ANÁLISE DO DESEMPENHO DE CLASSIFICADORES SUPERVISIONADOS NA DETECÇÃO DE FRAUDE EM TRANSAÇÕES POR CARTÕES DE CRÉDITO

Giovana Gualter Fioravante Teodoro[1] ; Mariana da Costa Lopes[2]; Yasmin Kamilly Christ[3]; Thiago Jose Lucas[4]; Carlos Eduardo Silva Bertazzoli[5]

Resumo
Diversos fatores no âmbito digital contribuem para o crescimento de fraudes em transações realizadas através do cartão de crédito, fato que acarreta prejuízos para pessoas e instituições financeiras. Em resposta a essa problemática, faz-se necessário encontrar medidas que identifiquem e reportem rapidamente esses desvios, tornando as compras digitais mais seguras. Portanto, esse estudo propõe a aplicação de técnicas capazes de, em grande quantidade de dados, extrair padrões e classificá-los como prováveis transações fraudulentas ou não, por meio da aplicação de Machine Learning e do aprimoramento do resultado com *Ensemble Learning*. O resultado obtido é uma melhoria significativa na performance do algoritmo de KNN combinado com a técnica de *stacking*, principalmente considerando a métrica *Recall*.
Palavras-chave: crédito; cartão de crédito; fraude; detecção; ensemble.

Abstract
Several factors in the digital realm contribute to the growth of fraud in credit card transactions, resulting in losses for people and financial institutions. In response to this issue, it is necessary to find measures that quickly identify and report these cheats, making digital purchases safer. Therefore, this study proposes the application of techniques capable of extracting patterns from large amounts of data and classifying them as likely fraudulent transactions or not, through the application of Machine Learning and improvement of the results with Ensemble Learning. The result obtained is a significant improvement in the performance of the KNN algorithm combined with the stacking technique, especially considering the Recall metric.
Keywords: Credit; credit card; fraud; detection; ensemble.

1 Introdução

Avanços tecnológicos, comércio eletrônico, globalização e inovações em produtos financeiros são alguns dos motivos que justificam o aumento das transações financeiras nos últimos anos. Apesar das medidas de segurança da informação, como criptografia e autenticação multifatorial, os consumidores continuam vulneráveis a uma variedade de fraudes

[1] Graduada em Análise e Desenvolvimento de Sistemas pela Faculdade de Tecnologia de Ourinhos-FATEC-Ourinhos. E-mail: giovana.teodoro@fatec.sp.gov.br.
[2] Graduada em Análise e Desenvolvimento de Sistemas pela Faculdade de Tecnologia de Ourinhos-FATEC-Ourinhos. E-mail: mariana.lopes@fatec.sp.gov.br.
[3] Graduada em Análise e Desenvolvimento de Sistemas pela Faculdade de Tecnologia de Ourinhos-FATEC-Ourinhos. E-mail: yasmin.christ@fatec.sp.gov.br.
[4] Doutor e Pós-Doutor em Ciência da Computação pelo Laboratório Avançado de Segurança de Redes da Universidade Estadual Paulista Júlio de Mesquita Filho (LARS/Unesp) de Bauru/SP, professor da Faculdade de Tecnologia de Ourinhos-FATEC Ourinhos e da Universitário na Saint Leo University-Flórida/USA. E-mail: thiagojlucas@gmail.com.
[5] Graduado em Análise e Desenvolvimento de Sistemas pela Faculdade de Tecnologia de Ourinhos-FATEC-Ourinhos. E-mail: eduardo.bertazzoli@fatecourinhos.edu.br.

ao realizar transações online, especialmente com o uso de cartões de crédito. Isso inclui roubos de informações, clonagem de cartões, transações não autorizadas e outros tipos de atividades fraudulentas. No Brasil, dados de uma pesquisa realizada por Sica (2022) revelam que dois em cada dez brasileiros já foram vítimas de fraudes de cartão de crédito, destacando a importância de medidas eficazes para combater esse problema.

Por outro lado, os avanços tecnológicos possibilitam também a utilização de Machine Learning, ou aprendizado de máquina, para promover aos computadores a capacidade de identificar padrões de grandes conjuntos de dados e realizar análise preditiva. Porém, nem sempre esses algoritmos são eficientes, principalmente em casos de conjuntos de dados complexos e variáveis. Nesse contexto, uma alternativa eficaz é o uso de ensembles de modelos, que consiste em combinar múltiplos classificadores para alcançar um desempenho superior em comparação com modelos individuais.

Diante dessa problemática, o presente trabalho teve como objetivo geral a aplicação de técnicas de Ensemble Learning junto a aprendizado de máquina na detecção de fraudes em transações por cartão de crédito através da plataforma Orange Data Mining. Em um viés específico, o escopo consistiu em utilizar o Dataset Kaggle para a preparação de uma base de dados de cartão de crédito, avaliar o desempenho de modelos individuais de Machine Learning, considerando métricas como acurácia, precisão e recall e desenvolver estratégias de ensemble para combinar os modelos de maneira eficaz. A análise dos resultados nos permitiu identificar as principais vantagens e limitações de cada abordagem, e proporcionou recomendações para instituições financeiras que buscam aprimorar suas medidas de segurança e proteção contra atividades fraudulentas no cenário abordado.

Os objetivos específicos foram:

• Realizar uma revisão sistemática da literatura para que fosse possível compreender o cenário atual desta área;

• Realizar o pré-processamento dos dados, para que os mesmos estivessem livres de ruídos discretizados e normalizados;

• Realizar a implementação dos classificadores;

• Realizar os testes dos classificadores analisando os retornos das variáveis da matriz de confusão.

• Analisar os resultados e elaborar as considerações finais deste texto científico.

A estrutura deste trabalho compreende cinco seções. A Seção Introdução apresenta esta introdução contextualizando o tema, expondo a problemática, justificando a relevância e definindo os objetivos. A Seção Revisão da Literatura, revisa a literatura existente sobre

aplicação de ensemble na detecção de fraudes em transações com cartões de crédito, expondo também o método de pesquisa. Na Seção Metodologia, detalhamos a metodologia utilizada para a realização do estudo, incluindo a descrição do conjunto de dados, as técnicas de pré-processamento e os algoritmos de Machine Learning utilizados. Na Seção Resultados, os resultados foram analisados e discutidos. Por fim, na Seção Conclusão, concluímos o estudo destacando suas contribuições e limitações.

2 Fundamentação Teórica

Para propor uma base teórica robusta para esse estudo, foi realizada uma análise de pesquisas anteriores identificando abordagens, metodologias e descobertas relevantes que enriqueceram o conhecimento existente sobre o assunto. Com o intuito de refinar melhor os artigos analisados neste trabalho, foi utilizada a base de artigos científicos *IEEE Xplore*, uma das maiores bases de dados de literatura técnica e científica digital para as áreas de engenharias e ciência da computação. Além da credibilidade em seus artigos publicados, a ferramenta permite recursos avançados de busca que possibilitam o refinamento por palavras-chave e datas.

A Figura 1 ilustra o processo sistemático para a obtenção de trabalhos correlatos:

Figura 1 - Representação da pesquisa de trabalhos relacionados

Fonte: Elaborada pelos autores.

Conforme pode ser observado na Figura 1, foi realizada a pesquisa por trabalhos contendo as palavras: "*credit*", "*card*", "*fraud*", "*detection*" e "*ensemble*". A pesquisa retornou um valor significativo de 102 trabalhos. Portanto, com o intuito de realizar uma filtragem mais específica, definimos o *range* de publicação em 2021 até 2024 e optamos por trabalhos de livre acesso. Dessa maneira, o número de resultados diminuiu para 9. Desses trabalhos, seguimos com o levantamento inicial das informações do *dataset*, técnicas e resultados.

2.1 Trabalhos correlatos

A presente subseção apresenta os principais detalhes obtidos através da análise dos trabalhos correlatos.

Mienye and Sun (2023) relatam que com a crescente importância do cartão de crédito para a economia digital, as fraudes com essa ferramenta crescem junto. Portanto, os autores utilizam as redes neurais Long Short-Term Memory (LSTM) e Gated Recurrent Unit (GRU) com uma Multilayer Perceptron (MLP) com balanceamento no conjunto de dados. Seus resultados apontam que a combinação do conjunto de aprendizado proposto com o método SMOTE-ENN alcançou uma sensibilidade e especificidade de 1,000 e 0,997, respectivamente, superando outros classificadores e métodos de Machine Learning amplamente utilizados na literatura.

No artigo de Ghaleb *et al*. (2023) é demonstrado que o aumento de fraudes em cartões de crédito tem causado grandes perdas financeiras. Nesse sentido, é proposto um modelo de detecção baseado em aprendizado de conjunto e uma rede generativa adversária (GAN) assistida por técnicas de oversampling de minoria sintetizada em conjunto (ESMOTE-GAN). Múltiplos subconjuntos foram extraídos e utilizados para treinar diversos conjuntos de modelos GAN, visando gerar subconjuntos sintetizados. Em seguida, um conjunto de classificadores Random Forest foi treinado com base nessa abordagem. Os resultados mostram que o modelo proposto obteve melhorias de 1,9\% e 3,2\% no desempenho geral e na taxa de detecção, respectivamente, com uma taxa de alarme falso de 0\%.

Kalid *et al*. (2024) expõe dois grandes desafios para a aplicação de Machine Learning na detecção de fraudes em pagamentos com cartões de crédito: distribuição desbalanceada dos dados e overlapping (quando há sobreposição entre as distribuições de características de diferentes classes). Com isso, foi dirigido um estudo utilizando a metodologia PRISMA para comparar 87 artigos provenientes de diferentes bases de dados com variação de ano de publicação. Como resultado, recomenda-se a aplicação de deep learning, ensemble learning e sampling methods para identificação dos desvios.

O estudo realizado por Esenogho *et al*. (2022) propõe uma abordagem inovadora para a detecção de fraudes de cartão de crédito, visando mitigar as perdas enfrentadas por empresas financeiras. Eles utilizam um classificador de conjunto de redes neurais e ensemble combinado com um método híbrido de reamostragem de dados, já que os algoritmos tradicionais têm dificuldade em lidar com o comportamento dinâmico de compras e clientes dos cartões. A base dessa abordagem é uma rede neural de memória de curto e longo prazo (LSTM), arquitetura

especializada em lidar com sequências de dados, como aquelas encontradas em transações de cartão de crédito. Essa rede é integrada a uma técnica de AdaBoost, que é um método de aprendizado de conjunto que combina múltiplos classificadores fracos para formar um classificador forte. Além disso, é utilizada uma reamostragem híbrida que combina a técnica de sobre amostragem minoritária sintética (SMOTE) com a técnica de undersampling (ENN). Os resultados do estudo demonstraram que essa abordagem superou as técnicas tradicionais, alcançando uma sensibilidade de 0,996 e uma especificidade de 0,998.

Embora o Machine Learning e o Deep Learning sejam usados para detectar fraudes com cartões de crédito, o desequilíbrio nos conjuntos de dados, onde os dados de fraude são muito menores do que os dados de transações normais, limita a eficácia dos algoritmos de classificação binária. Em vista disso, Ding *et al*. (2023) melhoraram a parte geradora da Rede Adversarial Generativa do Autoencoder Variacional (VAEGAN) e introduziram um novo método de sobreamostragem, que diversifica os dados gerados no conjunto de treinamento. Avaliado em um conjunto de dados de transações de cartão de crédito, o método de sobreamostragem utilizando o VAEGAN aprimorado superou técnicas convencionais, como GAN, VAE e SMOTE, em precisão, pontuação F1 e outras métricas.

Por conta desses crescentes números de casos de fraudes nas transações bancárias feitas de maneira digital, é possível ver um avanço nas implementações do que chamam de Sistemas Automáticos de Detecção de Fraudes (Fraud Detection Systems), softwares que são capazes de identificar com rapidez e com alta precisão possíveis fraudes nas transações. Mas percebe-se que por conta da aleatoriedade dos comportamentos, ocorre uma certa dificuldade no aprendizado e na adaptação. No artigo de Lebichot *et al*. (2021), é explicado que será implementada a técnica Transfer Learning, onde é treinado um modelo e ele é utilizado para treinar outros em atividades relacionadas, isso para obter uma análise de possíveis fraudes. Foi utilizado um conjunto de dados, com mais de dois milhões de transações de comércio eletrônico para o estudo. Também foram abordadas técnicas, e realizadas análises em diferentes tipos de ocorrências nas transações. Elas mostram que a precisão dos métodos Transfer Learning depende da quantidade de informações que foram aplicadas e mostram que métodos auto supervisionados e semi supervisionados foram aplicados para solução desse problema.

Baseando-se nos cenários atuais, Almazroi and Ayub (2023) adotaram de uma abordagem única de inteligência artificial que foi criada especificamente para processamentos de dados de transações financeiras em tempo real, a abordagem sistemática. Iniciam o processo com uma entrada e pré-processamento de dados de inteligência artificial, e com o SMOTE, uma técnica

utilizada principalmente em problemas de classificação desbalanceada, eles minimizam o desequilíbrio de dados.

O ponto principal da classificação de inteligência artificial vem no modelo RXT que foi ajustado com hiper parâmetros utilizando o algoritmo Jaya (RXT-J). O modelo de IA criado por eles passa por uma avaliação minuciosa que supera significativamente outros algoritmos existentes, por uma margem entre 10\% a 18\% em várias métricas de avaliação, mantendo sua eficiência computacional.

O seguinte estudo de Jemai *et al*. (2024) visa explorar os métodos da aprendizagem de máquina Ensemble Learning, que combinam múltiplos modelos visando melhorar a precisão nas suas previsões para a detecção de fraudes, utilizado dois conjuntos de dados diferentes: O Sparkov, conjunto de dados gerados artificialmente e uma base de dados reais de clientes da União Europeia. Foram utilizados modelos como XGBoost e Random Forest em ambas as bases para uma avaliação da sua performance. De acordo com os testes, teve maior êxito a base de dados real do que a base de dados fictícia. Como resultado, este estudo nos mostra que os algoritmos de detecção estão trazendo uma melhor eficácia em ambientes previsíveis, onde não há aleatoriedade de dados e que nesse caso é mais perigoso o vazamento de dados sensíveis.

Na pesquisa de Hashemi *et al*. (2023), é explorado o uso de hiper parâmetros de ajuste de peso de classe que são utilizados para lidar com dados desbalanceados e ajudam a melhorar o desempenho do modelo, tendo um controle das transações que são fraudadas ou as reais. Foi utilizada a Otimização Bayesiana para ajustar esses parâmetros e aplicando técnicas como CatBoost e XGBoost para o aprimoramento do método LightGBM, considerando o mecanismo de votação. Também foi utilizado no estudo o Deep Learning para maiores ajustes nesses parâmetros, com a intenção de melhorar seu desempenho. Foram feitas avaliações em dados reais, e uma validação cruzada separada nas técnicas utilizadas, e viram que os métodos LightGBM e XGBoost tiveram os melhores resultados nessas avaliações realizadas. Com a aplicação do Deep Learning e a Otimização Bayesiana, tiveram melhoras comparados com os métodos de referência.

O quadro 1 compila a taxonomia dos trabalhos correlatos que foram analisados, com foco em representar a Base de Dados, Técnicas e Resultados:

Quadro 1 - Taxonomia de trabalhos correlatos

Artigo	Dataset	Técnicas	Resultados
Mienye an Sun (2023)	Europeu	LSTM, GRU e MLP	Precisão: 98,40% | Sensibilidade: 1% | Especificidade: 0,997%
Ghaleb *et al*. (2023)	Europeu	ESMOTE-GAN e RF	AUC: 92,9%
Kalid *et al*. (2024)	Europeu e Chinês	Deep Learning,	True Positive Rate:

		Ensemble e Sampling métodos	99,6%
Esenogho *et al*. (2022)	Europeu	LSTM e AdaBoost	Sensibilidade: 0,996 | Especificidade: 0,998 | Acuracidade: 98,4% | Precisão: 97,34%
Ding *et al*. (2023)	Kaggle	VAEGAN e GRU	Precisão: 0,9298 | F1 Score: 0,8765
Lebichot *et al*. (2021)	Parceiro Industrial	Métodos de Transferência	-
Almazroi and Ayub (2023)	Europeu	GRU, RXT e SMOTE	F1 Score: 0,987 | Acuracidade: 0,979 | Precisão: 0,977
Jemai *et al*. (2024)	Sparkov	NB, RF, XGBoost	Acurácia: 0,86
Hashemi *et al*. (2023)	Mundo real (não especificado)	CatBoost, LightGBM e XGBoost	CatBoost: 0,99880 | LGBM: 0,99919 | XGB: 0,99923

Fonte: Elaborada pelos autores.

3 Metodologia

O objetivo deste capítulo é descrever os métodos escolhidos, incluindo os materiais e técnicas para o desenvolvimento da pesquisa.

Figura 2 - Representação do processo de metodologia

Fonte: Elaborada pelos autores.

Conforme o fluxograma descrito na Figura 2, a primeira etapa foi obter os dados utilizados como objeto de análise e a partir deles, iniciamos a classificação através dos algoritmos de Machine Learning. Em seguida, o modelo de validação cruzada foi comparado e validados os dados obtidos, e por fim, foi possível calcular os métodos de avaliação e definir os valores e resultados obtidos.

3.1 Conjunto de dados

O conjunto de dados escolhido está presente no site *Kaggle*, uma plataforma amplamente utilizada pela comunidade de cientistas de dados, pois o acesso é gratuito e com grande variedade. A escolha da base foi motivada por 3 parâmetros: data de atualização mais recente, classificação da usabilidade e a atribuição "*gold*", definida pela plataforma. O arquivo, denominado como "*creditcard_2023.csv*" contém 550.000 linhas de dados identificando se a transação é fraudulenta ou não, tornando-o ideal para a detecção através de *Machine Learning*.

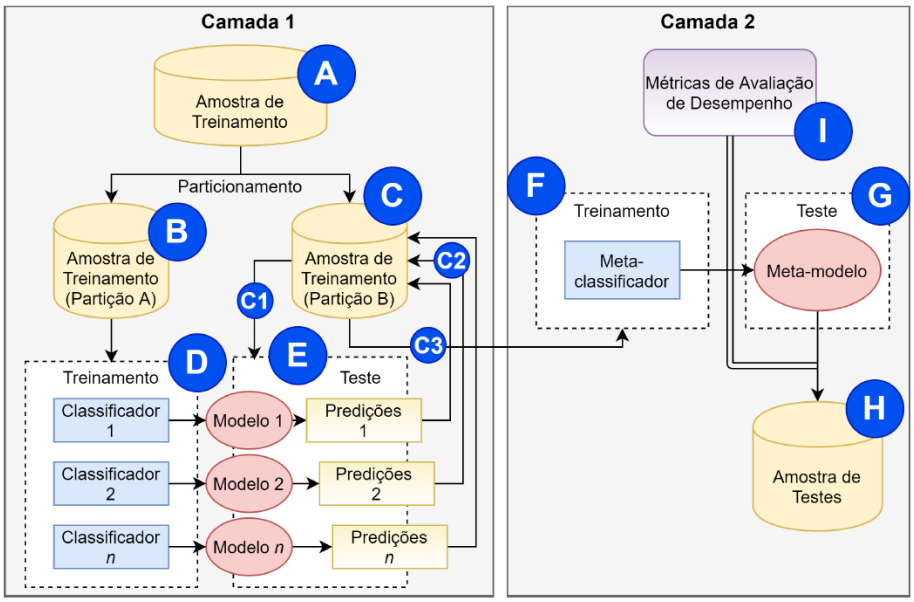
3.2 Ensemble Learning

A etapa de *Ensemble Learning* teve como intuito melhorar o desempenho do estudo, combinando múltiplos modelos para fazer previsões mais precisas, trabalhando com a ideia de grupos, dado que o *ensemble* compensa os pontos fracos individuais de cada classificador. Para isso, foram treinados independentes, agrupando suas previsões para obtermos um resultado mais confiável e com uma redução do *overfitting*, que acontece quando um algoritmo se ajusta muito bem aos dados de treinamento, mas não generaliza bem para novos dados.

Este trabalho teve foco no *Stacking*, uma técnica de aprendizado de *ensemble* que junta previsões de diversos modelos bases treinados que se combinam em uma segunda camada para no fim gerar uma previsão otimizada. Para isso, na primeira camada, foi necessário dividir a amostra de treino em duas partes $D_{Train}A$ e $D_{Train}B$, e treinar os classificadores escolhidos utilizando um desses conjuntos $D_{Train}A$. Depois com os classificadores treinados, foram feitas as predições utilizando o segundo conjunto $D_{Train}B$, o qual constrói um novo conjunto de dados.

Indo para a segunda camada, um meta-classificador foi escolhido e treinado com as predições e o conjunto anterior $D_{Train}B$, que ajusta e combina esses dados gerando uma previsão final com um melhor desempenho de classificação. A Figura 3 ilustra o funcionamento descrito acima com relação ao *Stacking*.

Figura 3 - Representação do fluxo de funcionamento do Stacking



Fonte: Lucas *et al*. (2023)

3.3 Algoritmos Classificadores

Os algoritmos de classificação baseados em *Machine Learning* foram treinados com uma amostra de dados fornecida pelo seu usuário, para que pudesse ser analisada e categorizada,

retornando previsões de resultados para certo tipo de situação. Para este trabalho, foram utilizados os seguintes algoritmos classificadores:

a) *K-Nearest Neighbors (KNN)*

Trata-se de um algoritmo de aprendizado supervisionado que pode ser usado para classificação e regressão. Ao classificar, o algoritmo olha para o ponto mais próximo ("k" vizinho) no conjunto em que trabalha e assim define a classe mais comum entre eles, em relação à classificação. Entretanto, para a regressão, o algoritmo utiliza a média dos valores dos vizinhos ("k").

b) *Support Vector Machine (SVM)*

Sendo um algoritmo de classificação, o SVM tenta encontrar um hiperplano que maximize a distância mais próxima entre cada ponto entre as classes, sendo assim, conseguindo separar as classes dos dados da melhor forma possível utilizando da técnica de Kernel para que os dados sejam transformados em um espaço dimensional mais ajustado, para que assim sejam separados linearmente.

c) *Rede Neural* (*Neural Network*)

Inspirada nas funcionalidades dos neurônios de um cérebro, a rede neural artificial conta com camadas de nós (neurônios), e os dados trabalhados passam por diversas camadas. A primeira camada conta com a entrada desses dados que em seguida são processados, transformados e ativados por uma função nas camadas ocultas, para no fim gerar previsões na camada de saída.

3.4 Modelo de Validação Cruzada

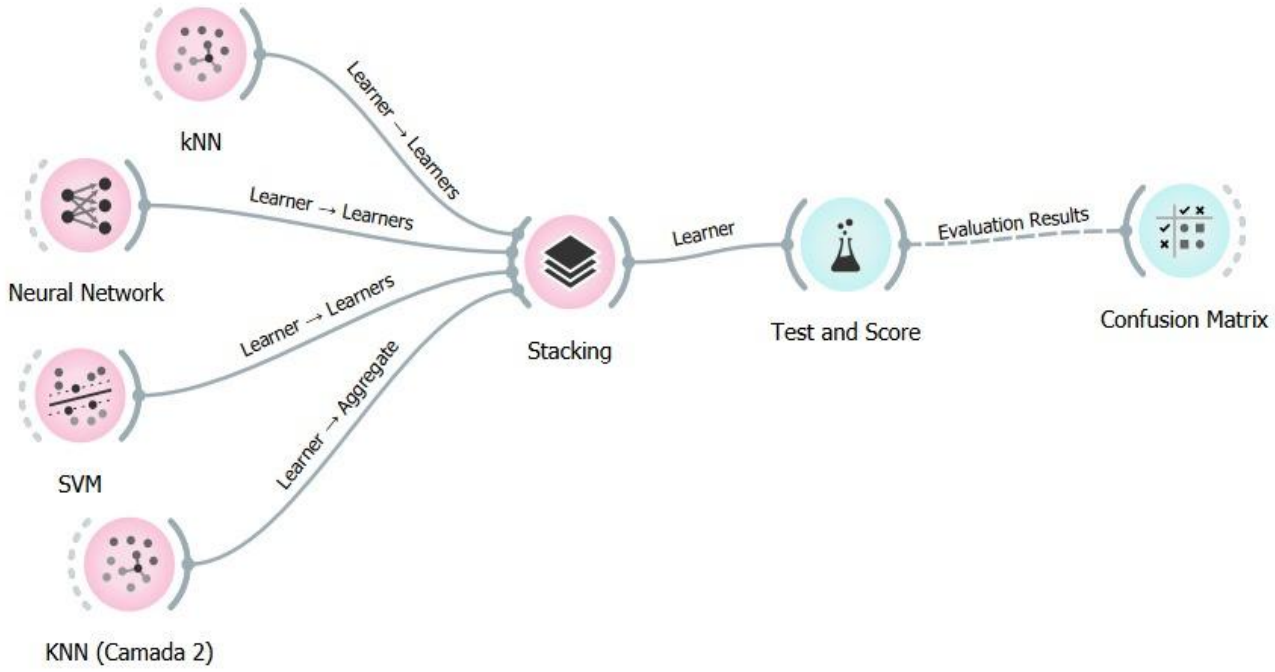
Para implementar a validação cruzada neste trabalho, a ferramenta utilizada é o *Orange*. Essa é uma técnica que avalia a performance de um modelo de treinamento analisando o conjunto de dados utilizados. Os dados de entrada são divididos em subconjuntos, chamados “dobras” e são realizados testes com estes modelos utilizando métricas de avaliação, e é fornecido um conjunto de estatísticas de precisão para cada uma dessas dobras. Depois, as estatísticas de todas as partes são comparadas para avaliar a qualidade do conjunto de dados e determinar o quanto o modelo está suscetível a variações.

3.5 Orange Data Mining

O Orange Data Mining é uma plataforma de código aberto criada para análise de dados e Machine Learning, que oferece uma variedade de ferramentas e serviços para criação, treinamento, avaliação e implantação de modelos. A escolha da utilização da plataforma neste

trabalho é pela popularização de ferramentas low-code, otimização do tempo e segurança das informações armazenadas em nuvem. A Figura 4 exemplifica como é a construção de fluxos na ferramenta.

Figura 4 - Representação do fluxo no Orange



Fonte: Lucas *et al*. (2023)

3.6 Modelo de Avaliação

As métricas escolhidas para avaliar os resultados das aplicações supracitadas foram acurácia, precisão, \textit{recall} e Matriz de Confusão. Abaixo segue um descritivo de cada métrica.

- **Acurácia:** Indica uma proporção da performance geral do modelo, indicando quantas classificações foram realizadas corretamente.

$$Acurácia = \frac{Verdadeiros\ Positivos + Verdadeiros\ Negativos}{Total\ de\ Exemplos} \quad (1)$$

- **Precisão:** Mede a proporção de exemplos positivos classificados corretamente, ou seja, a capacidade do modelo de não definir uma observação negativa como positiva.

$$Precisão = \frac{Verdadeiros\ Positivos}{Verdadeiros\ Positivos + Falsos\ Positivos} \quad (2)$$

- **Recall:** Demonstra a capacidade do modelo em encontrar todos os exemplos positivos.

$$Recall = \frac{Verdadeiros\ Positivos}{Verdadeiros\ Positivos + Falsos\ Negativos} \quad (3)$$

- **Matriz de Confusão:** Avalia o desempenho do modelo, criando um quadro que compara predições feitas pelo modelo com os dados reais do conjunto. Obtém quatro componentes principais que detalham as predições corretas e incorretas do sistema:
- **Verdadeiro Positivo (TP):** Quando o modelo prevê uma classe positiva, e o valor real era positivo.
- **Falso Negativo (FN):** Quando o modelo prevê uma classe negativa, mas o valor real era positivo.
- **Falso Positivo (FP):** Quando o modelo prevê uma classe positiva, mas o valor real era negativo.
- **Verdadeiro Negativo (TN):** Quando o modelo prevê uma classe negativa, e o valor real era negativo.

4 Resultados

Foram analisadas três perspectivas dos resultados: os números apresentados na matriz de confusão, as principais métricas de desempenho e o tempo de processamento. Essas análises foram representadas por um quadro e um gráfico, permitindo uma interpretação mais clara e detalhada dos dados e sua relação com a performance dos modelos.

Em todas as análises, é possível comparar os resultados dos modelos aplicados individualmente com a aplicação do método de *Stacking*, usando a combinação dos 3 algoritmos e variando a segunda camada que potencializa o *Stacking*.

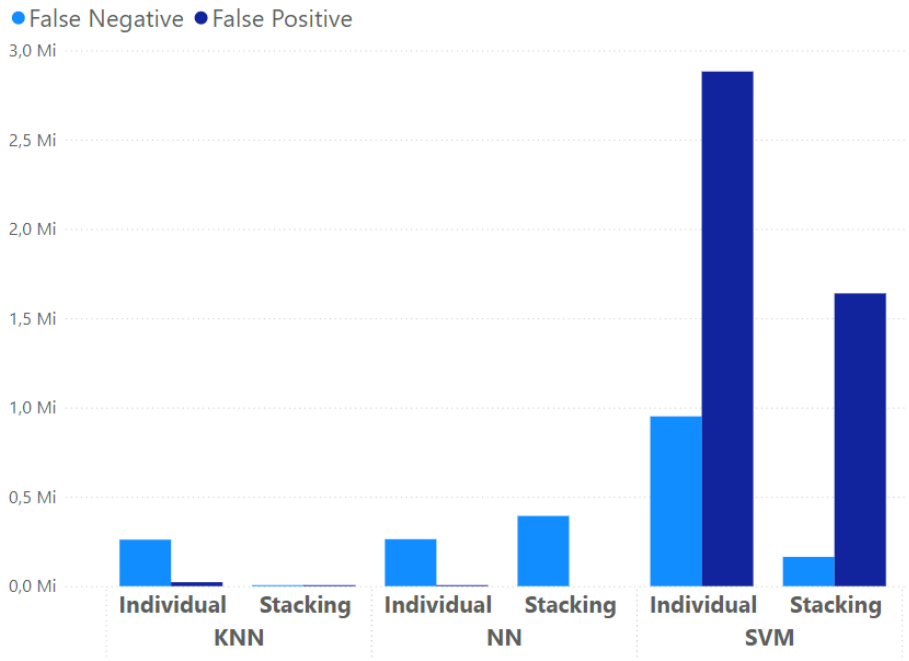
A primeira compreensão dos resultados é a proporção entre Verdadeiros Positivos, Verdadeiros Negativos, Falsos Positivos e Falsos Negativos.

Quadro 2 - Números Matriz de Confusão

Modelo Base	Camada 2	FN	FP	TN	TP
KNN		260471	22237	3810673	106619
SVM		950747	2882163	950747	324966
Neural Network		263043	4703	3828207	104047
Stacking KNN + SVM + NN	KNN	4421	3231	4103454	388894
Stacking KNN + SVM + NN	SVM	163724	1639611	2467074	229591
Stacking KNN + SVM + NN	NN	393315	0	4106685	0

Fonte: Elaborada pelos autores.

Figura 5 - Gráfico comparando os resultados da Matriz de Confusão



Fonte: Elaborada pelos autores.

Com esse resultado, foi possível analisar como o *Stacking* conseguiu aumentar o número de verdadeiros negativos ao passo que diminuiu o número de falsos positivos para o modelo de SVM, o que é um ponto positivo para a técnica. Por outro lado, há um princípio de *overfitting* nos resultados de NN.

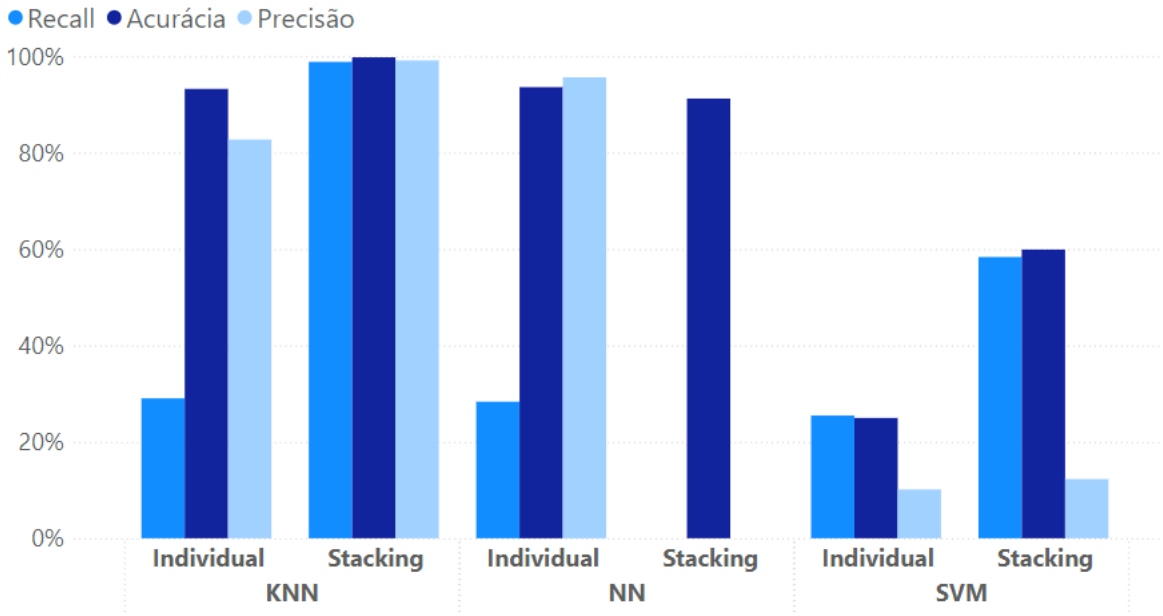
Em seguida, são apresentadas as métricas de Acurácia, Precisão e *Recall* que trazem os resultados mais interessantes para esse trabalho.

Quadro 3 - Indicadores de desempenho

Modelo Base	**Camada 2**	**Acurácia**	**Precisão**	**Recall**
KNN		93,27%	82,74%	29,04%
SVM		24,97%	10,13%	25,47%
Neural Network		93,63%	95,68%	28,34%
Stacking KNN + SVM + NN	KNN	99,83%	99,18%	98,88%
Stacking KNN + SVM + NN	SVM	59,93%	12,28%	58,37%
Stacking KNN + SVM + NN	NN	91,26%	0,00%	0,00%

Fonte: Elaborada pelos autores.

Figura 1 - Métricas de Desempenho



Fonte: Elaborada pelos autores.

Em um panorama geral, fica evidente que para KNN e SVM os resultados foram melhores utilizando *Ensemble Learning*. Em especial, cabe ressaltar a métrica de *Recall* para o KNN, que sobe de 29,04% para 98,88%, sendo essa uma métrica de suma importância para analisar o cenário de fraude de cartões de crédito.

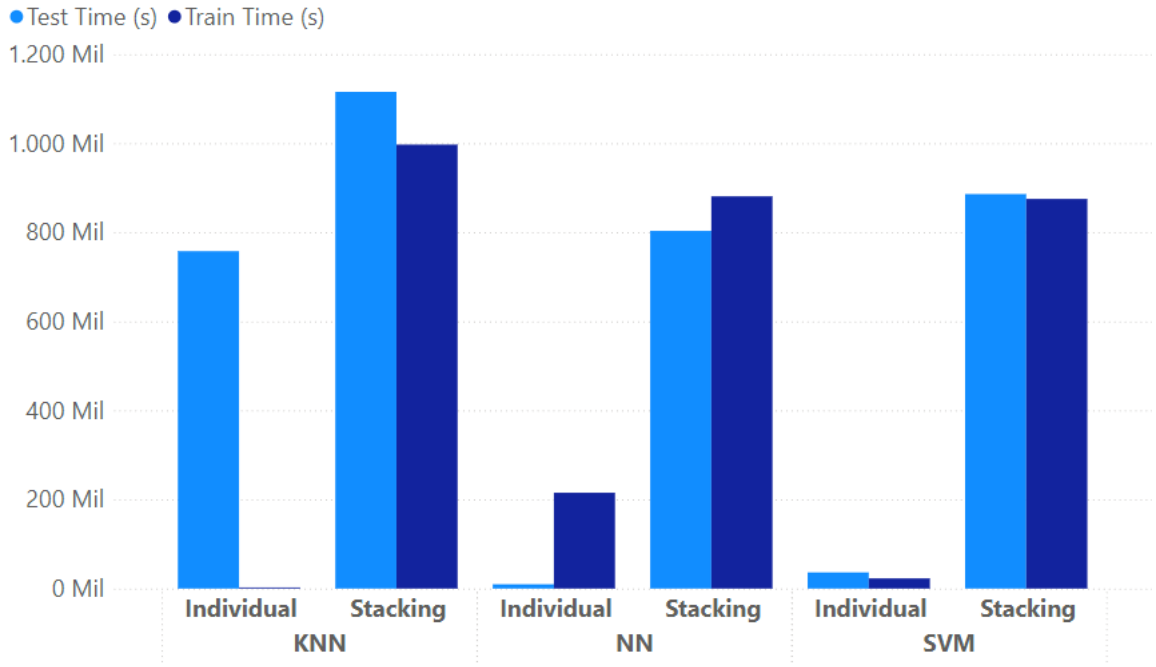
Por fim, verifica-se o tempo de processamento dos algoritmos.

Quadro 4 - Tempo de Processamento

Modelo Base	**Camada 2**	**Train Time (s)**	**Test Time (s)**
KNN		843	757.757
SVM		22.884	36.296
Neural Network		215.217	9.636
Stacking KNN + SVM + NN	KNN	997.171	1.115.534
Stacking KNN + SVM + NN	SVM	875.085	886.000
Stacking KNN + SVM + NN	NN	880.974	803.272

Fonte: Elaborada pelos autores.

Figura 2 - Tempo de Processamento



Fonte: Elaborada pelos autores

Ao utilizar a técnica de *Stacking*, já é esperado que o tempo de processamento seja maior quando comparado ao modelo individualmente. O tempo mais alto do *Stacking* é uma consequência da sua estrutura mais complexa que utiliza algoritmos de forma simultânea para otimizar as previsões.

5 Considerações finais

A proposta neste estudo de utilizar técnicas de *Ensemble Learning* na detecção de fraudes de cartão de crédito apresentou resultados interessantes, pois cada algoritmo se comportou de uma maneira diferente. Ao combinar os algoritmos com o método *Stacking*, as medidas de desempenho melhoraram significativamente, com exceção do modelo de *Neural Network*. Esse avanço foi notável na métrica *Recall* que tem uma importância ímpar para o cenário de análise de fraudes, tendo em vista que a classe positiva representa fraude e a classe negativa representa transação no *Dataset* utilizado, o *Recall* representa o quão bom é o algoritmo para detectar transações fraudulentas. Um exemplo claro é o algoritmo de KNN, cujo *recall* aumentou de 29,04\% para 98,88\% ao ser combinado com os outros algoritmos, evidenciando a eficácia da abordagem utilizada.

Além disso, a ferramenta *Orange* demonstrou grande potencial para implementar essas técnicas de forma mais intuitiva, facilitando o aprendizado e a prática de *Machine Learning*, o que a torna uma excelente escolha para estudos e experimentações nessa área.

A proposta para a continuação dessa pesquisa é a utilização de *Ensemble Pruning*, uma ferramenta que além de diminuir o custo computacional sem comprometer a qualidade das previsões também melhora o desempenho, o tempo de treinamento e a redução de *overfitting*, um supera juste, melhorando a acurácia, ajudando no equilíbrio entre os desempenhos e assim mantendo o modelo treinado, gerenciável.

Agradecimentos

Os autores agradecem ao Centro Estadual de Educação Tecnológica “Paula Souza” (CEETEPS) pelo suporte nas pesquisas desenvolvidas no Laboratório de Cibersegurança Defensiva e Inteligência Artificial (Detect.AI - https://detectai.fatecourinhos.edu.br/).

Referências

ALMAZROI, A. A.; AYUB, N. Online payment fraud detection model using machine learning techniques. **IEEE Access**, v. 11, p. 137188–137203, 2023. Disponível em:https://doi.org/10.1109/ACCESS.2023.3339226. Acesso em: 08 jul. 2025.

DING, Y. *et a*l. Credit card fraud detection based on improved variational autoencoder generative adversarial network. **IEEE Access**, v. 11, p. 83680–83691, 2023. Disponível em:https://doi.org/10.1109/ACCESS.2023.3302339. Acesso em: 08 jul. 2025.

ESENOGHO, E. *et al*. A neural network ensemble with feature engineering for improved credit card fraud detection. **IEEE Access**, v. 10, p. 16400–16407, 2022. Disponível em:https://ieeexplore.ieee.org/document/9698195. Acesso em: 08 jul. 2025.

GHALEB, F. A. *et al.* Ensemble synthesized minority oversampling-based generative adversarial networks and random forest algorithm for credit card fraud detection. **IEEE Access**, v. 11, p. 89694–89710, 2023. Disponível em:https://ieeexplore.ieee.org/document/10224552. Acesso em: 08 jul. 2025.

HASHEMI, S. K.; MIRTAHERI, S. L.; GRECO, S. Fraud detection in banking data by machine learning techniques. **IEEE Access**, v. 11, p. 3034–3043, 2023. Disponível em:https://ieeexplore.ieee.org/document/9999220. Acesso em: 08 jul. 2025.

JEMAI, J.; ZARRAD, A.; DAUD, A. Identifying fraudulent credit card transactions using ensemble learning. **IEEE Access**, v. 12, p. 54893–54900, 2024. Disponível em:https://ieeexplore.ieee.org/document/10477993. Acesso em: 08 jul. 2025.

KALID, S. N. *et al*. Detecting frauds and payment defaults on credit card data inherited with imbalanced class distribution and overlapping class problems: A systematic review. IEEE Access, v. 12, p. 23636–23652, 2024. Disponível em:https://ieeexplore.ieee.org/document/10423008. Acesso em: 08 jul. 2025.

LEBICHOT, B. *et al*. Transfer learning strategies for credit card fraud detection. **IEEE Access**, v. 9, p. 114754–114766, 2021. Disponível em:https://ieeexplore.ieee.org/document/9512084. Acesso em: 08 jul. 2025.

LUCAS, T. J. *et al*. A comprehensive survey on ensemble learning-based intrusion detection approaches in computer networks. IEEE Access, 2023. Disponível em:https://ieeexplore.ieee.org/document/10299619. Acesso em: 08 jul. 2025.

MIENYE, I. D.; SUN, Y. A deep learning ensemble with data resampling for credit card fraud detection. **IEEE Access**, v. 11, p. 30628–30638, 2023. Disponível em: https://ieeexplore.ieee.org/document/10081315. Acesso em: 08 jul. 2025.

SICA, N. **Impressões digitais e sua relação com as pessoas e as empresas**. Kaspersky Daily, 2022. Disponível em: https://www.kaspersky.com.br/blog/pesquisa-impressoes-digitais/18906/. Acesso em: 26 abr. 2025.