Revista Brasileira de Computação Aplicada, Abril, 2023

DOI: 10.5335/rbca.v15i1.13790
Vol. 15, Nº 1, pp. 1–11
Homepage: seer.upf.br/index.php/rbca/index

TUTORIAL

Detecção de fraude de cartão de crédito por meio de algoritmos de aprendizado de máquina

Credit card fraud detection using machine learning algorithms

Daniel H. M. de Souza[1], and Claudio J. Bordin Jr.[1]

[1]Universidade Federal do ABC (UFABC)

*daniel.henrique@aluno.ufabc.edu.br; claudio.bordin@ufabc.edu.br

Recebido: 20/08/2022. **Revisado:** 04/11/2023. **Aceito:** 25/04/2023.

Resumo

Neste trabalho, descreve-se um tutorial para resolução do problema de fraude sob um contexto de aprendizado supervisionado em aprendizado de máquina, sendo este tutorial composto por um conjunto de metodologias que possibilitam a construção de um modelo de reconhecimento de transações fraudulentas em pagamentos via cartão de crédito. Para isso, primeiramente é abordado o conceito de fraude em meios de pagamento, suas consequências, e a importância do reconhecimento deste tipo de transação para mitigação de risco. Em seguida, é descrito o problema de aprendizado supervisionado, a partir de uma revisão bibliográfica abordando os principais conceitos desta área, principais aplicações e métodos de avaliação de desempenho dos modelos utilizados para tarefas de classificação. É feita também uma revisão da literatura, descrevendo alguns trabalhos em que houve o uso de métodos clássicos e métodos híbridos para detecção de transações fraudulentas. Descrevem-se ainda as principais metodologias para balanceamento de conjuntos de dados que são aplicáveis ao problema em análise. Ao final do trabalho, são feitas as considerações finais, incluindo também algumas possibilidades de estudos para esta área.

Palavras-Chave: Detecção de fraude de cartão de crédito; estatística aplicada; inteligência artificial; mercado financeiro; machine learning.

Abstract

In this work, we describe a tutorial to solve the fraud problem under a supervised learning context in machine learning, and this tutorial consists of a set of methodologies that allow the construction of a model for recognizing fraudulent transactions in payments via card credit. For this, firstly, we explain the concept of fraud in means of payment, its consequences, and the importance of recognizing this type of transaction for risk mitigation is addressed. Then, we describe the supervised learning problem, based on a literature review covering the main concepts of this area, main applications and performance evaluation methods of the models used for classification tasks. Then, we do a literature review, describing some works in which classical and hybrid methods were used to detect fraudulent transactions. We also describe the main methodologies for balancing datasets that are applicable to the problem under analysis. At the end of the work, we bring the final considerations, also including some possibilities of studies for this area.

Keywords: Applied Statistics; artificial intelligence; credit card fraud detection; financial market; machine learning.

1 Introdução

O setor de comércio tem enorme importância no cenário econômico, representando cerca de 61,13% do Produto

Interno Bruto (PIB) no mundo em 2020 (O'NEILL, n.d.). Pela representatividade das operações relacionadas ao comércio, existe uma grande preocupação com a garantia de qualidade das transações, principalmente quando se trata de operações de crédito (Hu and Su, 2022). Com a expansão de crédito para públicos das mais diversas faixas de renda, houve um aumento no volume de operações bancárias, abrangendo parcela representativa da população. Esta expansão trouxe consigo um aumento significativo do número de transações realizadas, e consequentemente, um aumento na exposição destas operações a fraudes (Hu and Su, 2022).

Dentre os produtos de crédito, o cartão de crédito é um dos mais visados pelos fraudadores, vista a simplicidade de fraudar transações envolvendo altos valores financeiros, além da demora da descoberta da operação fraudulenta pela instituição financeiro e pelo cliente (Gupta et al., 2021). O número de operações fraudulentas relacionadas a cartão de crédito é tão grande que o impacto financeiro dessas em 2020 atingiu o valor de 32,39 Bilhões de dólares no mundo (Savy, 2020). Apesar da expressividade deste número, isto ainda não é o fator mais preocupante, mas sim a tendência de crescimento do número de casos de fraude na modalidade de cartão de crédito, e consequentemente, o impacto financeiro disto.

No estudo feito por Savy (2020), tem-se em 2020 um impacto financeiro 3 vezes maior do que o impacto desta mesma atividade em 2011, que foi de aproximadamente 9,84 bilhões de dólares. Além disso, é previsto que este número chegue a 40,63 bilhões em 2027, com uma proporção financeira de 5,68 centavos para cada 100 dólares, em relação à casos de fraude (Savy, 2020). Diante disso, modelos [1] estatísticos para reconhecimento de transações fraudulentas são de extrema importância para mitigar o risco neste tipo de operação (Cherif et al., 2022). O uso deste tipo de abordagem possibilita a análise e estimação de possibilidade de fraude em tempo real (Zhang et al., 2021). A garantia da autenticidade dessas operações é de extrema importância para o mercado financeiro, dado o impacto causado em caso de transações fraudulentas.

No atual cenário, existe uma predominância do uso de técnicas de *Machine Learning* (ML) com uma abordagem de aprendizado supervisionado para modelagem de fraude em operações com cartão de crédito, com destaque para modelos baseados em Árvores de Decisão (Sailusha et al., 2020), Redes Bayesianas (Itoo et al., 2021) e Redes Neurais (Roseline et al., 2022), além de métodos combinados entre estes classificadores (Feng et al., 2020).

Em vista desses fatos, este artigo tem como intuito trazer uma revisão da literatura, contendo um tutorial para detecção de fraude em operações realizadas via cartão de crédito a partir de algoritmos de aprendizado de máquina que correspondem ao estado da arte para este tipo de aplicação. Além disso, este artigo traz as principais metodologias utilizadas em cada etapa de modelagem a partir de algoritmos de aprendizado de máquina, incluindo os principais classificadores utilizados, abordagens alternativas com combinação de classificadores, métodos de balanceamento de classes para melhorar o desempenho dos algoritmos, além de métricas de avaliação de desempenho para medir o acerto dos algoritmos. O objetivo deste tutorial é trazer para pesquisadores de diversas áreas uma visão geral sobre o impacto do problema de fraude, além de uma descrição metodológica dos principais processos para construção de soluções de detecção de fraude, incluindo as dificuldades processuais destes.

O texto a seguir está organizado da seguinte forma: na Seção 2, é feita uma descrição sobre o conceito e funcionamento do cartão de crédito. Em seguida, na Seção 3, é feita uma breve descrição dos principais tipos de fraude de cartão de crédito. Na Seção 4, é feita uma descrição processual de como utilizar algoritmos de aprendizado de máquina, sendo essa também uma seção introdutória a alguns métodos que serão explicados posteriormente. Na Seção 5, é o descrito o problema de aprendizado supervisionado, com ênfase em reconhecimento de padrões a partir de classificadores. Além disso, descrevem-se os principais métodos para avaliação de desempenho de algoritmos de classificação. Na Seção 6, serão descritas as principais metodologias utilizadas para detecção de fraude, versando a abordagem clássica com algoritmos de classificação e o uso de métodos *ensembles*. Na Seção 7, descrevem-se os principais métodos de balanceamento de classes para mitigar o viés dos algoritmos clássicos em não detectar transações fraudulentas. Por fim, na Seção 8, são feitas considerações finais acerca do conteúdo exposto.

2 Cartão de Crédito

Dentre os diversos meios de pagamento, destacam-se as transações realizadas por meio do cartão de crédito. Esse serviço de pagamento é oferecido por inúmeras instituições financeiras, com destaque para os grandes bancos, existindo também diversas empresas do setor de serviços que oferecem cartões de crédito específicos para facilitar suas transações (Jachemet, 2018).

O cartão de crédito é um registro da intenção de pagamento do usuário, autorizado através de verificações de assinatura, cadastro, senha e outras informações pertinentes. Quando autorizado, o consumidor fica responsável por pagar os custos gerados em espécie, débito em conta, dentre outras formas. Existem diversas modalidades de cartões de crédito definidas quanto ao uso que, em linhas gerais, são combinações entre uso nacional, internacional, e programas de vantagens que possibilitam descontos em produtos domésticos, passagens aéreas, dentre outros (Jachemet, 2018).

Todo o mercado de cartão de crédito é baseado na interação entre diversos *stakeholders* (Jachemet, 2018), dentre os quais pode-se destacar o portador (*card holder*), estabelecimento (*merchant*), adquirente (*acquirent*), bandeira (*brand*) e emissor (*issuer*). Segundo Jachemet (2018), esses agentes podem ser descritos da seguinte forma: portador é a pessoa que possui o cartão de crédito e inicia a operação de pagamento ao estabelecimento recebedor do valor definido; a supervisão, gerenciamento e repasse do valor de pagamento é feita pelo adquirente ao estabelecimento; as operações das redes de comunicação e políticas de relacionamento entre emissores e adquirentes é responsabilidade das bandeiras (Visa, Mastercard, Amex, dentre outras); a

[1] Note que na literatura da área é comum o uso do termo "modelo" para se referir ao estimador utilizado, sendo este um classificador, regressor, ou agrupador.

emissão do cartão de crédito é realizada pelo órgão emissor, que em geral, são os grandes bancos.

O fluxo dos pagamentos via cartão de crédito funciona da seguinte forma: ao utilizar o cartão de crédito (presencialmente ou pela internet), é enviado um sinal para adquirente, que é repassado para a bandeira, e em seguida para o emissor do cartão. O emissor tem uma série de políticas de crédito para aprovar ou não a transação, como, por exemplo, uma verificação de consistência cadastral, checagem de limite, atrasos, histórico de operações fraudulentas, dentre outros aspectos. Ao final dessa checagem, o emissor envia o parecer de aprovação ou negação da transação ao estabelecimento, encerrando o fluxo de pagamento, que leva em torno de dez segundos (Jachemet, 2018).

A modelagem de fraude atua na etapa de checagem de políticas de crédito para aprovação ou não da transação, visando reconhecer padrões de comportamentos fraudulentos na operação em questão (Thomas et al., 2017). A seguir, são detalhes os tipos mais comuns de fraude de cartão de crédito.

3 Tipos de Fraude

A popularização do cartão de crédito e aumento exponencial de transações desse tipo torna cada vez mais difícil o controle dessas operações. Apesar dos inúmeros benefícios da expansão da internet, a consequência foi a popularização de pagamentos *online* pelo cartão de crédito, ou seja, compras sem a necessidade do cartão físico. Esse cenário torna muito mais simples a realização de operações fraudulentas (Sailusha et al., 2020).

Tratando-se de operações fraudulentas, existem dois tipos mais utilizados de fraude de cartão de crédito: *Application Fraud* e *Behaviour Fraud*. *Application Fraud* consiste na aplicação de fraude por meio de falsificação de informação. Para esse caso, os fraudadores falsificam informações legítimas dos titulares dos cartões, e assim, recebem novos cartões de crédito do emissor. Já no *Behaviour Fraud*, os criminosos roubam o cartão e senha de um titular de um cartão genuíno e utilizam-no para efetuar compras, geralmente até utilizar todo o limite disponível (Zhang et al., 2021).

Em geral, utilizam-se dois métodos de detecção de fraude (Zhang et al., 2021): detecção de uso indevido do cartão de crédito e detecção de anomalias quanto ao uso do mesmo. Em relação ao uso indevido do cartão de crédito, avaliam-se aspectos como dados cadastrais e outros particularidades que possibilitem investigar a legitimidade do usuário que realizou a operação. Já se tratando da detecção de anomalias, verificam-se questões como o perfil de compra do usuário, gasto médio em determinado período, além de outros pontos que possibilitem comparar a compra analisada com o histórico de compras constantes do banco de dados, visando detectar anormalidades.

Assim, é conveniente que os modelos utilizados na detecção de transações fraudulentas sejam treinados de modo a reconhecer padrões dos mais diversos tipos de fraude (Cherif et al., 2022).

4 Detecção de Fraude de Cartão de Crédito

Segundo MICHAELIS (n.d.), fraude consiste em "qualquer ato ardiloso, enganoso, de má-fé, com o intuito de lesar ou ludibriar outrem, ou de não cumprir determinado dever". Como mencionado na Seção 3, a cada ano há um aumento expressivo do número de operações fraudulentas relacionadas a cartão de crédito dada a facilidade de fraudar este tipo de transação. Diante desse cenário, existe um grande esforço por parte de grandes empresas e pesquisadores em desenvolver modelos para o reconhecimento de comportamentos fraudulentos (Hilal et al., 2022).

Existem diversas técnicas para detecção de fraude, com destaque para técnicas de aprendizado de máquina, especificamente em um contexto de classificação (Zhang et al., 2021).

Um procedimento básico para realização da tarefa de detecção de fraude pode ser resumido a partir do Algoritmo 1.

Algoritmo 1: Procedimento para Detecção de Fraude de Cartão de Crédito com classificadores

1. Selecione um conjunto de dados para treinamento do classificador a ser implementado (Randhawa et al., 2018)
2. Após o pré-processamento do conjunto de dados, aplique uma técnica de reamostragem (Sohony et al., 2018) para eliminar o viés do algoritmo em não encontrar transações fraudulentas
3. Aplique alguma metodologia para detecção de fraude (Husejinovic, 2020)
4. Valide o desempenho do algoritmo utilizado a partir de métricas de avaliação (Vujović, 2021)

As principais metodologias utilizadas para realização de cada um destes passos são descritas a seguir.

5 Aprendizado Supervisionado e o Problema de Classificação

As técnicas de ML são uma das grandes categorias dos métodos de Inteligência Artificial. Estas técnicas consistem na construção de sistemas que possuem certa capacidade para aprender de forma automática, de modo a conseguir tomar suas próprias decisões com o mínimo de intervenção humana (Vujović, 2021). A metodologia de ML baseia-se no paradigma de aprendizado indutivo, em que, a partir de inferência lógica, os modelos matemáticos desenvolvidos possam aprender padrões e generalizar um determinado conceito de interesse. O aprendizado indutivo pode ser segmentado conforme o mostrado na Fig. 1 (Bochie et al., 2020).

Nos métodos de aprendizado *supervisionado*, o aprendizado indutivo é realizado a partir de um conjunto de exemplos de treinamento com a saída conhecida, o que se define como *base rotulada* (Bochie et al., 2020), ou seja, é fornecida ao modelo uma base de treinamento com variáveis explicativas, bem como a variável de resposta com um rótulo conhecido para aquele cenário, com objetivo de que o modelo treinado possa generalizar e abstrair este tipo de conhecimento para exemplos não rotulados. Dentro dos métodos de aprendizado supervisionado, há uma se-

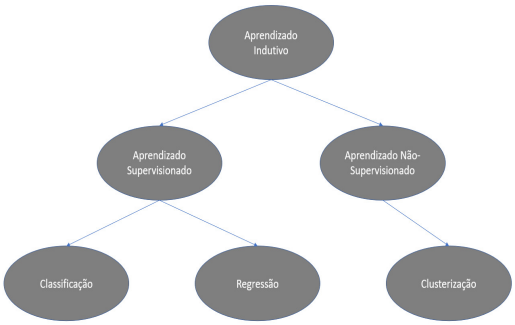

Figura 1: A hierarquia do aprendizado (Bochie et al., 2020)

gunda segmentação, podendo-se dividi-los em problemas de classificação (Vujović, 2021) e problemas de regressão (Nabipour et al., 2020). O que diferencia os dois é que, em problemas de classificação, o intuito é a predição de uma variável de resposta categórica e, em problemas de regressão, a variável de resposta é contínua.

Segundo Monard and Baranauskas (2003), um *conjunto de exemplos* é um conjunto de dados que possui os atributos, classes e seus respectivos valores associados. Este conjunto é utilizado para treinamento e análise de desempenho do classificador escolhido. Tome-se como exemplo a Tabela 1.

Tabela 1: Conjunto de exemplos no formato atributo-valor (Monard and Baranauskas, 2003)

	X_1	X_2	$\cdots$	X_m	Y
T_1	x_{11}	x_{12}	$\cdots$	x_{1m}	y_1
T_2	x_{21}	x_{22}	$\cdots$	x_{2m}	y_2
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
T_n	x_{n1}	x_{n2}	$\cdots$	x_{nm}	y_n

Na referida tabela, formam-se as tuplas $T_i = (x_{i1}, x_{i2}, \cdots, x_{im}, y_i) = (\vec{x_i}, y_i)$, em que a função $y_i = f(\vec{x_i})$ tenta predizer a classe y_i a partir do conjunto de atributos $\vec{x_i}$.

Diversos problemas práticos podem ser abordados através de métodos de aprendizado supervisionado. Em casos em que o interesse é a predição de uma variável categórica, há por exemplo a aplicação de algoritmos de classificação para diagnóstico de pacientes médicos (Ahmad et al., 2022), classificação de clientes por risco financeiro (Moscato et al., 2021), reconhecimento de mensagens de *spam* (Junnarkar et al., 2021), reconhecimento de imagens (Pengyu and Wanna, 2021), reconhecimento de voz (Ali et al., 2021) e detecção de fraude no mercado de crédito (Hilal et al., 2022), dentre outras possíveis aplicações. Há também a possibilidade de modelar problemas das mais diversas áreas da ciência por meio de problemas de regressão, com foco na estimação de séries temporais, como por exemplo a predição de preços de ações e outros índices do mercado financeiro (Nabipour et al., 2020), previsão de variáveis climáticas (Murugan Bhagavathi et al., 2021), previsão de demanda (Sharma et al., 2021), dentre outros.

Em contraste aos métodos supervisionados, os métodos de aprendizado *não supervisionado* tem o objetivo de treinar modelos com dados não-rotulados. Neste caso, espera-se que o modelo aprenda o comportamento das variáveis explicativas selecionadas e construa agrupamentos (*clusters*) a partir das observações conforme as suas similaridades (Jain et al., 2021). Esses métodos não serão abordados neste artigo.

Na próxima seção, são detalhados os principais métodos para a validação de algoritmos de classificação supervisionados.

5.1 Validação de Algoritmos de Classificação

Para a validação de algoritmos em aprendizado supervisionado, usualmente, é realizada a divisão do conjunto de dados rotulados em um conjunto de *treino* e um conjunto de *teste* (Costa, 2019). A ideia é que o algoritmo seja treinado a partir da base de treinamento e, na base de teste, seja medido o desempenho do modelo a partir de métricas de erro. Este método também é chamado de *validação cruzada*.

Em bases de dados muito grandes, o método de validação cruzada pode ser modificado com a adoção de um terceiro conjunto, definido como *conjunto de validação*. Neste caso, o conjunto de teste é utilizado para a escolha de modelos, seleção de variáveis e otimização de hiperparâmetros; o conjunto de validação, por sua vez, é usado para a otimização das métricas de erro do modelo escolhido (Moghaddam et al., 2016).

Ao se ajustar um modelo de ML, ao invés de ocorrer um ajuste adequado, podem ocorrer *overfitting* ou *underfitting* (Zhang et al., 2019). Quando há *overfitting*, o modelo memoriza os dados de treinamento e perde o poder de generalização para conjuntos de dados diferentes dos dados de treinamento. Em outras palavras, pode-se dizer que o modelo é mais flexível do que deveria ser. O mesmo pode ocorrer por outras razões como, por exemplo, um alto número de variáveis que não têm poder de explicação da variável resposta, poucos dados no conjunto de treinamento, ou, até, características particulares do modelo escolhido (Zhang et al., 2019). Uma característica muito comum nestes casos é um erro muito baixo no conjunto de treinamento e um erro alto nos dados de teste.

Já *underfitting* ocorre quando o modelo não foi capaz de estimar a variável resposta, ou seja, houve um baixo ajuste ao conjunto de treinamento, gerando assim um alto erro tanto no conjunto de treinamento, quando no conjunto de teste (Zhang et al., 2019).

O ajuste adequado ocorre em casos em que, apesar de alguns pontos não serem bem ajustados pelo modelo, este consegue generalizar e estimar bem a maioria das observações. Neste caso, os resultados apresentam um baixo erro tanto no conjunto de treinamento quanto no conjunto de teste.

Diante disso, pode-se dizer que os problemas de ML são problemas matemáticos de otimização, em que se busca otimizar uma função de custo, que no caso é o erro produzido aos se estimar os dados dos conjuntos de teste/validação (Bochie et al., 2020).

Em problemas de classificação, utiliza-se a chamada

matriz de confusão (Vujović, 2021), ilustrada na Tabela 2 para avaliações de desempenho. O elemento $M\left(C_i, C_j\right)$ da matriz de confusão é definido como a contagem do evento em que a classe verdadeira é C_i e a classe estimada é C_j, i.e.,

$$M(C_i, C_j) = \sum_{\{(\vec{x},y)\in T \mid y=C_i\}} \mathcal{I}\left\{h(\vec{x}) = C_j\right\}, \qquad (1)$$

em que T denota o conjunto de treinamento, $h(\vec{x})$ denota a decisão produzida pelo modelo para os atributos $\vec{x}$ e $I\{A\}$ denota a função indicador da condição A, que vale 1 caso esta seja verdadeira e 0 caso contrário.

Tabela 2: Matriz de Confusão (Vujović, 2021)

Classe	predita C_1	predita C_2	$\cdots$	predita C_k
verdadeira C_1	$M(C_1, C_1)$	$M(C_1, C_2)$	$\cdots$	$M(C_1, C_k)$
verdadeira C_2	$M(C_2, C_1)$	$M(C_2, C_2)$	$\cdots$	$M(C_2, C_k)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
verdadeira C_k	$M(C_k, C_1)$	$M(C_k, C_2)$	$\cdots$	$M(C_k, C_k)$

Observe que, na matriz de confusão, os acertos aparecem na diagonal principal e os erros de classificação nas entradas com $i \neq j$. Para problemas de classificação binária, os dois eventos possíveis podem ser rotulados como positivo (P) e negativo (N), resultando numa matriz de confusão com quatro elementos: verdadeiros positivos (V_P), verdadeiros negativos (V_N), falsos positivos (F_P) e falsos negativos (F_N) (Tabela 3).

Tabela 3: Matriz de Confusão Binária (Vujović, 2021)

Classe	preditiva C_+	preditiva C_-
verdadeira C_+	V_P	F_N
verdadeira C_-	F_P	V_N

A partir das entradas da matriz de confusão binária, podem-se definir métricas de erro úteis para avaliação de algoritmos, com destaque para sensitividade (sens), especificidade (spec) e *Balanced Classification Rate* (bcr), dadas pelas expressões:

$$\text{sens} = \frac{V_P}{V_P + F_N}, \qquad (2)$$

$$\text{spec} = \frac{V_N}{V_N + F_P}, \qquad (3)$$

$$\text{bcr} = \frac{1}{2}\left(\frac{V_P}{V_P + F_N} + \frac{V_N}{V_N + F_P}\right) = \frac{1}{2}\left(\text{sens} + \text{spec}\right). \qquad (4)$$

Estas métricas representam, para o problema em questão, a taxa de acerto da classe positiva (transações fraudulentas), a taxa de acerto da classe negativa (transações genuínas), e a taxa média de acerto entre ambas as classes.

6 Uso de Classificadores para Detecção de Fraude

Dentre os principais classificadores, os métodos mais explorados são modelos baseados em Árvores de Decisão (Sailusha et al., 2020), *Redes Bayesianas* (Itoo et al., 2021) e em Redes Neurais (Roseline et al., 2022).

Em Maes et al. (2002), realizou-se um estudo comparativo do desempenho de Redes Bayesianas e Redes Neurais na estimação de fraudes com cartão de crédito, utilizando para o estudo uma base com transações reais do mercado brasileiro. Neste estudo, as Redes Bayesianas apresentaram melhor desempenho em relação às Redes neurais, levando em consideração as principais métricas de análise de desempenho de classificadores.

Em Sailusha et al. (2020), realiza-se um estudo de fraude com cartão de crédito modelando o problema como um problema de classificação. Neste trabalho, o autor foca na modelagem por meio de métodos baseados em Árvores de Decisão, mostrando a viabilidade dos algoritmos *Random Forest* e *Gradient Boosted Tree* para tal fim.

As referências Roseline et al. (2022) e Zhang et al. (2021) mostraram a viabilidade do uso de Redes Neurais no reconhecimento de transações fraudulentas de cartões de crédito.

Em Makki et al. (2019), realizou-se uma análise comparativa entre os modelos de Rede Neural, Árvore de Decisão, *K-nearest Neighbors*, Regressão e *Support Vector Machines* para análise de operações fraudulentas de cartão de crédito. Nessa comparação, houve um melhor desempenho dos modelos *Support Vector Machines* e Regressão Logística.

Em Itoo et al. (2021), tratou-se o problema de fraude em operações de crédito utilizando os modelos *Naive Bayes*, Regressão Logística e KNN. Nesse trabalho, constatou-se o melhor desempenho do método *K-Nearest Neighbors* para esse tipo de problema.

Outra abordagem para este tipo de problema é a modelagem de fraude com cartão de crédito por meio do chamado "Aprendizado Não-Supervisionado", utilizando modelos de Clusterização/Agrupamento (*Clustering*) (Jain et al., 2021). Este tipo de abordagem é utilizada principalmente em cenários em que não há o rótulo de fraude/não-fraude nas transações disponíveis no conjunto de dados analisado, ou ainda quando o conjunto de dados rotulado é pequeno. Porém, vale ressaltar que quando se tem um grande conjunto de dados rotulados disponíveis, esta metodologia mostra um desempenho menor do que uma abordagem de aprendizado supervisionado, sendo este o principal fator para este método não ser tão explorado (Khatri et al., 2020).

Neste contexto, Jain et al. (2021) utiliza uma abordagem de clusterização para reconhecer anomalias em operações de cartão de crédito através do modelo *K-means*. Já Dharwa and Patel (2011) utiliza o algoritmo *Density-based spatial clustering of applications with noise* (DBSCAN) para uma abordagem parecida.

Ainda nesta abordagem de aprendizado não-supervisionado, Sabau (2012) propõe a utilização do modelo de agrupamento *Hierarchical Clustering* para reconhecer padrões em fraude com cartão de crédito, obtendo bons resultados e mostrando a viabilidade deste tipo de metodologia principalmente em casos em que não há um grande número de dados.

Em Maes et al. (2002), apontam-se as seguintes características como essenciais para um bom modelo de reconhecimento de fraude:

- Capacidade de lidar com distribuições assimétricas, dado que as operações fraudulentas são uma pequena parcela do total de operações realizadas;
- Capacidade de tratar *outliers*;
- Adaptabilidade do sistema para novos tipos de fraude;
- Bom custo benefício. O modelo deve ter um custo de implantação que faça sentido em relação ao valor agregado das predições do mesmo;
- Métricas adequadas para avaliação do desempenho do modelo.

6.1 Detecção de Fraude com Classificadores Agregados (*ensembles*)

Na seção anterior, foi explorada em alguns trabalhos a utilização de classificadores para detecção de fraude. Porém, não existem modelos perfeitos, visto que, devido às premissas assumidas na construção de cada modelo, é inevitável que estes apresentam certa fragilidade (Saxena et al., 2021). Além da fragilidade natural de cada modelo de aprendizado de máquina, a crescente complexidade dos problemas ao longo do tempo torna necessária a construção de métodos aprimorados para modelagem de estruturas de dados (Saxena et al., 2021). Uma alternativa para minimizar este efeito é a utilização de métodos combinados (Wang and Liu, 2021)(Forough and Momtazi, 2021)(do Amaral et al., 2021), os quais serão detalhados a seguir.

Uma proposta para melhoria de desempenho no processo de aprendizado de máquina é a abordagem combinada de modelos de aprendizado de máquina, também conhecida como *ensembles* (Wang and Liu, 2021). Este tipo de metodologia visa fortalecer o poder preditivo e diminuir o viés de uma classe para o caso em questão (Sudha and Akila, 2021).

Este tipo de metodologia é bem comum utilizando modelos da mesma classe. Têm-se por exemplo os métodos *ensembles* baseados em árvores de decisão , em que são construídas árvores de decisão paralelamente e a classificação é feita por meio de voto majoritário em relação à classificação de cada árvore (*Bagging Classification Trees* e *Random Forest*), ou têm-se árvores de decisão construídas sequencialmente, a partir do resíduo da árvore anterior (*Gradient Boosted-Tree* (Sailusha et al., 2020)). Têm-se ainda as próprias redes neurais, que consistem em conjuntos de neurônios artificiais combinados, visando a resolução de problemas de difícil abstração (Roseline et al., 2022).

Segundo do Amaral et al. (2021), a utilização de métodos combinados originados da agregação de classificadores idênticos/parecidos leva a um erro de predição em massa quando o método em questão não é o mais adequado para a estimação do objeto em análise. Para minimizar este efeito, existe também a abordagem de combinação de modelos distintos, como por exemplo Regressão logística e métodos baseados em Árvores de Decisão, Redes Neurais e KNN, Redes *Bayesianas* e Redes Neurais, dentre outras possibilidades. Para este fim, a classificação final é realizada por meio de algum método de agregação entre as classificações de cada método, sendo o mais utilizado a classificação por voto majoritário (Sudha and Akila, 2021).

O funcionamento da agregação de classificadores por meio de voto majoritário pode ser descrito pelo Algoritmo 2.

Algoritmo 2: Agregação por Voto Majoritário

1. **for** $i = 1$ **to** n:
1.1. **for** $j = 1$ **to** m:
1.1.1. Classifique o objeto i utilizando o classificador M_j, sendo o resultado definido como y_{ij}
1.2 **end for**
1.3. y_i = Classe mais frequente entre todas as classificações para i, ou seja, $y_i = MODA\{y_{i1}, y_{i2}, \cdots, y_{im}\}$
2. **end for**
3. **return** $y = \{y_1, y_2, \cdots, y_n\}$

7 Amostragem e *Data Augmentation*

Os algoritmos de aprendizado de máquina baseiam-se no reconhecimento de padrões a partir de dados, possibilitando assim a resolução de problemas complexos. Porém, neste cenário, existe uma grande limitação quanto à disponibilidade de dados para uso no treinamento dos algoritmos em questão (Shorten and Khoshgoftaar, 2019).

Este problema se agrava em casos em que os dados possuem naturalmente característica de desbalanceamento entre classes, como por exemplo problemas de detecção de anomalias em redes elétricas, anomalias em tráfego de rede de internet, sequenciamento de DNA, diagnóstico médico, problemas de segurança, problemas de imagem, detecção de fraude, dentre outras possibilidades (Tarekegn et al., 2021).

O desbalanceamento de classes gera diversas consequências negativas no treinamento/desempenho em grande parte dos algoritmos de aprendizado de máquina, principalmente em problemas de classificação binária, sendo a principal consequência a geração de um viés em classificar a maioria das instâncias com o rótulo da classe majoritária (Tarekegn et al., 2021).

Diante deste contexto, existem diversas abordagens para geração de dados da classe minoritária, visando treinar os modelos com uma proporção de 50% entre as classes, e assim, reduzir o erro na classificação (Tarekegn et al., 2021). A seguir, serão descritas de forma breve as principais metodologias para este fim.

7.1 Metodologias Clássicas de Amostragem

As abordagens clássicas se baseiam na superamostragem da classe minoritária e subamostragem da classe majoritária (Sohony et al., 2018). Na subamostragem, realiza-se uma amostragem aleatória da classe majoritária sem reposição, até que se tenha uma proporção igual entre esta e a classe minoritária (Sohony et al., 2018). Já na superamostragem, são utilizadas técnicas para gerar dados a partir da classe minoritária de forma consistente. As mais conhecidas dentre estas são *Synthetic Minority Over-sampling Technique* (SMOTE) e *Random Over Sampling* (Sohony et al.,

2018).

O método SMOTE é uma técnica de aumento de dados (*Data Augmentation* (Shorten and Khoshgoftaar, 2019)), para geração de novas instâncias da classe minoritária a partir da combinação entre instâncias desta classe a partir do algoritmo KNN, construindo assim amostras sintéticas. A ideia é que se escolha aleatoriamente uma amostra da classe minoritária, e a partir deste exemplo, selecionam-se os k vizinhos mais próximos. Dentre estes, escolhe-se um vizinho aleatoriamente, e gera-se um novo exemplo a partir de uma combinação entre a amostra inicial e o vizinho selecionado. Isto é feita até que as classes estejam balanceadas (Claro et al., 2020). Já na metodologia *Random Over Sampling*, as amostras da classe minoritária são replicadas aleatoriamente, gerando cópias destas instâncias (Sohony et al., 2018).

Quanto ao desempenho destas, Sohony et al. (2018) mostra que as técnicas de superamostragem da classe minoritária (*Random Over Sampling* e SMOTE) apresentam como vantagem um ganho de desempenho em relação à sub-amostragem da classe majoritária (*undersampling*). Porém, Sohony et al. (2018) defende que, ao utilizar o *undersampling*, o treinamento dos classificadores é expressivamente mais rápido.

Já a referência Baesens et al. (2021) defende que não há um consenso quanto ao melhor desempenho entre o SMOTE e o *Random Over Sampling*, tendo em vista que quase não há diferença em desempenho quando estes são comparados. Baesens et al. (2021) mostra um desempenho ligeiramente melhor de um método ou outro a depender das características dos conjuntos de dados utilizados, classificadores, etc.

7.2 Metodologias Alternativas de *Data Augmentation* com Redes Neurais

Com o aumento do poder computacional, algoritmos baseados em redes neurais têm sido desenvolvidos para diversos fins, visando a resolução de problemas de difícil resolução (Bochie et al., 2020). Em meio a este cenário, pesquisadores desenvolveram metodologias para geração de dados sintéticos (*Data Augmentation*) a partir de redes neurais. Em trabalhos de detecção de fraude, o intuito do uso deste tipo de metodologia é a utilização de algoritmos que possam aprender e gerar dados de fraude, possibilitando aumentar a quantidade de amostras da classe minoritária no conjunto de dados de treinamento, e consequentemente, mitigar o desbalanceamento de classes (Tingfei et al., 2020)(Fiore et al., 2019).

Dentre as metodologias para este fim, destacam-se os algoritmos *Generative Adversarial Network*/Redes Adversárias Generativas (GAN) e *Variational Autoencoder* (VAE) (Burks et al., 2019), os quais serão descritos a seguir.

7.2.1 *Variational AutoEncoder*

AutoEncoders são técnicas de aprendizado de máquina baseadas em redes neurais que operam sob uma abordagem não-supervisionada, visando o aprendizado de distribuições dos dados originais e posterior reprodução destas distribuições (Burks et al., 2019).

Este tipo de rede opera copiando as entradas para as saídas, compactando estas entradas em uma representação de espaço latente, e posteriormente, reconstruindo a saída desta representação. Isto é feito a partir de três componentes principais (Dong et al., 2018):

- Codificador (*Encoder*): conjunto de filtros lineares *feed-forward* com o intuito de compactar a entrada em uma representação de espaço latente;
- Ativação: um mapeamento não linear que transforma os coeficientes codificados para números no intervalo [0,1];
- Decodificador (*Decoder*): Conjunto de filtros lineares reversos que têm como objetivo reconstruir a entrada da representação do espaço latente.

Este tipo de rede possui diversas variações, e possui aplicações tanto para tarefas de Redução de Dimensionalidade (Alsenan et al., 2020), quanto para *data augmentation* (Mu and Chen, 2022). Para projetos que demandam estes dois tipos de tarefas, destaca-se uma modalidade de *AutoEncoder* chamada *Variational Auto Encoder* (Dong et al., 2018).

Em linhas gerais, pode-se dizer que o algoritmo VAE é um *AutoEncoder* que utiliza métodos explícitos de regularização para evitar *overfitting* e garantir que o espaço latente seja consistente, gerando assim bons conjuntos de dados. A grande mudança em relação aos *AutoEncoders* convencionais é que ao invés de codificar uma entrada como um único ponto, esta entrada é codificada como uma distribuição do espaço latente, seguindo os seguintes passos para treinamento do modelo (Dong et al., 2018):

i. Codificação da entrada como uma distribuição no espaço latente;
ii. Amostragem de um ponto desta distribuição;
iii. Decodificação do ponto amostrado;
iv. Cálculo do erro de reconstrução, e retropagação deste pela rede.

Este fluxo pode ser representado pela Fig. 2 (Dong et al., 2018), em que X representa os dados de treinamento, X_hat os dados gerados pelo VAE, $N(\mu, \sigma)$ uma distribuição normal com média μ e desvio padrão σ.

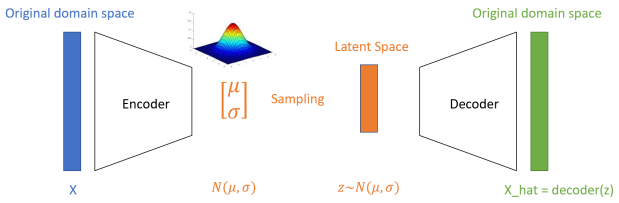

Figura 2: Funcionamento de uma Rede Neural do tipo VAE (Dong et al., 2018)

Na etapa de regularização, as distribuições codificadas são restritas de modo a estarem próximas de uma distribuição normal padrão, em termos de média e variância (Dong et al., 2018). Além disso, tem-se que a busca dos

codificadores e decodificadores que minimizam o erro de reconstrução é feita por meio da técnica de gradiente descendente (Bochie et al., 2020).

Dentre as inúmeras possibilidades de aplicações dos *AutoEncoders* do tipo VAE para resolução de problemas de *Data Augmentation*, destacam-se trabalhos voltados para detecção de vírus em redes de computadores (Burks et al., 2019), análise de sentimentos (Luo et al., 2020), processamento de áudio e texto (Sun et al., 2020), além de resoluções de problemas voltados para detecção de imagem na área da medicina (Zhou et al., 2019).

7.2.2 *Generative Adversarial Network*

As redes neurais do tipo GAN foram abordadas pela primeira vez no artigo de Goodfellow et al. (2014). A ideia por trás deste algoritmo é que uma rede neural seja capaz de aprender e reproduzir uma distribuição de dados qualquer, possibilitando assim a geração de dados fictícios quase que idênticos aos dados de treinamento do algoritmo em questão.

As redes GAN baseiam-se nos conceitos de Treinamento Adversarial, que apoia-se em duas arquiteturas de redes neurais com objetivos distintos (Claro et al., 2020). A primeira delas visa aprender a distribuição dos dados de treinamento e gerar dados com esta mesma distribuição. Já a segunda, é uma "rede rival" que recebe dados reais e dados gerados pela primeira rede, classificando-os em Falso ou Real. Isto é feito visando melhorar o desempenho da rede geradora a partir da detecção de pontos fracos (Moosavi-Dezfooli et al., 2016). O funcionamento deste tipo de rede pode ser representado pela Fig. 3 (Oinar, n.d.).

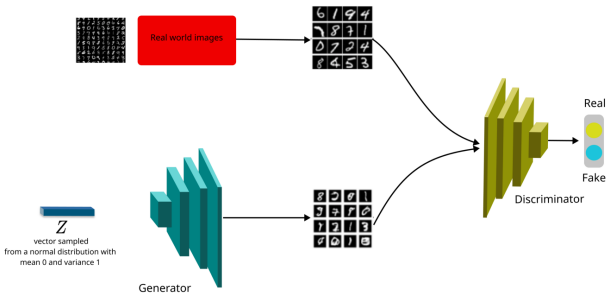

Figura 3: Funcionamento de uma Rede Neural do tipo GAN (Oinar, n.d.)

Este tipo de arquitetura vem sendo utilizado em diversos casos de classificação, como por exemplo para detecção de anomalias em redes de internet (Usama et al., 2019) e problemas de segurança no geral (Yinka-Banjo and Ugot, 2020), análise de sentimento (Shang et al., 2021), além de várias possibilidades de aplicações em computação visual (Peng et al., 2022).

8 Conclusões

Este artigo teve como intuito trazer ao leitor um tutorial para detecção de fraude em operações realizadas via cartão de crédito, sendo este tutorial composto por um conjunto de metodologias que correspondem ao estado da arte para esta aplicação. O objetivo aqui foi de despertar o interesse no tema, além de direcionar pesquisadores quanto aos principais métodos existentes para detecção de fraude e dificuldades para execução desta tarefa com algo grau de desempenho.

Existem ainda diversas dificuldades presentes no reconhecimento de transações fraudulentas, com destaque para a necessidade em mitigar a característica de desbalanceamento presente nos conjuntos de dados de fraude, que traz um viés para os modelos em classificar erroneamente as transações de fraude. Isto proporciona oportunidades de desenvolvimento tanto metodologias de balanceamento de classes para melhoria do treinamento de algoritmos de aprendizado de máquina, quanto a exploração de modelos alternativos para detecção de fraude, sendo estes da categoria de aprendizado supervisionado, aprendizado não-supervisionado, ou ainda, combinações de diferentes formas de aprendizado.

Diante do exposto, tem-se que apesar dos avanços, esta área se mostra muito promissora tanto para a pesquisa científica, quanto para a aplicação em problemas do mercado financeiro, considerando que além da complexidade matemática e computacional dos recursos envolvidos para detecção de fraude, a mitigação da ocorrência de transações fraudulentas possibilita um ganho financeiro para diversas esferas, sendo ainda um aliado no combate a este tipo de crime.

Referências

Ahmad, G. N., Fatima, H., Saidi, A. S. et al. (2022). Efficient medical diagnosis of human heart diseases using machine learning techniques with and without gridsearchcv, *IEEE Access* . https://doi.org/10.1109/access.2022.3165792.

Ali, A. T., Abdullah, H. S. and Fadhil, M. N. (2021). Voice recognition system using machine learning techniques, *Materials Today: Proceedings* . https://doi.org/10.1016/j.matpr.2021.04.075.

Alsenan, S., Al-Turaiki, I. and Hafez, A. (2020). Autoencoder-based dimensionality reduction for qsar modeling, *2020 3rd International Conference on Computer Applications & Information Security (ICCAIS)*, IEEE, pp. 1–4. https://doi.org/10.1109/iccais48893.2020.9096747.

Baesens, B., Höppner, S., Ortner, I. and Verdonck, T. (2021). robrose: A robust approach for dealing with imbalanced data in fraud detection, *Statistical Methods & Applications* **30**(3): 841–861. https://doi.org/10.1007/s10260-021-00573-7.

Bochie, K., da Silva Gilbert, M., Gantert, L., Barbosa, M. d. S. M., de Medeiros, D. S. V. and Campista, M. E. M. (2020). Aprendizado profundo em redes desafiadoras: Conceitos e aplicações, *Sociedade Brasileira de Computação* . https://doi.org/10.5753/sbc.5033.7.4.

Burks, R., Islam, K. A., Lu, Y. and Li, J. (2019). Data augmentation with generative models for improved

malware detection: A comparative study, *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, IEEE, pp. 0660–0665. https://doi.org/10.1109/uemcon47517.2019.8993085.

Cherif, A., Badhib, A., Ammar, H., Alshehri, S., Kalkatawi, M. and Imine, A. (2022). Credit card fraud detection in the era of disruptive technologies: A systematic review, *Journal of King Saud University-Computer and Information Sciences* .

Claro, M., Vogado, L., Santos, J. and Veras, R. (2020). Utilização de técnicas de data augmentation em imagens: Teoria e prática, *Sociedade Brasileira de Computação* . https://doi.org/10.5753/sbc.11.5.3.

Costa, P. J. d. S. (2019). *Sistema de deteção de fraude em pagamentos eletrónicos. Estudo e Implementação usando software de distribuição livre*, PhD thesis.

Dharwa, J. N. and Patel, A. R. (2011). A data mining with hybrid approach based transaction risk score generation model (TRSGM) for fraud detection of online financial transaction, *Intl. J. of Computer Applications* **16**(1): 18–25. https://doi.org/10.5120/1977-2651.

do Amaral, L. R., da Silva Alves, A. H., de Lima Mendes, R., de Souza Gomes, M., Bertarini, P. L. L. and Hruschka, E. R. (2021). Applying never-ending learning (nel) principles to build a gene ontology (go) biocurator, *2021 IEEE Congress on Evolutionary Computation (CEC)*, IEEE, pp. 458–465. https://doi.org/10.1109/cec45853.2021.9504981.

Dong, G., Liao, G., Liu, H. and Kuang, G. (2018). A review of the autoencoder and its variants: A comparative perspective from target recognition in synthetic-aperture radar images, *IEEE Geoscience and Remote Sensing Magazine* **6**(3): 44–68. https://doi.org/10.1109/mgrs.2018.2853555.

Feng, H., Wang, W., Chen, B. and Zhang, X. (2020). Evaluation on frozen shellfish quality by blockchain based multi-sensors monitoring and SVM algorithm during cold storage, *IEEE Access* **8**: 54361–54370. https://doi.org/10.1109/access.2020.2977723.

Fiore, U., De Santis, A., Perla, F., Zanetti, P. and Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection, *Information Sciences* **479**: 448–455. https://doi.org/10.1016/j.ins.2017.12.030.

Forough, J. and Momtazi, S. (2021). Ensemble of deep sequential models for credit card fraud detection, *Applied Soft Computing* **99**: 106883. https://doi.org/10.1016/j.asoc.2020.106883.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative adversarial nets, *Advances in neural information processing systems* **27**. https://doi.org/10.1145/3422622.

Gupta, A., Lohani, M. and Manchanda, M. (2021). Financial fraud detection using naive bayes algorithm in highly imbalance data set, *Journal of Discrete Mathematical Sciences and Cryptography* **24**(5): 1559–1572. https://doi.org/10.1080/09720529.2021.1969733.

Hilal, W., Gadsden, S. A. and Yawney, J. (2022). Financial fraud:: A review of anomaly detection techniques and recent advances.

Hu, Y. and Su, J. (2022). Research on credit risk evaluation of commercial banks based on artificial neural network model, *Procedia Computer Science* **199**: 1168–1176. https://doi.org/10.1016/j.procs.2022.01.148.

Husejinovic, A. (2020). Credit Card Fraud Detection Using Naive Bayesian and C4.5 Decision Tree Classifiers, *Periodicals of Engineering and Natural Sciences, ISSN 2303-4521* **8**(1): 1–5. https://doi.org/10.21533/pen.v8i1.300.g480.
URL: *https://ssrn.com/abstract=3521283*

Itoo, F., Singh, S. et al. (2021). Comparison and analysis of logistic regression, naïve bayes and knn machine learning algorithms for credit card fraud detection, *International Journal of Information Technology* **13**(4): 1503–1511. https://doi.org/10.1016/j.dajour.2022.100071.

Jachemet, B. (2018). *A regulação dos pagamentos eletrônicos: interoperabilidade e desafios jurídicos*, PhD thesis.

Jain, A., Purwar, A. and Yadav, D. (2021). Credit card fraud detection using k-means and fuzzy c-means, *Handbook of Research on Innovations and Applications of AI, IoT, and Cognitive Technologies*, IGI Global, pp. 216–240. https://doi.org/10.4018/978-1-7998-6870-5.ch016.

Junnarkar, A., Adhikari, S., Fagania, J., Chimurkar, P. and Karia, D. (2021). E-mail spam classification via machine learning and natural language processing, *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, IEEE, pp. 693–699. https://doi.org/10.1109/icicv50876.2021.9388530.

Khatri, S., Arora, A. and Agrawal, A. P. (2020). Supervised machine learning algorithms for credit card fraud detection: a comparison, *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, IEEE, pp. 680–683. https://doi.org/10.1109/confluence47617.2020.9057851.

Luo, Y., Zhu, L.-Z., Wan, Z.-Y. and Lu, B.-L. (2020). Data augmentation for enhancing eeg-based emotion recognition with deep generative models, *Journal of Neural Engineering* **17**(5): 056021. https://doi.org/10.1088/1741-2552/abb580.

Maes, S., Tuyls, K., Vanschoenwinkel, B. and Manderick, B. (2002). Credit card fraud detection using Bayesian and neural networks, *Proc. of the 1st Intl. Naiso Congr. on Neuro Fuzzy Technologies*, pp. 261–270.

Makki, S., Assaghir, Z., Taher, Y., Haque, R., Hacid, M.-S. and Zeineddine, H. (2019). An experimental study with imbalanced classification approaches for credit card fraud detection, *IEEE Access* **7**: 93010–93022. https://doi.org/10.1109/access.2019.2927266.

MICHAELIS (n.d.). Dicionário brasileiro da língua portuguesa.

Moghaddam, A. H., Moghaddam, M. H. and Esfandyari, M. (2016). Stock market index prediction using artificial neural network, *Journal of Economics, Finance and Administrative Science* **21**(41): 89–93. https://doi.org/10.1016/j.jefas.2016.07.002.

Monard, M. C. and Baranauskas, J. A. (2003). Conceitos sobre aprendizado de máquina, *Sistemas inteligentes-Fundamentos e aplicações* **1**(1): 32.

Moosavi-Dezfooli, S.-M., Fawzi, A. and Frossard, P. (2016). Deepfool: a simple and accurate method to fool deep neural networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2574–2582. https://doi.org/10.1109/cvpr.2016.282.

Moscato, V., Picariello, A. and Sperlí, G. (2021). A benchmark of machine learning approaches for credit score prediction, *Expert Systems with Applications* **165**: 113986. https://doi.org/10.1016/j.eswa.2020.113986.

Mu, G. and Chen, J. (2022). Developing a conditional variational autoencoder to guide spectral data augmentation for calibration modeling, *IEEE Transactions on Instrumentation and Measurement* **71**: 1–8. https://doi.org/10.1109/tim.2022.3142060.

Murugan Bhagavathi, S., Thavasimuthu, A., Murugesan, A., George Rajendran, C. P. L., Raja, L. and Thavasimuthu, R. (2021). Weather forecasting and prediction using hybrid c5. 0 machine learning algorithm, *International Journal of Communication Systems* **34**(10): e4805. https://doi.org/10.1002/dac.4805.

Nabipour, M., Nayyeri, P., Jabani, H., Shahab, S. and Mosavi, A. (2020). Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis, *IEEE Access* **8**: 150199–150212. https://doi.org/10.1109/access.2020.3015966.

Oinar, C. (n.d.). Generative adversarial networks (gan): Introduction and example. Disponível em https://medium.com/mlearning-ai/generative-adversarial-networks-gan-introduction-and-example-3b66f5f235e9.

O'NEILL, A. (n.d.). Share of economic sectors in the gross domestic product (gdp) of selected global regions in 2020. Disponível em https://www.statista.com/statistics/256580/share-of-economic-sectors-in-the-gross-domestic-product-by-global-regions/.

Peng, J., Zou, B. and Zhu, C. (2022). Combining external attention gan with deep convolutional neural networks for real–fake identification of luxury handbags, *The Visual Computer* pp. 1–12. https://doi.org/10.1007/s00371-021-02378-x.

Pengyu, W. and Wanna, G. (2021). Image detection and basketball training performance simulation based on improved machine learning, *Journal of Intelligent & Fuzzy Systems* **40**(2): 2493–2504. https://doi.org/10.3233/jifs-189243.

Randhawa, K., Loo, C. K., Seera, M., Lim, C. P. and Nandi, A. K. (2018). Credit card fraud detection using AdaBoost and majority voting, *IEEE Access* **6**: 14277–14284. https://doi.org/10.1109/access.2018.2806420.

Roseline, J. F., Naidu, G., Pandi, V. S., alias Rajasree, S. A. and Mageswari, N. (2022). Autonomous credit card fraud detection using machine learning approach, *Computers and Electrical Engineering* **102**: 108132.

Sabau, A. S. (2012). Survey of clustering based financial fraud detection research, *Informatica Economica* **16**(1): 110.

Sailusha, R., Gnaneswar, V., Ramesh, R. and Rao, G. R. (2020). Credit card fraud detection using machine learning, *2020 4th international conference on intelligent computing and control systems (ICICCS)*, IEEE, pp. 1264–1270. https://doi.org/10.1109/iciccs48265.2020.9121114.

Savy, M. (2020). Global Payment Fraud Statistics, Trends & Forecasts.
URL: *https://www.merchantsavvy.co.uk/payment-fraud-statistics/*

Saxena, A., Brault, N. and Rashid, S. (2021). *Big Data and Artificial Intelligence for Healthcare Applications*, Big Data for Industry 4.0, CRC Press. https://doi.org/10.1201/9781003093770.

Shang, Y., Su, X., Xiao, Z. and Chen, Z. (2021). Campus sentiment analysis with gan-based data augmentation, *2021 13th International Conference on Advanced Infocomm Technology (ICAIT)*, IEEE, pp. 209–214. https://doi.org/10.1109/icait52638.2021.9702068.

Sharma, V., Cali, Ü., Sardana, B., Kuzlu, M., Banga, D. and Pipattanasomporn, M. (2021). Data-driven short-term natural gas demand forecasting with machine learning techniques, *Journal of Petroleum Science and Engineering* **206**: 108979. https://doi.org/10.1016/j.petrol.2021.108979.

Shorten, C. and Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning, *Journal of big data* **6**(1): 1–48. https://doi.org/10.1186/s40537-019-0197-0.

Sohony, I., Pratap, R. and Nambiar, U. (2018). Ensemble learning for credit card fraud detection, *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*, pp. 289–294. https://doi.org/10.1145/3152494.3156815.

Sudha, C. and Akila, D. (2021). Majority vote ensemble classifier for accurate detection of credit card frauds, *Materials Today: Proceedings* . https://doi.org/10.1016/j.matpr.2021.01.616.

Sun, G., Zhang, Y., Weiss, R. J., Cao, Y., Zen, H., Rosenberg, A., Ramabhadran, B. and Wu, Y. (2020). Generating diverse and natural text-to-speech samples using a quantized fine-grained vae and autoregressive prosody prior, *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, pp. 6699–6703. https://doi.org/10.1109/icassp40776.2020.9053436.

Tarekegn, A. N., Giacobini, M. and Michalak, K. (2021). A review of methods for imbalanced multi-label classification, *Pattern Recognition* **118**: 107965. https://doi.org/10.1016/j.patcog.2021.107965.

Thomas, L., Crook, J. and Edelman, D. (2017). *Credit scoring and its applications*, SIAM. https://doi.org/10.1137/1.9781611974560.bm.

Tingfei, H., Guangquan, C. and Kuihua, H. (2020). Using variational auto encoding in credit card fraud detection, *IEEE Access* **8**: 149841–149853. https://doi.org/10.1109/access.2020.3015600.

Usama, M., Asim, M., Latif, S., Qadir, J. et al. (2019). Generative adversarial networks for launching and thwarting adversarial attacks on network intrusion detection systems, *2019 15th international wireless communications & mobile computing conference (IWCMC)*, IEEE, pp. 78–83. https://doi.org/10.1109/iwcmc.2019.8766353.

Vujović, Ž. Đ. (2021). Classification model evaluation metrics, *International Journal of Advanced Computer Science and Applications* **12**(6): 599–606. https://doi.org/10.14569/ijacsa.2021.0120670.

Wang, R. and Liu, G. (2021). Ensemble method for credit card fraud detection, *2021 4th International Conference on Intelligent Autonomous Systems (ICoIAS)*, IEEE, pp. 246–252. https://doi.org/10.1109/icoias53694.2021.00051.

Yinka-Banjo, C. and Ugot, O.-A. (2020). A review of generative adversarial networks and its application in cybersecurity, *Artificial Intelligence Review* **53**(3): 1721–1736. https://doi.org/10.1007/s10462-019-09717-4.

Zhang, H., Zhang, L. and Jiang, Y. (2019). Overfitting and underfitting analysis for deep learning based end-to-end communication systems, *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, IEEE, pp. 1–6. https://doi.org/10.1109/wcsp.2019.8927876.

Zhang, X., Han, Y., Xu, W. and Wang, Q. (2021). Hoba: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture, *Information Sciences* **557**: 302–316. https://doi.org/10.1016/j.ins.2019.05.023.

Zhou, T., Ruan, S. and Canu, S. (2019). A review: Deep learning for medical image segmentation using multi-modality fusion, *Array* **3**: 100004. https://doi.org/10.1016/j.array.2019.100004.