International Journal of Advanced Research and Multidisciplinary Trends (IJARMT)
An International Open Access, Peer-Reviewed Refereed Journal
Impact Factor: 6.4 Website: https://ijarmt.com ISSN No.: 3048-9458

Advanced Machine Learning and Deep Learning Approaches for Credit Card Fraud Detection: A Comprehensive Review and Performance Analysis

Prachi Singh
Research Scholar, Computer Science & Engineering, Sam Global University Bhopal, Madhya Pradesh
prachisingh335@gmail.com

Abstract

The rapid growth of online transactions and digital payment systems has significantly increased the risk of credit card fraud, posing serious financial and security challenges for banking institutions and customers. Traditional machine learning techniques often struggle to detect complex, evolving, and highly imbalanced fraud patterns, resulting in reduced detection accuracy and delayed responses. Recent advancements in deep learning, federated learning, attention mechanisms, and graph-based models have shown promising improvements in fraud detection performance while addressing privacy, scalability, and adaptability issues. This study presents a comprehensive review of state-of-the-art credit card fraud detection techniques, including federated deep learning models, attention-enhanced neural networks, prototype-based learning, ensemble methods, and feature selection strategies. The analysis highlights key contributions, datasets, and performance outcomes of existing approaches, emphasizing their strengths and limitations. Furthermore, the review identifies persistent challenges such as computational complexity, concept drift, class imbalance, and feature redundancy. The findings suggest that hybrid and privacy-preserving deep learning models offer superior accuracy, robustness, and practical applicability for real-world fraud detection systems, paving the way for future research toward efficient and adaptive fraud prevention solutions.

Keywords: Credit Card Fraud Detection; Machine Learning; Deep Learning; Federated Learning; Graph Attention Networks; Class Imbalance; Feature Selection; Financial Security

I INTRODUCTION

The detection of fraudulent transactions is a critical concern in the cyber security industry due to the imbalance between legitimate and fraudulent transactions. Various methods have been evaluated to improve fraud detection accuracy, including random under-sampling and SMOTE, where under-sampling increases recall while SMOTE ensures better overall accuracy and F1-score. Ensemble models combining techniques such as Bagging, Boosting, Random Forest, SVM, and KNN have been shown to outperform traditional methods in terms of accuracy, precision, recall, and F1-score. Compact Data Learning (CDL) has been introduced to streamline machine learning training while maintaining high accuracy in fraud detection. Meta-Heuristic Optimization (MHO) techniques have been applied to reduce feature size significantly while achieving high detection accuracy. Hybrid machine learning models have also been explored, with certain combinations, such as Adaboost and LGBM, performing best.

The challenges of fraud detection in online transactions have been highlighted, emphasizing the importance of feature engineering and dataset balance. Novel data augmentation models, such as K-CGAN, have been proposed to enhance precision, recall, and F1-score beyond traditional SMOTE techniques. The integration of Generative Adversarial Networks (GANs) with RNN architectures has demonstrated improved sensitivity and specificity in fraud detection. Additionally, models designed to address sparse data and evolving fraud patterns have been developed to mitigate financial losses. Collectively, these studies emphasize the need for innovative techniques to address data imbalance, improve detection accuracy, and adapt to evolving fraud strategies. Existing approaches for detecting credit card fraud were covered in this chapter. These methods are categorized into Deep learning Models, Machine learning Models, Hidden Markov Models, and Artificial Immune System Models. [1]

Literature on Credit Card Fraud

It is the most common type of fraud that occurs in banking and various financial sector service industries. The fraudster performs illegitimate transactions using someone else’s card. The owner of the cardholder will not get any doubt or alert in such a way their card will be used to perform fraudulent transactions. The fraudsters will take the advantage of cardholder’s inexperience and unawareness and perform many transactions. By the time, the owner realizes the fact, by that time his/her card might be used many times. These two ways are discussed below.

a) Offline Credit Card Fraud

The fraudsters, somehow gets the physical card and perform transactions at various places such offline stores or online stores, this type of usage is called Offline Fraud. In this scenario, the owner of the cardholders understands the loss of his card and report to their bank, by that time the fraudster manages to perform illegitimate and illegal transactions, subsequently substantial financial loss will happen. Whatever policies and procedures imposed by various banking and financial institutions enforce during the card issue time, the cardholders have to pay the amount even though he did not perform the transaction.

b) Online Credit Card Fraud

In Online credit card fraud, the card’s sensitive information will be stolen and not the physical card. It is can be referred to as virtual card theft. The stolen sensitive information will be used to perform transactions where the presence of the physical card is not essential, such as online transactions and international transactions. This type of cheating is very serious and highly challenging task. When attacker performs one or more illegitimate transactions, then only this type of fraud can be identified. Fraudsters use different ways to steal the sensitive information of the credit card. Few of the approaches are furnished here.

Skimming

Fraudsters use advanced tools, generally referred as skimmers, to obtain sensitive information from the card's magnetic stripe electronically. This process involved copying of card’s magnetic stripe electronically. After that, they will create forged physical cards from the stolen information, then they will use the forged cards to perform transactions. It can be performed by an untrustworthy employee who may use and swipe the customers card in the skimming device. These days, the same type of skimmers has also been introduced in ATMs to get the debit cards and credit card data when customers came for withdrawing money. Micro-cameras are installed in ATMs to capture the pass code or PIN of the cardholder during ATM transactions. [2]

Cloning websites

Intruder clones genuine and most famous websites to mislead customers. As the deceitful website appears like a genuine one, some customers will use their credit card data to perform the transaction. Once intruders get the customer's sensitive information, they will perform fraudulent transactions whenever they wish to do so.

Credit card Fraud detection based on deep learning

In this part, we will go over the ways that deep learning can help detect credit card fraud. An artificial learning method known as deep learning has the ability to learn and make judgments independently. Here are some ways that deep learning can help to spot fraudulent charges on a credit card:

[3] used a Generative Adversarial Network simulation to identify fraudulent charges on credit cards. This method was successful in identifying the con artists because it used an inference technique that was based on both the discriminator and generative models. The inference method was able to differentiate between the samples that meet the criteria. The inference model learned the probabilistic relationship using a deep denoising auto encoder. There was a very low percentage of misclassification and the approach had good accuracy. We were able to differentiate between the positive and negative data by constructing a minimax game using adversarial training. The method's efficiency dropped as the number of parameters increased, and it couldn't solve the complicated classification issues.

[4] came up with a data mining technique for detecting credit cards. This technique established a risk scoring methodology that identified the Fraud Suspicion score. If the threshold value was greater than the risk scoring system's predetermined value, the order would be automatically authorized; otherwise, human approval would be required. The risk score was recognized and the manual revision process was improved by this strategy. But giving solutions in e-tail for fraud detection was a difficulty with this strategy.

[5] developed an approach to fraud detection using semantic fusion. By combining the k-means algorithm with the artificial bee colony algorithm (ABC), this method created a semantic fusion method. Combining global and neighborhood searches in the ABC solved the problem of real cluster handling incompetence. The ABC optimizer, k-mean classifier, and optimized classifier were combined to form the unified frame after the rule engine was employed to determine the crucial features. Considerations such as the customer's location, account balance, and usage trends helped determine whether they were being dishonest. Though computationally demanding, this approach enhanced classification accuracy and convergence speed.

[6] created a hybrid approach to credit card fraud detection by combining GANN with Genetic Algorithms. They trained the neural network to identify features such network type, weight, node count, and layer count using the BPN. Because of the high success rate for the gifted individual, this strategy took use of that feature. This approach successfully identified instances of credit card fraud. But this approach couldn't anticipate the credit card purchases.

[7] Created a CNN to identify credit card fraud. In this approach, fraud behavior was determined by learning inherent patterns from tagged data. The available transaction data was then used to build the feature matrix. Then, nous utilized CNN to get the latent pattern set for the samples. A big worry is the method's computational complexity, although it successfully predicted credit card fraud.

[8] devised a HOBA system to identify credit card fraud. I utilized HOBA to construct the feature variables that are used to detect credit card fraud. In order to analyze the transaction's behavior, deep learning approaches were employed. Taking the false-positive rates into account allowed us to measure performance. This strategy worked well since it cut down on fraud losses and regulatory expenses. On the other hand, the method in question was unable to alleviate the rising computational cost caused by the expansion of the variable set.

[9] used a hybrid ensemble approach with deep learning to identify credit card fraud. Both approaches were compared using the champion-challenger framework, where the hybrid ensemble method is the champion and the deep learning method is the challenger. Creating the challenger and champion technique required data analysis in addition to the existing standard. In order to assess both models, we looked at the limitations of the fraud detection system. The framework for selecting the victorious model made use of both post-launch and offline testing.

[10] Examined a technique for detecting credit card fraud by means of Generative Adversarial Networks. To make the classifier more accurate, The goal was to train Generative Adversarial Networks to mimic the actions of members of the underrepresented group. The training data was supplemented with examples from the minority class to produce a better training set. In addition to addressing the issue of class imbalance, this method enhanced the performance of detecting credit card fraud. Nonetheless, the claimed maximum sensitivity was not achieved using this approach.

[11] developed a Whale Algorithm - optimized BP neural network for the detection of fraud in credit cards. This method optimized the back propagation network weights through the whale swarm optimization algorithm. They used the whale swarm optimization process to find the starting point, and the backpropagation algorithm to fix the erroneous value. This approach was able to detect credit card fraud with good accuracy and a faster convergence speed. Nevertheless, the system's stability was weak, and there were issues with local optimization with this strategy.

[12] developed a new fraud detection method based on deep networks. The deep network helped in the recognition of the data by the complicated distributions. The best features were extracted using the deep autoencoder and the instances were classified. The class labels were determined using the SoftMax network. The data was mapped into high dimensional space by employing the autoencoder. The discriminative space was used in the sparse models for classification. This method had a low variance in the detection of credit card fraud but it failed to implement the recurrent convolutional neural network (RCNN) and sequence of the transaction for improving the accuracy.

[13] Modeled a fraud detection method in the credit card using the hybrid method by integrating the unsupervised and supervised method. The unsupervised method detected the new types of fraud whereas the supervised method used the behaviour of the past fraudulent for the detection. The different granularity levels were determined by considering the outlier scores. The effectiveness of detecting credit card fraud was enhanced by this strategy. Using so many outlier ratings degraded the accuracy; hence this strategy was unsuccessful in improving it.

Credit card Fraud detection based on machine learning

This section explains the various fraud detections methods that used machine learning approaches. The machine learning models, take decisions based on the learned data. The credit card fraud detection methods of this kind are given below,

[14] developed an algorithm for detecting fraud using machine learning. There were a lot of fraudulent transactions found using this strategy. Modifying the cost matrix yielded the accessible transaction amount. In terms of detecting fraud, this strategy performed admirably. Unfortunately, this approach didn't take tiny frauds into account and didn't incorporate prior bankcard transactions, thus it couldn't accurately detect fraudsters.

[15] built a model for detecting fraud using Bayes with a minimum risk. Using quantifiable tradeoffs, Bayes minimal risk made recommendations for credit card fraud detection. For a budget-conscious system, even if this approach reduced costs, even more significant savings are required.

[16] built a model of an OlightGBM (Optimized Light Gradient Boosting Machine) to identify credit card fraud. For this method, we combined LightGBM with a Bayesian hyper-parameter optimization algorithm to fine-tune the parameters. Using this strategy, both valid and fraudulent credit card transactions were treated differently. Function selection followed data pre-processing. The features were chosen with the intention of lowering dimensionality via the practice of knowledge gathering. Prior to assigning weights, the information gain approach was used to determine the similarities. In order to determine the top features, we looked at the heaviest features first. Furthermore, OLightGBM was employed to process and manage a substantial quantity of data. Thanks to its feature-scanning technique, OLightGBM was able to group similar features together and generate histograms based on these bundles. But while this strategy did improve the prediction method's performance, it did little to simplify processing.

[17] examined the potential for fraud by means of a Federated Learning for Fraud Detection (FFD) concept. By combining the common Fraud Detection System (FDS), the updates were calculated locally. Instead of transmitting sensitive bank information to a data center, this approach used federated fraud detection. Data that was not centrally stored had an effect on how sensitive and unavailable the dataset was. Learning parameters from the global shared technique and individual bank updates provided the data. Overcoming the skew distribution problem, this approach discovered credit card fraud. But it wasn't able to efficiently aggregate the model changes using this way.

[18] Modeled a Deep Learning–based Auto-Encoder–based fraud detection approach utilizing Restricted Boltzmann Machine. The data was cleaned up by organizing and converting the data properties into principal component analysis (PCA) using XLSTAT. The technique established the input and output equivalents utilizing the auto-encoder by means of the backpropagation procedure. This approach successfully identified the highest number of system frauds. The hyperbolic tangent function was used to encrypt and decode the input, which was then translated to the output. Using the parameter gradients and an auto-encoder, the backpropagation was executed. Although the computing cost was great, this approach offered accurate detection.

developed a method for identifying fraudulent transactions that is grounded in game theory. To measure the rule's impact on the pool's performance, normalized scores were awarded to each selection rule. The power-index was generated by combining the Shapley Value with Coalitional Game Theory, which described the performance. The top-ranked rules were selected using the compact rule-set, and the score was utilized to predict the rule's maintenance in the periodic rule assessment procedure. By choosing the rules for keeping the operation running in the NRT pool, we were able to determine the particular rule's past performance. Although it had a good recall rate and moderate precision, this technique couldn't figure out how to summarise performance by assigning rules with a normalized score.

used Dataset Shift Quantification to construct a detection algorithm. Daily transaction classification allowed us to quantify the covariate shift in the dataset. Covariate shift and different days were present for efficient categorization. To assess the distance matrix between the days, the agglomerative clustering method was employed, which involved synchronizing the calendar with the shift pattern for predetermined durations. Going forward, the detection approach was enhanced by factoring in the knowledge of the dataset shift as the significant characteristic. Although performance was improved, this method only slightly improved recall and precision. Credit card theft can be reduced if Thennakoon, A.et al.[147] created a machine learning-based system. This approach detected fraudulent transactions and notified the end-user through the GUI using the API module. The following stage in the investigation of fraud had to be taken when the suspicious transaction was detected. While this technique was successful in detecting data distribution skewness, it was unable to pinpoint exactly where fraud had taken place.

[19] developed a fusion model to detect credit card transaction fraud. Data pre-processing initially involved the under-sampling method. This approach predicts credit card fraud by combining and optimizing the XGBoost and Lasso-Logistic algorithms. While Lasso-Logistic was quicker, the XGBoost technique had the limitation of keeping extraneous variables. They were able to overcome these issues using the XGBoost and Lasso-Logistic methods by combining them. Improved prediction classification accuracy was a result of this method's strong extrapolation capabilities. But the computational complexity increases due to the merging of the methods.

II ELECTRONIC PAYMENT SYSTEM

Internet access and initial registration with the relevant payment service provider are prerequisites for using the electronic payment system by both the customer and the retailer. Gateway for accepting payments: If it's a private interbank clearing network or the public network, the vendor will provide with a payment gateway choice. The gateway is a connecting point between the conventional and electronic payment systems; in other words, it's a bridge. They also have bank accounts with a bank that is part of a clearing network, which is the other side of things. The payment instrument used by the customer was issued by the customer's bank (the "issuer bank") (the "issuer bank"). People who buy things from a bank get their hands on transactional data from that bank. The buyer opts to pay with his debit or credit card when he purchases the products and services. The vendor verifies the buyer's identity and payment details with the payment gateway before shipping the product. The payment gateway asks the issuer bank to make sure everything is okay. A payment gate will take money from a customer's account, deposit it in the merchant's account and tell them about the transaction if everything goes well.

The main reason to build an electronic payment system is that it allows businesses and customers to mix and match different commercial services in an electronic market place. Lynch and Lundquist suggest that the e-payment system will help both businesses and consumers, according to Tae-Hwan, 1998.Virtual markets are good for businesses because they make it easier and cheaper to communicate with each other.

[20] Identifying fraudulent transactions is a top priority for the cyber security industry due to the vast majority of transactions being legitimate. Random under-sampling and over-sampling are two approaches they test for gauging the precision of identifying fraudulent transactions in unbalanced real-world datasets. In this investigation, they employ the SMOTE—synthetic minority over-sampling technique. Although this strategy sacrifices accuracy for recall, random under-sampling aims for fairness by excluding examples from the majority class. To ensure fairness and reach statistical significance, the researchers used SMOTE to generate synthetic instances of the minority class. Compared to SMOTE, which has a little lower recall but greater accuracy (86.75%) and a more consistent F1 score (73.47%), random under-sampling has a high recall (92.86%) but isn't that accurate. They looked at a plethora of machine learning algorithms until we found one that satisfied our requirements for accuracy, F1 score, and recall; this is necessary because genuine fraudulent transactions necessitate two stages of verification. Cyber security experts may now better assess their company's unique requirements thanks to our comparison, which clarifies the nuances and consequences of each method. The need of addressing class imbalances in cyber security fraud detection, maintaining ongoing monitoring and exploring novel ways to boost applicability are all emphasized in this paper[41].

[21] The rise of credit card fraud in this age of technological developments highlights the critical need for more effective fraud detection systems. This study aims to examine the efficacy of ensemble methods, a type of machine learning model, in improving the accuracy of credit card fraud detection. A thorough literature analysis revealed multiple issues with existing fraud detection algorithms, such as data imbalance, concept drift, false positives and negatives, restricted generalizability, and challenges in real-time processing. This is a distinctive ensemble model that integrates Bagging, Boosting, Random Forest, Support Vector Machine, and K-Nearest Neighbors to address various issues. The credit card dataset is recognized for its imbalance; however, this ensemble model employs a combination of under-sampling and the Synthetic Minority Over-sampling Technique (SMOTE), with several other machine learning techniques, to address this problem. To evaluate the model in a real-world setting, a dataset comprising information on credit card transactions conducted by Europeans is employed. This model's methodology encompasses data pre-processing, feature engineering, model selection, and evaluation, while leveraging the computational resources provided by Google Colab. This facilitates effortless training and testing of the model. The proposed ensemble method markedly decreases the complexity of credit card fraud detection when contrasted with traditional machine learning techniques and independent classifiers. All metrics, including accuracy, precision, recall, and F1-score, indicate that the ensemble surpasses current models. This research emphasizes the efficacy of ensemble methods in combating fraudulent transactions. The provided findings provide a foundation for future research into building more robust and flexible fraud detection systems, which is essential because credit card theft methods are always changing.

[22] The dramatic growth in credit card usage over the past decade, spurred on by developments in global trade, e-commerce and FinTech, has made credit card fraud an urgent topic of study. The critical requirement for efficient fraud detection systems is defined by the worldwide loss estimates that are expected to surpass USD 400 billion in the next ten years. A new method for detecting credit card fraud is introduced in our paper, which makes use of ML. To overcome the problem of data imbalance, which frequently reduces the efficacy of learning, they trained our models using the European cardholders dataset. The novel approach relies on compact data learning (CDL), a potent method for streamlining ML training without compromising accuracy. This allows us to use smaller and simpler datasets. Our CDL-adapted feature reduction beats out a number of other ML algorithms and techniques for feature reduction in experimental comparisons. The results of this study have both theoretical and practical significance for the financial industry, which stands to gain significantly from an improved fraud detection system and for the field as a whole [43].

[23] a This study's overarching goal is to rectify the data mismatch in CCFD, a major obstacle to trustworthy fraud prediction algorithms. Due to the low frequency of fraudulent transactions and the ever-changing tactics used by fraudsters, FD is an extremely difficult problem to resolve. Efficiently detecting fraud is crucial for minimizing financial losses and maintaining secure transactions. The study improves the efficiency and reliability of FD mechanisms by changing the paradigm for moving from imbalanced to balanced data. Datasets from Kaggle's CCF benchmarks were studied to strategically apply Meta-heuristic optimization (MHO) approaches. Customers in Europe who used credit cards were the subjects of this dataset. They measured the impact on prediction accuracy, fitness values, feature selection, and processing time to assess their capacity to identify the most critical minimal set of features. The research compares fifteen different MHO methods, with the use of nine different transfer functions (TFs) that determine which features are most important for fraud prediction. To determine how the selected features affect prediction accuracy, two ML classifiers—SVM and random forest (RF) are employed. While reducing the feature size by as much as 90%, they were able to attain a classification accuracy of 97%. The result suggested a major improvement in model efficiency. Further, it demonstrated how important feature selection is for improving FDSs and adjusting to data imbalance problems. In addition, this study demonstrates how machine learning is still developing, which is radically improving FDSs through new methods.

[24] The negative effects of financial crimes on banks have increased at an alarming rate. Various single-and hybrid-learning methods have been employed to identify fraudulent activities like credit card transactions. Nevertheless, these methods are severely lacking since they did not examine alternative hybrid algorithms for a specific dataset. Using a real-world dataset, this study suggests and examines seven hybrid ML models for fraud detection. previous, cutting-edge ML algorithms were employed to identify instances of credit card fraud. The second stage involved building hybrid approaches using the most effective algorithm from the previous stage. According to our results, the best model is the hybrid Adaboost + LGBM, since it achieved the best overall performance. Research in this area should move further with an emphasis on credit card hybridization and algorithm kinds

[25] The arrival of the COVID-19 pandemic further accelerated the already steady expansion of online purchasing. Credit cards are, without a doubt, the most popular method of payment for internet purchases. As the damages incurred from fraud continue to mount, the issue of credit card fraud detection has taken on more significance. The majority of studies that have examined this issue have done so in silos, with a singular emphasis on improving the data mining models. Businesses are understandably wary of embracing new models, particularly black-box ones and there seems to be a large chasm between the conclusions drawn from academic studies and these concerns. In this paper, we looked at the issue from a broader academic and business viewpoint, identifying features engineering and imbalanced datasets as obstacles to fraud detection and identifying the most profitable areas to invest in when improving fraud detection systems. Our results are grounded in actual data pertaining to CNP (card not present) fraud transactions, the most common kind of fraud. An worldwide card-processing company, who is also our industrial partner, supplied the data. In order to determine the most cost-effective way to enhance their fraud detection system, we examined various data mining models and methods to the stated difficulties to their current production systems.

[26] Credit cards are among the most popular ways to pay for things online in many countries, both developed and developing. The invention of the credit card has simplified, made easier and improved online transactions. On the other hand, it has increased the rate of fraud by providing criminals with additional possibilities to conduct fraud. Many companies and regular people have lost millions of dollars due to credit card fraud, which is a worrying global impact. Many organizations and enterprises depend significantly on machine learning techniques to detect and automatically categorize fraudulent transactions due to the high volume of transactions. An imbalance in the data set is not a minor concern since machine learning technique success is heavily dependent on training data quality. In most cases, the data only shows a small fraction of transactions that were fraudulent. This has a significant impact on how well machine learning classifiers work. Addressing the imbalanced data problem and dealing with the rare incidences of fraud, this study offers a novel data augmentation model, K-CGAN, for credit card fraud detection and analyzes several data augmentation strategies. In order to measure how well the augmentation methods worked, we compared them against some of the most popular categorization algorithms. After comparing several augmentation methods, these results demonstrate that B-SMOTE, K-CGAN, and SMOTE possess the highest Precision and Recall. In terms of F1 Score and Accuracy, K-CGAN is at the top.

[27] The detection of credit card fraud is an important problem for the banking sector because of the large monetary stakes involved. When faced with imbalanced datasets or changing fraud patterns, traditional ML methods frequently fall short. This research suggests a deep learning framework that combines RNNs with GANs to improve the capacity to detect fraud. In order to improve the training set and fix data imbalance, the GAN component creates synthetic fraudulent transactions that look realistic. To determine if a transaction is fraudulent or valid, the discriminator is first trained to distinguish between actual and fake ones. It makes use of a number of DL architectures, including GRUs, Simple RNNs and LSTM networks. Both the sensitivity and specificity of the GAN-GRU model were 0.992 and 1.000, respectively, on the European credit card dataset. demonstrating substantial improvements over conventional approaches according to the experimental results. This study demonstrates how GANs, when integrated with deep learning architectures, can improve credit card fraud detection while also making it more flexible.

[28] financial organizations face enormous financial losses and a decline in consumer trust as a result of credit card theft, making the detection of this crime a key concern. In addition to having an immediate effect on profits, fraud erodes consumer trust in financial institutions and can have far-reaching consequences for their image. When faced with sparse data, novel fraud tactics, or complicated patterns, traditional machine learning approaches often fail to provide adequate detection accuracy. We offer FEDGAT-DCNN, a model that incorporates dilated convolutions and a GAT into a federated learning framework, to tackle these issues. With FEDGAT-DCNN's federated learning, banks can work together to train models on local datasets, improving accuracy and robustness without compromising data privacy. With a GAT, institutions may update their models in real-time, allowing them to respond rapidly to changing fraud trends. By extending the size of the model's receptive field without raising computing overhead, enlarged convolutions improve the detection of complex and subtle fraudulent activities. When tested against both traditional models and other federated learning methods, FEDGAT-DCNN outperforms them with ROC-AUC values of 0.9712 and 0.9992, respectively, on the 2018CN and 2023EU datasets. These outcomes demonstrate the reliability, precision and practicality of FEDGAT-DCNN in detecting fraud in real-life situations.

[29] The banking, insurance, government and law enforcement industries are just a few of the many modern-day uses for fraud detection systems. With the current spike in fraud attempts, fraud detection has become more important in protecting sensitive information or personal details. Other forms of fraud abound, such as card-not-present (CNP) fraud, check forgeries, deceptive accounting practices and stolen credit cards. This article provides an overview of the

CCFDP approach for detecting and preventing credit card not present fraud using big data analytics. The two-step process that makes up the proposed CCFDP for dealing with suspicious activity is the fraud prevention process (FPP) and the fraud detection process (FDP). After the FDP has detected potentially harmful conduct, the FPP steps in to put a stop to it. The FDP stage makes use of five cutting-edge methods: t-distributed stochastic neighbor embedding (RU), PCA, logistic regression learning (LRL), and singular value decomposition (SVD). The FDP can't conduct experiments without first balancing the dataset. Random Undersampling is employed to circumvent this problem. In addition, FDP needs to reduce the dimensionality characteristics for better data presentation. This method uses the t-SNE, PCA and SVD algorithms, which makes training data faster and more accurate. The FPP determines the likelihood of CNP fraud success or failure using the LRL model. The proposed CCFDP mechanism is implemented using Python. Based on the outcomes of the tests, we confirm that the proposed CCFDP mechanism works. [30]

Credit card transactions have skyrocketed alongside the expansion of online shopping and other digital payment methods. Analyzing consumer data for the purpose of detecting and preventing fraud has been greatly facilitated by machine learning (ML). Nevertheless, ML classifiers perform worse in real-world credit card data due to the existence of irrelevant and redundant characteristics. This study proposes a hybrid feature-selection methodology that integrates both filter and wrapper methods to refine the feature set utilized by machine learning to only the most pertinent variables. A genetic algorithm (GA) wrapper incorporates characteristics evaluated using the information gain (IG) technique and uses an ELM as its learning algorithm. In order to address the issue of uneven classification, the fitness function of the proposed GA wrapper is adjusted by replacing the conventional accuracy measure with the geometric mean (G-mean). With a sensitivity of 0.997 and a specificity of 0.994, the suggested strategy outperformed prior baseline procedures and methodology in the most current literature. [31]

[32-33] Some category features in credit card transactions may have high-cardinality, or broad domains, with hundreds of potential values. Due to findings with weaker generalization and increased resource utilization, the inclusion of such features makes analysis tougher. Thus, it is usual practice to disregard such qualities and delete them, even though doing so wastes the information they supplied. On the flip side, we detail in this article how high-cardinality features improve credit card fraud detection. Therefore, they introduce a novel domain reduction approach that maintains the ability to identify fraud. Implementing a deep feedforward neural network using actual statistics sourced from a major Brazilian bank, researchers were able to demonstrate that adding these features does improve the quality of fraud detection, as assessed by the F-1 metric. The suggested algorithm's ability to decrease attribute cardinality improved model training times without sacrificing predictive skills; this was the key contribution.

III CONCLUSION

Despite significant advancements in credit card fraud detection, several challenges persist that hinder the development of highly efficient and real-time fraud detection systems. One major issue is the computational complexity of existing models, as many state-of-the-art machine learning and deep learning approaches require substantial processing power, making them impractical for real-world applications with stringent latency requirements. Additionally, fraudulent transaction patterns evolve dynamically, and most existing methods struggle to adapt to concept drift, leading to decreased accuracy over time. While techniques like Synthetic Minority Over-sampling Technique (SMOTE) are commonly used to address dataset imbalance, they can introduce noise and overfitting, reducing the overall reliability of fraud detection systems. Another critical gap lies in feature selection and dimensionality reduction, where many studies still rely on manual engineering rather than automated feature extraction methods that could leverage deep learning for improved efficiency. Moreover, federated learning has emerged as a promising approach to enhance privacy in fraud detection by enabling decentralized training, but challenges related to security protocols, communication overhead, and scalability remain unresolved. Addressing these gaps requires a holistic approach that integrates adaptive learning techniques, efficient data preprocessing, and privacy-preserving mechanisms to develop robust, real-time, and scalable fraud detection systems for electronic payment platforms.

Future Scope

The increasing sophistication and dynamic nature of financial fraud necessitate the development of highly efficient, adaptive, and real-time fraud detection systems. One of the major challenges in this domain is the computational complexity of existing models, as many state-of-the-art machine learning and deep learning approaches require substantial processing power and memory, making them impractical for deployment in real-world environments with strict latency and resource constraints. Future research should therefore focus on designing lightweight and optimized models that balance detection accuracy with computational efficiency.

Another critical limitation of current fraud detection systems is their inability to effectively handle concept drift, as fraudulent transaction patterns evolve continuously over time. Most existing models are trained on static datasets and fail to adapt to changing behaviors, resulting in performance degradation. Future studies should explore adaptive and online learning mechanisms that can dynamically update detection models in response to evolving fraud patterns without requiring complete retraining.

Although techniques such as the Synthetic Minority Over-sampling Technique (SMOTE) are widely used to address class imbalance in fraud datasets, they may introduce noise and increase the risk of overfitting, thereby reducing model generalization. Future research can investigate more robust imbalance-handling strategies, including cost-sensitive learning, ensemble-based sampling methods, and generative approaches that preserve the underlying data distribution.

Feature selection and dimensionality reduction remain another important research avenue. Many existing studies rely heavily on manual feature engineering, which is time-consuming and may overlook complex feature interactions. Future work should emphasize automated feature extraction techniques using deep learning architectures and attention mechanisms to improve efficiency and detection performance while reducing reliance on domain-specific manual intervention.

Furthermore, federated learning has emerged as a promising paradigm for privacy-preserving fraud detection by enabling decentralized model training across multiple institutions without sharing sensitive data. However, challenges related to communication overhead, scalability, security vulnerabilities, and model convergence remain largely unresolved. Future research should focus on enhancing federated learning frameworks through secure aggregation protocols, communication-efficient model updates, and robust defense mechanisms against adversarial attacks.

To explore various approaches and methodologies developed to detect fraudulent credit card activities, focusing on their applicability, effectiveness, and adaptability in dynamic transaction environments. In the modern banking industry, assessing credit risk is a critical task for ensuring financial stability and minimizing loan defaults. Existing approaches for creditworthiness evaluation often struggle to provide high accuracy and robust predictions due to the complex and non-linear relationships among diverse features in client data, such as income, age, employment and previous credit history. Moreover, many traditional models fail to strike a balance between precision and computational efficiency, resulting in suboptimal performance in predicting credit risk.

Objectives of further work

Develop a robust data pipeline to handle the Credit Dataset in multiple formats (.csv, .xlsx, .txt), enabling pre-processing steps such as handling missing values, performing label encoding and dropping unnecessary columns.

Design and implement an Artificial Neural Network (ANN)-based approach to effectively detect fraudulent credit card transactions with high accuracy.

Utilize data preprocessing techniques, such as Standard Scaling, Principal Component Analysis (PCA), and Synthetic Minority Over-sampling Technique (SMOTE), to improve model performance on highly imbalanced datasets.

Implement regularization techniques such as Batch Normalization and Dropout to prevent overfitting and ensure the model generalizes well to unseen fraud patterns.

Evaluate the performance of ANN, VGG16, and VGG19 models in terms of accuracy, precision, recall, and F1-score to determine the most efficient approach.

References

1. Fiore U, De Santis A, Perla F, Zanetti P, Palmieri F, “Using generative adversarial networks for improving classification effectiveness in credit card fraud detection”, Information Sciences, vol.479, pp.448-55, April 2019.
2. Wang, C., Wang, Y., Ye, Z., Yan, L., Cai, W., and Pan, S., "Credit Card Fraud Detection Based on Whale Algorithm Optimized BP Neural Network," In proceedings of 13th International Conference on Computer Science & Education (ICCSE), 2018.
3. Kazemi, Z., and Zarrabi, H., "Using deep networks for fraud detection in the credit card transactions," In proceedings of 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), IEEE, 2017

4. Carcillo F, Le Borgne Y A, Caelen O, Kessaci Y, Oblé F, Bontempi G, “Combining unsupervised and supervised learning in credit card fraud detection”, Information Sciences, May 2019.
5. Zhou, H., Chai, H., and Qiu, M, “Fraud detection within bankcard enrollment on mobile device based payment using machine learning,” Frontiers of Information Technology & Electronic Engineering, vol.19, no.12, pp.1537–1545, 2018.
6. Bahnsen, A.C., Stojanovic, A., Aouada, D. and Ottersten, B.,” Cost sensitive credit card fraud detection using Bayes minimum risk”, In proceedings of 12th international conference on machine learning and applications, vol. 1, pp. 333-338, December 2013.
7. Taha A and Malebary S J, “An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine”, IEEE Access, vol.8, pp.25579-87, February 2020.
8. Yang W, Zhang Y, Ye K, Li L, Xu C Z, “FFD: A Federated Learning Based Method for Credit Card Fraud Detection”, In International Conference on Big Data, pp.18-32, June 2019.
9. Pumsirirat A and Yan L, “Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine”, International Journal of advanced computer science and applications, vol.9, no.1, pp.18-25, January 2018.
10. Gianini G, Fossi L G, Mio C, Caelen O, Brunie L, Damiani E, “Managing a pool of rules for credit card fraud detection by a Game Theory based approach”, Future Generation Computer Systems, vol.102, pp.549-61, January 2020.
11. Lucas, Y., Portier, P.-E., Laporte, L., Calabretto, S., He-Guelton, L., Oble, F., & Granitzer, M., "Dataset Shift Quantification for Credit Card Fraud Detection," In proceedings of second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), IEEE, 2019.
12. Li, D., "Credit card fraud identification based on unbalanced data set based on fusion model," In proceedings of 1st International Conference on Civil Aviation Safety and Information Technology (ICCASIT), IEEE, 2019.
13. Abdul Rehman Khalid, Nsikak Owoh, Omair Uthmani, Moses Ashawa, Jude Osamor, John Adejoh (2023) “Enhancing Credit Card Fraud Detection: An Ensemble Machine Learning Approach” 2024, 8(1), 6; https://doi.org/10.3390/bdcc8010006, 3 January 2024
14. Xiaomei Feng, Song-Kyoo Kim (2024) “Novel Machine Learning Based Credit Card Fraud Detection Systems” 2024, 12(12), 1869; https://doi.org/10.3390/math12121869, 15 June 2024
15. Diana T. Mosa, Shaymaa E. Sorour, Amr A. Abohany, Fahima A. Maghraby (2024) “CCFD: Efficient Credit Card Fraud Detection Using Meta-Heuristic Techniques and Machine Learning Algorithms” 2024, 12(14), 2250; https://doi.org/10.3390/math12142250, 19 July 2024

16. Esraa Faisal Malik, Khai Wah Khaw, Bahari Belaton, Bahari Belaton, Wai Peng Wong, XinYing Chew (2022) “Credit Card Fraud Detection Using a New Hybrid Machine Learning Architecture” 2022, 10(9), 1480; https://doi.org/10.3390/math10091480, 28 April 2022
17. Igor Mekterović, Mladen Karan, Damir Pintar, Ljiljana Brkić (2021) “Credit Card Fraud Detection in Card-Not-Present Transactions: Where to Invest” 021, 11(15), 6766; https://doi.org/10.3390/app11156766, 23 July 2021
18. Emilija Strelcenia, Simant Prakoonwit (2023) “Improving Classification Performance in Credit Card Fraud Detection by Using New Data Augmentation” 2023, 4(1), 172-198; https://doi.org/10.3390/ai4010008, 31 January 2023
19. Ibomoiye Domor Mienye, Theo G. Swart (2024) “A Hybrid Deep Learning Approach with Generative Adversarial Network for Credit Card Fraud Detection” 2024, 12(10), 186; https://doi.org/10.3390/technologies12100186, 2 October 2024
20. Mengqiu Li, John Walsh (2024) “FEDGAT-DCNN: Advanced Credit Card Fraud Detection Using Federated Learning, Graph Attention Networks and Dilated Convolutions” 2024, 13(16), 3169; https://doi.org/10.3390/electronics13163169, 11 August 2024
21. Abdullah Alharbi, Majid Alshammari, Ofonime Dominic Okon, Amerah Alabrah, Hafiz Tayyab Rauf, Hashem Alyami, Talha Meraj (2022) “A Novel text2IMG Mechanism of Credit Card Fraud Detection: A Deep Learning Approach” 2022, 11(5), 756; https://doi.org/10.3390/electronics11050756, 1 March 2022
22. Boyu Liu, Longrui Wu, Shengdong Mu (2024) “Research on Small-Sample Credit Card Fraud Identification Based on Temporal Attention-Boundary-Enhanced Prototype Network” 2024, 12(24), 3894; https://doi.org/10.3390/math12243894, 10 December 2024
23. Shanshan Jiang, Ruiting Dong, Jie Wang, Min Xia (2023) “Credit Card Fraud Detection Based on Unsupervised Attentional Anomaly Detection Network” 2023, 11(6), 305; https://doi.org/10.3390/systems11060305, 13 June 2023
24. Abdul Razaque, Mohamed Ben Haj Frej, Gulnara Bektemyssova, Fathi Ams (2022) “Credit Card-Not-Present Fraud Detection and Prevention Using Big Data Analytics Algorithms” 2023, 13, 57. https://doi.org/ 10.3390/app13010057, 21 December 2022
25. Ibomoiye Domor Mienye, Yanxia Sun (2023) “A Machine Learning Method with Hybrid Feature Selection for Improved Credit Card Fraud Detection” 2023, 13(12), 7254; https://doi.org/10.3390/app13127254, 18 June 2023
26. Emanuel Mineda Carneiro, Carlos Henrique Quartucci Forster, Lineu Fernando Stege Mialaret, Luiz Alberto Vieira Dias, Adilson Marques da Cunha (2022) “High-Cardinality Categorical Attributes and Credit Card Fraud Detection” 2022, 10(20), 3808; https://doi.org/10.3390/math10203808, 15 October 2022
27. Li, Z., Liu, G., & Jiang, C. (2020). Deep representation learning with fulcenter loss for credit card fraud detection. IEEE Trans. Comput. Soc. Syst., 7(2), 569- 579

28. Almhaithawi, D., Jafar, A., & Aljnidi, M. (2020). Example-dependent costsensitive credit cards fraud detection using SMOTE and Bayes minimum risk. SN Appl. Sci, 2(9), 1-12.
29. Rtayli, N., & Enneya, N. (2020). Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization. J. Inf. Secur. Appl., 55, 102596.
30. Taha, A. A., & Malebary, S. J. (2020). An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine. IEEE Access, 8, 25579-25587.
31. Husejinovic, Admel. (2020). Credit card fraud detection using naive Bayesian and C4.5 decision tree classifiers. 8. 1-5. 10.21533/pen. v%25vi%25i.300. Kajal, D., & Kaur, K. (2021). Credit card fraud detection using imbalance resampling method with feature selection. Int J., 10(3).
32. Carta, S., Fenu, G., Recupero, D.R. and Saia, R., "Fraud detection for E-commerce transactions by employing a prudential Multiple Consensus model," Journal of Information Security and Applications, vol.46, pp.13-22, 2019.
33. Dornadula, V. N., and Geetha, S., "Credit Card Fraud Detection using Machine Learning Algorithms", Procedia Computer Science, vol.165, pp.631–641, 2019.