	0	0.73	0.84	0.78	
	1	0.86	0.76	0.81	0.79
LGBM	0	0.68	0.78	0.73	
	1	0.80	0.72	0.76	0.74

Os resultados obtidos revelaram que o modelo de Floresta Aleatória *(Random Forest)* e o Perceptron Multicamadas apresentaram os melhores desempenhos na detecção de fraudes em cartão de crédito. Ambos os modelos alcançaram uma precisão de 90 a 100%

respectivamente, ou seja, classificaram corretamente todas as amostras positivas. Ambos os modelos também alcançaram valores elevados de F1-score, o que evidencia um bom equilíbrio entre as métricas de precisão e *recall*, garantindo uma capacidade sólida de classificação em ambas as direções. Isso pode ser especialmente útil para a detecção precoce de fraudes, permitindo a tomada de ações rápidas para evitar prejuízos. Esse resultado corrobora nossa hipótese de que a combinação de diferentes algoritmos pode ser benéfica para a detecção de fraudes.

Por outro lado, a Regressão Logística e o *Gradient Boosting Classifier* apresentaram desempenho inferior em relação à precisão e *recall*. Embora esses algoritmos ainda tenham desempenhos aceitáveis, eles podem requerer ajustes ou a utilização de técnicas adicionais para melhorar sua eficácia na detecção de fraudes.

As duas hipóteses investigadas neste estudo visavam aprimorar a detecção de fraudes em transações de cartão de crédito, abordando aspectos cruciais da segurança financeira. A primeira hipótese sugeriu que o custo associado à verificação de falsos positivos é consideravelmente maior do que o prejuízo causado pelas fraudes não detectadas. Em outras palavras, o impacto financeiro e operacional causado por transações legítimas erroneamente marcadas como fraudulentas pode ser significativo, mas provou não ser maior do que o prejuízo resultante de fraudes não identificadas, uma vez que conseguimos por meio dos algoritmos identificar corretamente 100% das transações consideradas fraudes verdadeiras.

A análise dos resultados confirmou essa hipótese, enfatizando a importância de minimizar falsos positivos para otimizar o desempenho dos sistemas de detecção de fraudes. Um modelo de detecção de fraudes excessivamente conservador, que rotula muitas transações legítimas como fraudulentas, pode levar a custos desnecessários em investigações e perdas de negócios. Portanto, é fundamental encontrar um equilíbrio entre a precisão na detecção de fraudes e a minimização de falsos positivos.

Por outro lado, a segunda hipótese explorou a eficácia da combinação de diferentes tipos de algoritmos e técnicas de Aprendizagem de Máquina para potencializar a detecção de fraudes em cartões de crédito. Essa abordagem permitiu aumentar a robustez do sistema de detecção de fraudes, possibilitando que diferentes técnicas trabalhassem em conjunto para identificar padrões e anomalias que poderiam passar despercebidas por algoritmos individuais. A combinação dessas proporcionou um equilíbrio entre a precisão na detecção de fraudes e a capacidade de recuperar um maior número de transações fraudulentas, aumentando o *recall*.

Desta maneira, conclui -se que os resultados confirmaram a eficácia das duas hipóteses propostas. Os modelos de Floresta Aleatória e Perceptron Multicamadas se destacaram na detecção de fraudes em transações de cartão de crédito, contribuindo para a segurança e confiabilidade dos sistemas financeiros. Além disso, a utilização de múltiplos algoritmos mostrou-se uma abordagem bem-sucedida para melhorar a capacidade de identificar transações fraudulentas. No entanto, é importante ressaltar que a detecção de fraudes é uma tarefa contínua e desafiadora, uma vez que os fraudadores estão constantemente desenvolvendo novas táticas para evadir os sistemas de segurança. Portanto, é essencial continuar aprimorando e atualizando os modelos de detecção, incorporando técnicas avançadas de Aprendizagem de Máquina e mantendo-se atualizado com as tendências e desenvolvimentos em segurança financeira. Somente assim será possível garantir a eficácia contínua dessas soluções no enfrentamento dos desafios impostos pelas atividades fraudulentas.

Referências

ABECS. 2022. Disponível em: https://api.abecs.org.br/wp-content/uploads/2023/02/Abecs-Apresentacao-2022.pdf.

AL, Q. B. et. What is machine learning? a primer for the epidemiologist. v. 188, 2019. Https://doi.org/10.1093/aje/kwz189.

BHATLA, T. P.; PRABHU, V. C.; DUA, A. Understanding credit card frauds. In: . [S.l.: s.n.], 2003.

BOLTON, R. J.; HAND., D. J. Statistical fraud detection: A review. v. 17, n. 3, p. 235–255, 2002.

HASTIE T.; TIBSHIRANI, R. F. J. The elements of statistical learning: Data mining, inference and prediction. v. 2, 2008. Https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.158.8831rep=rep1type=pdf.