Detection of Fraudulent Transaction Issues in the Payment Card Industry using Machine Learning: A Comprehensive Survey

M. R. Kalideen

Abstract The increasing prevalence of online payment card transactions has brought with it a surge in fraudulent activities, posing significant challenges to the financial industry. This comprehensive survey examines the utilization of machine learning methodologies to identify fraudulent transactions within the payment card industry. This review examines a diverse array of machine learning algorithms, including advanced deep learning architectures such as Deep Neural Networks, Recurrent Neural Networks, Autoencoders, and Generative Adversarial Networks. The strengths and limitations of these models in the context of fraud detection challenges are thoroughly discussed. The review also analyzes key issues such as imbalanced datasets, model interpretability, scalability, security concerns, and the importance of privacy preservation. Furthermore, we highlight emerging trends such as explainable AI, privacy-preserving machine learning, including federated learning, and the potential of blockchain technology for enhancing fraud detection systems. Finally, we offer actionable recommendations for practitioners and identify promising directions for future research, emphasizing the need for robust, scalable, and ethical AI solutions to combat evolving fraudulent activities in the payment card industry. This research endeavor aims to offer a comprehensive overview of the most recent advancements in machine learning-based methodologies for detecting fraudulent transactions, providing insights into the strengths and limitations of diverse approaches in addressing the distinctive challenges encountered within the payment card industry. This study serves as a valuable academic resource by examining emerging trends and future research directions, providing insights that can aid academics, researchers, and industry professionals in developing more effective and ethical solutions to combat the persistent threat of payment card fraud.

***Index Terms* — Payment Card Fraud Detection, Machine Learning, Deep Learning, Imbalanced Data, Interpretability, Privacy-Preserving AI**

I. INTRODUCTION

THE payment card industry, encompassing credit, debit, and prepaid cards, has witnessed exponential growth, becoming the backbone of global commerce. This surge in digital transactions has unfortunately been paralleled by a rise in fraudulent activities. As the payment landscape evolves, so do the tactics employed by fraudsters, demanding sophisticated countermeasures to safeguard financial systems and stakeholders[1][2][3].

Traditional fraud detection methods often prove inadequate in this dynamic environment. Therefore, the industry has turned to machine learning, a powerful tool adept at analyzing vast datasets and identifying complex patterns indicative of fraudulent behaviour[2][4]. This shift towards machine learning-driven solutions is revolutionizing fraud detection, enabling more accurate and proactive identification of suspicious transactions[5][6]. The ability to learn and adapt from new data makes machine learning a particularly potent weapon in the ongoing battle against payment card fraud[6][7].

MR. Kalideen is a Senior Lecturer at Department of Information and Communication Technology, Faculty of Technology, South Eastern University of Sri Lanka, Oluvil, Sri Lanka (kmr@seu.ac.lk)

Despite the advancements in machine learning-based fraud detection, effectively combating fraudulent activities in the payment card industry remains a significant challenge[3][4][8]. The immense scale of transaction volumes, the continuously shifting nature of fraudulent schemes, and the requirement for real-time detection capabilities collectively contribute to the intricate and challenging nature of this problem domain[3][9]. Fraudsters continuously adapt their tactics to exploit vulnerabilities in existing systems, often outpacing the development and deployment of new security measures[2]. This necessitates a comprehensive understanding of the latest machine learning techniques and their applicability to the dynamic landscape of payment card fraud, considering both their strengths and limitations in addressing the unique challenges posed.

This research aims to explore emerging trends and potential future directions in leveraging machine learning for more robust and adaptive fraud detection systems. This exploration will consider advancements in areas such as deep learning, ensemble methods, and explainable AI, highlighting their potential to enhance fraud prevention strategies. A key focus will be on explainable AI's (XAI) role in increasing transparency and trust in fraud detection models.

ISSN: 2961-5992

Furthermore, this research will analyze the strengths and limitations of various machine learning approaches in addressing the unique challenges associated with payment card fraud. These challenges include imbalanced datasets, real-time detection requirements, model interpretability, and the ability to adapt to evolving fraud patterns. The objective is to offer a clear understanding of the most appropriate machine learning methods for specific fraud detection scenarios.

Finally, this research study will provide a comprehensive review of the existing machine learning techniques that have been utilized for the purpose of identifying fraudulent transactions within the payment card industry. The survey will encompass both conventional and novel approaches, covering supervised, unsupervised, and hybrid learning methodologies. This review will serve as a foundation for understanding the current state of the art and identifying areas for future research and development.

A systematic literature search was conducted, searching for relevant studies in prominent academic databases such as IEEE Xplore, Scopus, and PubMed. The keywords used in the search included terms related to credit card fraud detection, machine learning algorithms, payment card fraud, fraudulent transactions, and anomaly detection. The search was limited to English-language studies published between 2010 and 2024. Initial screening based on title/abstract review was followed by full-text assessment. Studies outside finance or lacking empirical findings were excluded. Of the initial 500 studies, 450 remained after duplicate removal. Abstract screening yielded 150 relevant studies, with 49 meeting the final inclusion criteria. Most of the selected articles were published within the last five years.

II. Literature Review

A. Overview of Payment Card Fraud

Payment card fraud, a pervasive issue within the financial landscape, encompasses any unauthorized use of a payment card, including credit, debit, and prepaid cards, to illicitly obtain funds or goods[3][10]. The exponential expansion of e-commerce and digital payment transactions has been accompanied by a concomitant rise in fraudulent activities directed towards this payment modality[4][7]. Fraudsters employ various tactics, ranging from basic scams like counterfeit cards to more sophisticated schemes involving stolen card data and online account takeovers[7].

The ramifications of payment card fraud are extensive, affecting financial institutions, individual consumers, and merchants alike. Financial losses, eroded trust in payment systems, and the escalating costs of fraud prevention measures represent significant challenges for all stakeholders involved [2][5].

Several factors contribute to the vulnerability of payment cards to fraudulent activities. The rise of e-commerce and card-not-present transactions has created more opportunities for fraudsters, as physical card possession is no longer necessary. This shift to online transactions makes it easier for criminals to operate remotely and target a wider range of victims[3][4].

Large-scale data breaches expose sensitive cardholder information, making it easier for criminals to create counterfeit cards or conduct unauthorized transactions. These breaches can compromise millions of card details at once, providing a wealth of information for fraudsters to exploit[7][8].

Additionally, fraudsters continuously adapt their methods, employing techniques like phishing, malware, and social engineering to compromise card details and exploit system vulnerabilities. The increasing sophistication of these techniques makes it more challenging for individuals and institutions to protect themselves from fraud [3][4][11].

The dynamic landscape of payment card fraud necessitates a proactive and adaptable approach to detection and prevention. Traditional rule-based systems frequently struggle to keep up with emerging fraud patterns, underscoring the need for more advanced solutions like machine learning.

B. Traditional Fraud Detection Models

Conventional fraud detection approaches predominantly leverage rule-based systems and manual review processes. These methods frequently incorporate expert systems, where rules are established based on subject matter experts' knowledge of recognized fraud patterns and anomalies. Transactions flagged by these rules are then investigated further. Statistical analysis, such as anomaly detection techniques, is also used to identify transactions that deviate significantly from expected patterns. Finally, suspicious transactions are reviewed manually by fraud analysts, who make decisions based on their experience and intuition[3][4].

While these traditional methods have been employed for some time, they often prove inadequate in addressing the evolving landscape of payment card fraud. They are often static and inflexible; rule-based systems face difficulties in adapting to emerging fraud patterns, necessitating frequent updates that can be both time-consuming and resource-intensive. Traditional methods are also prone to high false positives, leading to unnecessary investigations and customer inconvenience[2][3][4]. Furthermore, the sheer volume of transactions in today's digital age overwhelms manual review processes and limits the effectiveness of traditional methods. The shortcomings of conventional fraud detection methods underscore the necessity for more advanced and flexible approaches, thereby paving the way for the integration of machine learning techniques within this domain[4][5].

C. Introduction to Machine Learning in Fraud Detection

Machine learning provides a transformative approach to fraud detection, facilitating the development of more precise, adaptable, and efficient systems. In contrast to traditional rule-based methods, machine learning algorithms can learn intricate patterns and relationships from extensive datasets, identifying subtle indicators of fraudulent activity that may otherwise evade detection[2][5][6]. The capacity of machine learning to learn from data and adapt accordingly renders it particularly well-suited for addressing the dynamic and ever-evolving landscape of payment card fraud.

Machine learning algorithms can analyze various data points associated with transactions to identify potentially fraudulent activity. These data points include transaction amount and

ISSN: 2961-5992

frequency, as unusual spending patterns or sudden spikes in transaction volume can be indicative of fraud. The location and time of the transaction are also relevant; transactions originating from unusual locations or at odd hours might raise red flags. Furthermore, analyzing historical data related to merchants and customers can help identify suspicious behavior. Finally, details about the device and network used for the transaction can provide valuable insights into potential fraud[2][3].

By leveraging these data points, can construct predictive models that evaluate the probability of a transaction being fraudulent in real-time. This enables financial institutions to take immediate action, such as declining suspicious transactions or flagging them for further investigation, minimizing potential losses and enhancing overall security.

D.Traditional vs. Machine Learning in Fraud Detection

The following TABLE I highlights the key differences between traditional and machine learning approaches to payment card fraud detection:

TABLE I
TRADITIONAL FRAUD DETECTION VS. MACHINE LEARNING IN FRAUD DETECTION

Feature	Traditional Methods	Machine Learning Methods
Approach	Rule-based, expert systems, manual reviews	Algorithm-driven, data-driven, predictive modeling
Adaptability	Static, inflexible, requires manual updates	Adaptive, learns from data, can identify new patterns
Accuracy	Prone to high false positives, limited by human bias	Potentially higher accuracy, can detect subtle anomalies
Scalability	Difficult to scale with increasing data volumes	Highly scalable, can handle massive datasets
Real-time Detection	Often delayed due to manual review processes	Enables real-time or near real-time fraud detection
Maintenance	Labor-intensive, requires constant rule updates	Requires model training and validation, but can be automated

Conventional fraud detection methods are frequently inflexible and encounter difficulties in keeping up with the dynamic and ever-changing nature of fraudulent activities. On the other hand, machine learning offers a more flexible and scalable approach, enabling the detection of complex patterns and real-time fraud prevention.

However, it is crucial to note that the application of machine learning in fraud detection also presents several challenges. These include the need for extensive, well-labeled datasets, the potential for bias in the training data, and the requirement for interpretable model outputs.

III. CLASSIFICATION OF MACHINE LEARNING TECHNIQUES

A. Supervised Learning Approaches

Supervised learning is a prominent category of machine learning algorithms that leverage labeled data, where historical transactions have been previously classified as either fraudulent or legitimate, to learn the underlying patterns and relationships that distinguish fraudulent activities from genuine ones[2][8]. This approach enables the algorithm to learn the inherent patterns and associations that differentiate fraudulent transactions from legitimate ones.

TABLE II presents a selection of widely used supervised learning algorithms commonly employed for payment card fraud detection:

TABLE II
WIDELY USED SUPERVISED LEARNING ALGORITHMS FOR FRAUD DETECTION

Algorithm	Description	Advantage	Ref.
Logistic Regression	a statistical model that estimates the likelihood of a transaction being fraudulent based on the input features	relatively straightforward technique to implement and interpret, rendering it a common choice for foundational models	[1]
Support Vector Machines	SVMs seek to determine the optimal hyperplane that can distinctly categorize fraudulent and non-fraudulent transactions within a high-dimensional feature space	They are recognized for their capability to manage complex datasets and deliver robust generalization performance	[1]
Decision Trees	construct a hierarchical model of decisions based on various input features, ultimately classifying the transaction as either fraudulent or legitimate.	They are effective for identifying important features and decision rules because they are straightforward to understand and interpret visually.	[5]
Random Forests	Random forests leverage an ensemble of decision trees to enhance predictive accuracy and mitigate the risk of overfitting the data.	Random forests exhibit robust behavior in the presence of outliers and can effectively handle high-dimensional datasets.	[1]
Neural Networks	Complex models modelled after the human brain, neural networks are made up of interconnected nodes that process and learn from data.	They are very effective at managing big and complicated information and are able to capture non-linear correlations	[7]

The selection of the most appropriate supervised learning algorithm hinges on various factors, including the dataset's size and quality, the complexity of the fraud patterns, and the desired equilibrium between accuracy, interpretability, and computational efficiency.

B. Unsupervised Learning Approaches

In contrast to supervised learning, which relies on datasets with labeled outcomes, unsupervised learning algorithms are trained on unlabeled data where the target variable is unknown a priori.

ISSN: 2961-5992

These algorithms seek to uncover latent patterns, anomalies, and relationships within the data without explicit direction or supervision. This characteristic makes unsupervised learning particularly valuable in fraud detection for identifying previously unknown fraud patterns that might not be captured by labeled data [12][13].

Unsupervised learning algorithms commonly employed for fraud detection include clustering techniques. These algorithms group similar transactions together based on their shared characteristics and attributes. Transactions that fall outside of these clusters, or into clusters known to be associated with fraudulent behavior, can be flagged as suspicious [3][4]. Anomaly detection algorithms aim to identify data points that exhibit significant deviations from the expected or typical pattern within the data [6][10]. In the domain of payment card fraud detection, anomalies might represent transactions with unusual amounts, locations, or spending patterns [3][10]. Dimensionality reduction methodologies such as Principal Component Analysis can diminish the complexity of data by determining the principal features that account for the predominant variance within the information[2][14]. Dimensionality reduction techniques like Principal Component Analysis can assist in data visualization, the identification of latent patterns, and the enhancement of other machine learning algorithms' performance.

Unsupervised learning approaches offer several advantages in fraud detection. They can uncover new and evolving fraud schemes that might not be present in labeled datasets. Unsupervised learning reduces the need for extensive data labeling, which can be time-consuming and expensive. It can also complement supervised learning by being used to pre-process data, identify features, or generate labels for supervised learning algorithms, enhancing their overall effectiveness [4][13]. However, unsupervised learning also presents challenges, such as the difficulty in evaluating the performance of models without ground truth labels and the potential for false positives due to the inherent nature of anomaly detection[15][16].

C. Semi-supervised and Hybrid Methods

Apart from supervised and unsupervised learning, semi-supervised and hybrid approaches provide different ways to capitalise on the advantages of both approaches, particularly in situations where labelled data is hard to come by or prohibitively expensive.

Semi-Supervised Learning: By using both labelled and unlabelled data during training, semi-supervised learning fills the gap between supervised and unsupervised learning. Given the abundance of unlabelled data and the rarity of labelled fraudulent transactions, this method is especially helpful in fraud detection [2][16].

One way to apply semi-supervised learning is by pre-training with unsupervised learning. Based on their innate qualities, comparable transactions can be grouped together using an unsupervised learning method like clustering [4][13]. Then, a limited set of labeled transactions can be used to assign labels to the clusters, effectively propagating the known labels to a larger portion of the unlabeled data[4][6][14].

Hybrid Methods: Several machine learning techniques are combined in hybrid approaches to improve the robustness and accuracy of fraud detection. This could involve integrating supervised and unsupervised algorithms or combining different types of supervised algorithms. For example, ensemble approaches build a more reliable and effective fraud detection system by combining predictions from several supervised learning models, including decision trees, support vector machines, and neural networks[17][18]. Another hybrid approach involves unsupervised feature learning, where techniques like autoencoders learn compressed representations of the data, which can then be used as input features for a supervised learning algorithm[19].

By leveraging the strengths of different learning paradigms, semi-supervised and hybrid methods offer promising avenues for improving fraud detection accuracy, particularly in situations with limited labeled data or complex, evolving fraud patterns.

D. Deep Learning Techniques

Deep learning, because of its capacity to automatically recognise intricate patterns and process large volumes of data, this subset of machine learning has become increasingly popular in the field of fraud detection. Multiple layers of interconnected nodes are used by deep learning models, such deep neural networks, to extract increasingly abstract representations of input. This allows them to spot complex relationships and abnormalities that could be signs of fraud [20][21].

Several prominent deep learning techniques have been employed for fraud detection, as outlined in TABLE III.

TABLE III
DEEP LEARNING TECHNIQUES USED IN FRAUD DETECTION

Technique	Description	Use in Fraud Detection	Example
Deep Neural Networks	Multiple hidden layers excel at capturing non-linear relationships	Detecting complex fraud patterns	[8][22] suggested a DNN method that, when tested on real-world datasets, achieved great accuracy in detecting credit card fraud. A DNN, for example, might examine different transaction characteristics (amount, location, and time) to find minute patterns suggestive of fraud that more basic models could overlook
Recurrent Neural Networks	Designed to manage sequential data	Analyzing transactions over time	[14]used a GRU-centered sandwich-structured model (a type of RNN) for transaction fraud detection. This model can efficiently examine transaction sequences to find irregularities that could point to fraud, including an abrupt spike in expenditure or odd transaction locations
Autoencoders	Neural networks trained to reconstruct input data	Learn a compressed representation of normal transactions;	[4] discusses the use of autoencoders for fraud detection. An example would be training an autoencoder on a dataset of legitimate

ISSN: 2961-5992

		deviations are flagged as potential fraud	transactions. When presented with a fraudulent transaction, the autoencoder would likely have a higher reconstruction error, flagging it as potentially fraudulent
Generative Adversarial Networks	Two networks (generator and discriminator) trained adversarially	Generator creates synthetic fraudulent transactions; discriminator distinguishes real from synthetic fraud. Used to augment training data.	[9] created fictitious fraudulent transactions using GANs to supplement the training data fo a fraud detection model. This method can enhance model performance, particularly when working with datasets that are imbalanced and have a lower frequency of fraudulent transactions than valid ones.

Deep learning models have exhibited exceptional performance in fraud detection, outperforming conventional methods. These models possess the ability to automatically extract salient features from raw data, reducing the need for laborious manual feature engineering[8]. Finally, their proficiency in capturing intricate, non-linear correlations allows them to identify subtle patterns of deception [14]. However, deep learning also presents challenges. For training, these models usually need enormous volumes of labelled data, which might be challenging to find in fraud detection. Training may require specialised hardware and be computationally costly. Furthermore, because of their intricacy, deep learning models are frequently referred to as "black boxes" because it is difficult to comprehend their predictions[4][16]. Despite these challenges, deep learning continues to drive advancements in fraud detection, offering promising avenues for enhancing accuracy, efficiency, and the ability to combat evolving fraud techniques[4][6].

E. Ensemble Methods

In machine learning, ensemble methods combine several learning algorithms to achieve higher predicted performance than any one of the individual learning algorithms could. The theory is that by integrating several models, each with unique advantages and disadvantages, the ensemble can get beyond the drawbacks of each model and improve accuracy, resilience, and generalisation capacity.

Ensemble approaches have demonstrated great potential in the domain of fraud detection because of their capacity to manage intricate data patterns and enhance prediction accuracy.

TABLE IV briefly explains some key ensemble methods used in fraud detection.

Ensemble methods in machine learning offer several key advantages. Higher overall accuracy is frequently achieved by combining predictions from several models, particularly when the base models are varied and capture various facets of the data. This diversity helps to mitigate the weaknesses of individual models and leverage their strengths[23][24]. In comparison to single models, ensembles are typically more resilient to data noise and outliers. Individual model errors typically cancel each other out, producing a prediction that is more solid and trustworthy [24]. Finally, by reducing overfitting, ensembles typically generalize better to unseen data. When a model learns the training data too thoroughly, including its noise and specificities, it is said to be overfitting and performs poorly on new data. Ensembles mitigate this by averaging out the idiosyncrasies of individual models [23].

TABLE IV
ENSEMBLE METHODS USED IN FRAUD DETECTION

Ensemble Method	Description	Advantages in Fraud Detection
Bagging	Creates multiple training subsets by random sampling with replacement. Combines predictions from individual models that have been trained on each subset. (majority vote or averaging).	Reduces variance, improves stability.
Boosting	Sequentially builds an ensemble, weighting misclassified instances more in each iteration. New models correct errors of previous models. Examples: AdaBoost, Gradient Boosting, XGBoost.	Creates a strong learner with high accuracy.
Random Forests	Extends bagging with random feature selection at each decision tree split. Further decorrelates trees and improves generalization.	Improves generalization.
Stacking	uses base model predictions as a meta-model's input to combine models. Combining base model predictions is something the meta-model learns to do.	Captures higher-order interactions between models.

Despite their benefits, ensemble methods also have some limitations. Training and assessing ensembles can be computationally expensive, particularly when working with large datasets and complex base models. The cost increases with the number of models and their individual complexity[23]. Furthermore, ensembles can be more complex to build, train, and deploy than individual models. Managing multiple models, ensuring their diversity, and combining their predictions adds complexity[24]. Lastly, while individual models within an ensemble might be interpretable, understanding the ensemble's predictions as a whole can be challenging. The combined decision-making process can obscure the reasoning behind the final prediction, making it difficult to explain or analyze[25].

Overall, ensemble methods provide a powerful approach to enhance fraud detection accuracy and robustness. In order to create more efficient fraud detection systems, practitioners can take use of the advantages of various learning algorithms by carefully choosing and combining suitable base models.

IV. COMPARISON OF MACHINE LEARNING TECHNIQUES FOR FRAUD DETECTION

TABLE V offers a comparative summary of the relative advantages and disadvantages of several machine learning algorithms that are applicable to fraud detection.

ISSN: 2961-5992

TABLE V
COMPARISON OF MACHINE LEARNING TECHNIQUES USED IN FRAUD DETECTION

Technique	Description	Advantages	Disadvantages
Supervised Learning	Learns from labeled data to predict outcomes for unseen data.	High accuracy when trained on sufficient labeled data. Relatively easy to interpret and understand.	Requires significant quantities of labelled data, which can be costly and time-consuming to acquire. If the training data is not representative, it might not generalise well to new data.
Unsupervised Learning	Learns patterns and structures from unlabeled data without explicit guidance	can find abnormalities and hidden patterns in data without the requirement for labels. beneficial for feature engineering and exploratory data analysis	Results can be difficult to interpret and may not always be directly applicable to fraud detection. Evaluation of performance can be challenging without ground truth labels
Semi-Supervised / Hybrid	Combines aspects of both supervised and unsupervised learning, leveraging both labeled and unlabeled data	Can achieve good performance with limited labeled data by leveraging unlabeled data. Offers flexibility in combining different learning paradigms to address specific challenges	Model complexity can increase, making interpretation and training more challenging. Requires careful selection and integration of appropriate techniques
Deep Learning	uses multi-layered deep neural networks to extract intricate patterns and representations from input	reaches cutting-edge results in a variety of fraud detection tasks. Excels are adept at managing intricate, non-linear relationships in data and can automatically extract pertinent elements from raw data	huge volumes of labelled data are needed for training. Computationally expensive to train, requiring specialized hardware and significant processing time. Model interpretability can be challenging due to complex architectures
Ensemble Methods	Combines multiple learning algorithms to improve prediction accuracy, robustness, and generalization ability	makes use of the advantages of many algorithms to frequently surpass individual models. more resilient to data noise and anomalies. "-" Better at generalising to unknown data.	Can be more complex to build, train, and deploy compared to individual models. Interpretability of the ensemble as a whole can be challenging. Training and evaluating ensembles can be computationally expensive

V. EVALUATION METRICS AND DATASETS

A. Common Evaluation Metrics

It is essential to evaluate fraud detection algorithms' performance in order to determine their efficacy and make well-informed deployment decisions. However, conventional measurements like accuracy might be deceptive because fraud datasets are imbalanced (fraudulent transactions usually make up a small portion of the total data). Therefore, to give a more thorough picture of a model's effectiveness in fraud detection, particular evaluation criteria are employed. These are a few widely used measures for evaluation [26]:

Precision: Out of all transactions that are projected to be fraudulent, precision quantifies the percentage of accurately predicted fraudulent transactions. It answers the question: "Of all the transactions the model flagged as fraud, how many were actually fraudulent?"

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Recall: The proportion of actual fraudulent transactions that the model properly identifies is known as recall. It answers the question: "Of all the actual fraudulent transactions, how many did the model correctly identify?"

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

F1-Score: The F1-score offers a fair assessment of precision and recall since it is the harmonic mean of the two variables. It is especially helpful in situations when the distribution of classes is not uniform.

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Area Under the Receiver Operating Characteristic Curve: The trade-off between the true positive rate (sensitivity) and the false positive rate (1 - specificity) at different classification thresholds is shown graphically by the AUC-ROC curve. Better model performance is indicated by a higher AUC-ROC value.

Average Precision: By delivering a single value that represents the average of precision values at various recall levels, AP condenses the precision-recall curve.

Matthews Correlation Coefficient: The confusion matrix's four values—true positives, true negatives, false positives, and false negatives—are all taken into account by the balanced MCC metric. It goes from -1 to +1, where +1 denotes a perfect forecast, 0 a random guess, and -1 a total discrepancy between the prediction and the observation.

These assessment metrics offer a thorough understanding of a fraud detection model's performance by taking into account both the rate of false positives and the model's capacity to accurately identify fraudulent transactions. The particular needs and goals of the application will determine which statistic is best. For instance, in high-stakes scenarios where minimizing false negatives is paramount, recall might be prioritized. Conversely, if minimizing false positives is crucial to avoid disrupting legitimate transactions, precision might be more important.

ISSN: 2961-5992

B. Comparing Evaluation Metrics for Credit Card Fraud Detection

It is necessary to carefully analyse a variety of evaluation measures when assessing the effectiveness of machine learning models for credit card fraud detection. Relying only on accuracy can be deceptive because fraud datasets are inherently imbalanced (fraudulent transactions are far less common than valid ones)[27][2]. TABLE VI depicts the comparison of different evaluation metrics commonly used in this domain:

TABLE VI
COMPARISON OF DIFFERENT EVALUATION METRICS USED IN MACHINE LEARNING

Metric	Description	Strengths	Weaknesses
Accuracy	The proportion of accurately classified transactions to all transactions	Easy to understand. High accuracy can be achieved by simply classifying all transactions as non-fraudulent for imbalanced datasets.	Not recommended for imbalanced fraud datasets
Precision	The proportion of fraudulent transactions that were accurately predicted out of all those that were projected to be fraudulent	aims to reduce false positives, or the quantity of valid transactions that are mistakenly reported as fraudulent	May not capture all actual fraudulent transactions, especially if the model is too conservative in flagging fraud. When minimizing false positives is crucial, such as in scenarios where incorrectly flagging a legitimate transaction as fraud can lead to significant customer dissatisfaction
Recall	proportion of actual fraudulent transactions that the model accurately detected	Focuses on minimizing false negatives (i.e., reducing the quantity of fraudulent transactions that are not discovered). This approach may result in a greater number of false positive classifications if the model is overly aggressive in identifying potential fraud instances	Prioritizing the minimization of false negatives, even if it results in a higher number of false positives, is crucial. Prioritizing the minimization of false negatives is critical in high-stakes scenarios where the failure to detect a fraudulent transaction can result in severe repercussions
F1-Score	A balanced assessment of both measures is provided by the harmonic mean of precision and recall..	is appropriate for imbalanced datasets since it offers a single score that strikes a balance between precision and recall	May not be as intuitive to interpret as precision or recall individually. When achieving a balanced measure that takes into account both false positives and false negatives.
AUC-ROC	Plotting the genuine positive rate against the false positive rate at different categorisation thresholds is done via the Receiver Operating Characteristic curve's area under the curve.	enables a trade-off analysis between true positive and false positive rates by offering a thorough picture of model performance across various thresholds.	Can be less intuitive to interpret than other metrics, especially for business stakeholders. When a trade-off analysis between true positive and false positive rates is required and a thorough understanding of model performance across various thresholds is needed
Average Precision	provides a single value that represents the average of precision values at various recall levels, summarising the precision-recall curve.	offers a single score that considers the precision-recall trade-off across different thresholds.	Can be less intuitive to interpret than precision or recall individually. When a single score that summarizes the precision-recall trade-off is desired

The particular objectives and limitations of the fraud detection system determine which metric is best. For instance, if minimizing customer inconvenience caused by false positives is a top priority, precision would be a key metric. Conversely, if detecting as many fraudulent transactions as possible is crucial, even at the cost of some false positives, recall would be more important. Often, a more thorough evaluation of model performance can be obtained by combining metrics like F1-score and AUC-ROC.

VI. DATASETS USED IN RESEARCH

Research on credit card fraud detection frequently uses a number of datasets. Labelled credit card transactions that distinguish between fraudulent and lawful operations are included in one popular dataset that is accessible on Kaggle. However, it suffers from class imbalance and lacks certain real-world features due to privacy concerns[28]. Some research papers, such as [29] and [5], likely utilize relevant datasets, but the specifics aren't always detailed in readily available information. Another resource, the [30] paper, highlights the significance of realistic datasets and evaluation processes by introducing a specific benchmark for fraud detection; nevertheless, it does not name a specific dataset for credit card fraud detection.

It is important to keep in mind that publicly available datasets often undergo anonymization and feature engineering to protect sensitive information, potentially limiting their representativeness of real-world fraud patterns[16]. Researchers are exploring techniques like Generative Adversarial Networks to create synthetic datasets that mimic real-world distributions while preserving privacy[9]. Selecting an appropriate dataset is essential for developing and evaluating fraud detection models. Consider factors like the dataset's size, features, class distribution, and relevance to the specific fraud detection task[16].

ISSN: 2961-5992

A. Challenges and Limitations

Imbalanced Datasets

The imbalanced nature of datasets is one of the biggest obstacles to detecting credit card fraud. Less than 1% of all transactions are fraudulent, which is a very small percentage. This imbalance can lead to several issues. When trained on such data, machine learning models are typically biassed in favour of the majority class (legitimate transactions). This occurs because the model can achieve high accuracy by simply classifying most transactions as legitimate, even if it misses many fraudulent ones. In addition to not detecting fraudulent transactions that show patterns distinct from the few fraudulent examples in the training data, models built on imbalanced data may also perform poorly when applied to unseen data[31][32][33][34].

These issues can be resolved in a number of ways. The class ratio can be altered via sampling strategies such as oversampling, which replicates or creates synthetic instances for the minority class, and undersampling, which eliminates instances from the majority class. But whereas oversampling might result in overfitting, undersampling can result in information loss[35]. By penalising false negatives more severely than false positives, cost-sensitive learning encourages the model to focus more on the minority class by allocating distinct misclassification costs to various classes[10]. Algorithmic approaches, such as ensemble methods like Random Forest and XGBoost, are inherently better suited for handling imbalanced datasets compared to traditional algorithms like Logistic Regression[16].

Addressing imbalanced datasets is crucial for developing effective fraud detection models. Techniques to lessen the effects of class imbalance and increase the precision and dependability of fraud detection systems are still being investigated and improved by researchers and practitioners. Many studies highlight these challenges and discuss various mitigation techniques, emphasizing the importance of carefully considering the class distribution and employing appropriate methods for building robust and effective models.

Model Interpretability

Although deep learning models in particular have demonstrated impressive accuracy in detecting credit card fraud, they frequently lack transparency. The fundamental mechanisms and decision-making procedures of these "black box" models are difficult for humans to understand. This lack of interpretability is a major problem, particularly in the financial industry where trust and regulatory compliance depend on knowing the reasoning behind a model's prediction[25][23].

Model interpretability is crucial for a number of reasons. Financial institutions must have trust in their fraud detection models' predictions. Understanding why a model flags a transaction as fraudulent is crucial for investigators and for customers to understand declined transactions. Regulations often require financial institutions to provide explanations for decisions made by automated systems, especially those impacting customers. Additionally, interpretability aids in finding possible biases or errors in the model's judgement, directing model enhancement and producing stronger and more dependable systems [36][37].

Several techniques can enhance model interpretability. More transparency can be achieved by using models that are naturally interpretable, like decision trees or linear models, but they might not be as accurate as more sophisticated ones[12][25]. Model-agnostic interpretability methods, applied after training and usable with any model type, include feature importance analysis (identifying important features), partial dependence plots (visualizing relationships between features and predictions), and surrogate models (training simpler models to mimic complex ones)[12][25].

One important factor to take into account is the trade-off between interpretability and accuracy. While complex models might offer slightly better accuracy, their lack of transparency can hinder trust and limit their practical use in finance. Finding the correct balance is a constant challenge.

Scalability and Real-Time Detection

Systems for detecting credit card fraud must be able to process enormous amounts of transactions instantly, posing significant challenges in terms of scalability and processing speed. In order to handle the constantly increasing number of transactions that financial institutions process, fraud detection methods must be scalable. This involves handling large datasets for training and making predictions quickly on massive incoming data streams. Traditional models may struggle to scale to handle the demands of real-time detection in high-volume settings. Fraud detection systems must also operate in real-time to prevent fraudulent transactions from being authorized, necessitating models that can process incoming data and make predictions with minimal latency. Delays in detection can result in significant financial losses[38].

Several approaches can address these challenges. Leveraging distributed computing frameworks like Apache Spark enables processing large datasets and training models across multiple machines, improving scalability. Employing stream processing technologies like Apache Kafka facilitates real-time data ingestion and processing, enabling the detection of fraudulent transactions as they occur. Optimizing machine learning models for faster prediction times is crucial for real-time detection. Model compression and quantisation are two methods that can lower the computing complexity of the model without appreciably compromising accuracy [39].

Balancing scalability and real-time performance with accuracy remains an ongoing challenge. Researchers and practitioners continue to explore innovative approaches, including distributed computing, stream processing, and model optimization techniques, to create systems that can manage enormous amounts of data and efficiently identify fraudulent transactions in real time.

Security and Adversarial Attacks

Machine learning models, while effective in detecting known fraud patterns, are vulnerable to adversarial attacks. In order to trick the model into producing inaccurate predictions, these

ISSN: 2961-5992

attacks entail altering input data. In credit card fraud detection, adversaries could craft transactions that appear legitimate but are, in fact, fraudulent.

Security and adversarial attacks are significant concerns due to evolving fraud tactics. Fraudsters constantly adapt their techniques to circumvent detection mechanisms, exploiting vulnerabilities in models by understanding the features the model relies on and crafting transactions that evade detection. The field of adversarial machine learning specifically focuses on developing techniques to attack and defend machine learning models. As research advances, fraudsters can leverage these techniques to develop more sophisticated attacks. Many machine learning models are not inherently robust to adversarial perturbations in the input data[40][41]. Even small, carefully crafted changes can lead to misclassifications.

Adversarial and security threats necessitate a multifaceted strategy. Using adversarial samples to train the model is known as adversarial training, exposing it to various attack scenarios during training to make it more robust. Various defensive mechanisms, such as input sanitization, anomaly detection, and ensemble methods, can be employed to prevent adversarial examples from reaching the model or mitigate their impact[41]. To stay up with changing fraud strategies and hostile attacks, constant observation and adjustment are required. To do this, the model must be updated frequently with fresh data, retraining it on new attack patterns, and incorporating new defensive mechanisms [3][4].

The security of machine learning-based fraud detection systems is seriously threatened by the growing complexity of adversarial attacks. Addressing these security and adversarial challenges is crucial for building trustworthy and reliable fraud detection systems. As machine learning plays an increasingly vital role in combating financial fraud, ensuring the security and resilience of these systems against adversarial attacks is still a topic of current investigation and development.

VII. Emerging Trends and Future Directions

A. AI and Future Directions

The use of artificial intelligence in fraud detection is becoming more and more important. Deep learning and reinforcement learning are two examples of advanced machine learning models that are increasingly being used. These models can analyze vast and complex datasets to identify subtle patterns indicative of fraudulent activity, often outperforming traditional rule-based systems[39]. AI makes it possible to process transactional data in real-time, which enables prompt fraud identification and prevention. This competence is essential for reducing losses and preserving financial systems' integrity [42][43]. Furthermore, AI is facilitating the integration of biometric authentication methods, such as facial recognition and voice recognition, into fraud prevention strategies[42]. These methods provide an additional layer of security by verifying user identities more robustly. The continuous evolution of AI promises to further revolutionize fraud detection strategies, enabling more accurate, efficient, and secure systems for combating financial crime.

B. Explainable AI (XAI)

Explainable AI is crucial for transparency and interpretability in AI-driven fraud detection. As models become more complex, understanding their predictions is paramount to building trust and confidence. XAI boosts system confidence by enabling human analysts to comprehend the reasons behind a transaction being detected as fraudulent. Furthermore, it helps meet regulatory requirements in many financial sectors that demand transparency in decision-making processes involving sensitive financial data. Finally, understanding the model's reasoning can reveal biases or limitations in the data or the model itself, leading to improvements in accuracy and effectiveness[36].

XAI has several real-world applications in fraud detection. It enhances decision support by providing analysts with insights into why a transaction is flagged as suspicious, allowing for more informed decisions and reducing false positives[44][45]. XAI also facilitates regulatory compliance and auditability by providing auditable trails of how fraud detection models arrive at their conclusions. Additionally, it aids in model debugging and improvement by uncovering biases or limitations, allowing data scientists to refine models and improve accuracy[12].

However, ethical implications of XAI in fraud detection must be considered. Bias and fairness are paramount, ensuring that explanations are unbiased and do not perpetuate societal biases[46]. Privacy concerns must be addressed by balancing transparency with privacy protection, avoiding the inadvertent revelation of sensitive information[47]. Finally, overreliance and automation bias should be mitigated by maintaining human oversight and judgment in the fraud detection process[4][16]. The design, implementation, and continuous monitoring of XAI systems must be carefully considered in order to address these ethical issues, and cooperation between data scientists, ethicists, and subject matter experts is crucial.

C. Privacy-Preserving Machine Learning and Fraud Detection

The increasing use of sensitive personal data for fraud detection necessitates a strong emphasis on privacy preservation. Privacy-preserving machine learning (PPML) is becoming critical for future fraud detection systems. Fraud detection often involves highly sensitive financial and personal data, raising significant privacy concerns. Strict data protection laws, such the CCPA and GDPR, require that personal information be protected when processing data [26][45]. Using PPML strategies can increase client trust by showcasing a dedication to safeguarding their privacy.

Several PPML techniques are relevant for fraud detection. Without exchanging raw data, federated learning allows cooperative model training across several dispersed devices or computers[48][13]. In fraud detection, this could allow banks to train models on a larger pool of data without compromising customer privacy[2]. Differential privacy ensures that individual data points cannot be deduced from the findings while maintaining the dataset's overall statistical characteristics by adding precisely calibrated noise to the data or model parameters[47]. Training fraud detection models on sensitive data without jeopardising privacy is made possible by homomorphic

ISSN: 2961-5992

encryption, which permits calculations on encrypted data without the need for decryption [47][49].

However, future directions and challenges remain. Balancing privacy and utility, finding the best balance between between privacy preservation and model accuracy, is an ongoing challenge. Scalability and efficiency are also concerns, as implementing PPML techniques in real-world systems with massive datasets requires addressing computational efficiency. Finally, standardization and adoption are crucial for wider adoption and interoperability, requires the creation of best practices and industry standards for PPML in fraud detection.

VIII. Discussion and Implications

A. Synthesis of Findings

Conventional rule-based fraud detection systems, while easily understood, struggle to adapt to the ever-changing tactics employed by fraudsters, often leading to high rates of false positives. Machine learning offers a powerful alternative by enabling the detection of subtle patterns and anomalies that traditional methods may miss. However, applying machine learning to fraud detection has its own set of challenges, including imbalanced datasets, model interpretability, scalability, and security concerns, all of which require careful consideration.

Emerging trends like explainable AI and privacy-preserving machine learning offer promising solutions to these challenges. XAI enhances transparency and trust in fraud detection models, while PPML addresses growing privacy concerns associated with the use of sensitive financial data. Federated learning, for example, allows for collaborative model training without directly sharing sensitive data.

Furthermore, integrating blockchain technology holds significant potential for revolutionizing fraud detection. Blockchain can provide immutable audit trails, enhance Know Your Customer and Anti-Money Laundering compliance, and enable secure data sharing. Despite these advancements, a significant research gap remains in addressing the scalability and efficiency of these emerging technologies for real-world deployment.

Future research should prioritize several key areas: developing robust PPML techniques that effectively balance privacy and model accuracy; improving blockchain-based fraud detection systems' scalability and computational effectiveness; and creating industry best practices and standards for the moral and responsible application of AI and machine learning in fraud detection.

B. Implications for Practice

The knowledge gained from this assessment of the literature has several important ramifications for professionals creating and implementing fraud detection systems. Organizations should consider a gradual transition from purely rule-based systems to hybrid models that combine rules and machine learning, leveraging the strengths of both approaches. Explainable AI techniques should be prioritized to foster trust and ensure responsible decision-making by providing clear explanations for model predictions.

Adopting privacy-preserving machine learning strategies like federated learning, differential privacy, and homomorphic encryption is essential given the sensitive nature of the data involved. These methods assist in striking a compromise between the necessity of precise fraud detection and the requirement to safeguard consumer privacy. Practitioners should also explore integrating blockchain technology to enhance data security, streamline Know Your Customer/Anti Money Launering (KYC/AML) compliance, and enable secure data sharing.

Finally, addressing the complex challenges of fraud detection requires a collaborative effort. Industry stakeholders, including financial institutions, technology providers, and researchers, should actively share knowledge, best practices, and lessons learned to drive innovation and improve collective defenses against fraud. By embracing these implications, practitioners can contribute to developing and deploying more robust, transparent, and privacy-preserving fraud detection systems that effectively combat evolving fraudulent activities while maintaining customer trust and adhering to ethical considerations.

C. Implications for Research

This literature review highlights several promising avenues for future research in fraud detection. One key area is enhancing Privacy-Preserving Machine Learning for real-world deployment. This involves developing more efficient and scalable PPML techniques capable of handling massive datasets, exploring novel approaches to balance privacy-utility trade-offs, reducing computational overhead, and addressing limitations of existing techniques like federated learning.

Another crucial research area is robustness and adversarial learning. As fraudsters become more sophisticated, developing robust fraud detection models resilient to adversarial attacks is essential. This entails investigating methods for anomaly detection, model hardening, and adversarial training.

Explainable AI for fraud detection is also a critical area for future research. Developing XAI methods tailored for the complexities of fraud detection, creating user-friendly explanations for both technical and non-technical stakeholders, and developing evaluation metrics for XAI are all important research directions.

Integrating blockchain and AI for fraud detection is another promising avenue. Research should explore how blockchain can enhance data security, provenance, and transparency in AI-driven fraud detection systems, including investigating the use of smart contracts for secure data sharing and decentralized identity verification.

Finally, it is critical to conduct study on the moral and societal ramifications of AI in fraud detection. This involves investigating potential biases in training data and model predictions, ensuring fairness and non-discrimination, and developing guidelines for responsible AI use in fraud detection. Addressing these research implications can significantly advance the field and lead to more effective, secure, and ethical solutions for combating financial fraud.

ISSN: 2961-5992

IX. CONCLUSION

This study offers a thorough analysis of fraud detection, looking at its development, the rise of machine learning, and the potential and problems that go along with it. It provides a comprehensive resource for comprehending the background, present situation, and potential future paths of fraud detection, covering conventional approaches, machine learning strategies, and cutting-edge technologies.

The review critically examines the limitations of existing approaches, pointing out important issues such as imbalanced data sets, interpretability of the model, scalability, and security issues, and the need for privacy preservation. It also provides valuable insights into the potential of emerging trends like explainable AI, privacy-preserving machine learning (specifically highlighting techniques like federated learning), and blockchain technology to address current challenges and shape the future of fraud detection.

Furthermore, the review offers actionable recommendations for industry practitioners, guiding them in transitioning towards hybrid systems, embracing explainable AI, prioritizing privacy-preserving techniques, exploring blockchain integration, and fostering collaboration. Finally, it suggests promising directions for further study, promoting the creation of more reliable and scalable PPML methods, strong models resistant to hostile attacks, explainable AI techniques specifically designed for fraud detection, the cooperative integration of blockchain and AI, and an emphasis on the moral and societal ramifications of AI in this field.

For researchers, practitioners, and policymakers interested in preventing financial fraud, this study is an invaluable resource because it summarises current knowledge and identifies areas that require more investigation.

There is a continuous arms race in the fight against financial fraud. Since scammers are always changing their strategies, staying ahead of the curve requires a dynamic and multifaceted approach. While machine learning offers a powerful arsenal for combating fraud, it's not a complete solution on its own.

The future of fraud detection lies in a holistic strategy combining human expertise, advanced technologies like AI and blockchain, and a commitment to ethical considerations. Building safe and reliable financial systems will require more study in fields like Privacy-Preserving Machine Learning, especially federated learning, and the creation of strong, explicable AI models. Everyone can benefit from a safer and more robust financial ecosystem if researchers, practitioners, and legislators work together.

REFERENCES

[1] G. S. Sowmya and H. K. Sathisha, "Detecting Financial Fraud in the Digital Age: The AI and ML Revolution," Sep. 04, 2023. doi: 10.36948/ijfmr.2023.v05i05.6139.

[2] W. Yundong, Alexander Zhulev, and Omar G. Ahmed, "Credit Card Fraud Identification using Logistic Regression and Random Forest ", WJCMS, vol. 2, no. 3, pp. 1–8, Sep. 2023, doi: 10.31185/wjcms.184.

[3] Akarshan Kumar, Bhagyashri R Hanji, Akash Roy, Ayman Saleem, Ayush Chandak. "Credit Card Fraud Detection using Machine Learning", International Journal of Engineering Development and Research (IJEDR), ISSN:2321-9939, Vol.12, Issue 3, pp.e220-e225, March 2024, URL :http://www.ijedr.org/papers/IJCRT2403513.pdf.

[4] D. R. M. and S. N. Jagadeesha, "Credit Card Fraud Detection using Machine Learning and Data Mining Techniques - a Literature Survey," Jul. 28, 2023. doi: 10.47992/ijaeml.2581.7000.0186.

[5] Md. A. Talukder, R. Hossen, Md. A. Uddin, M. N. Uddin, and U. K. Acharjee, "Securing transactions: a hybrid dependable ensemble machine learning model using IHT-LR and grid search," Nov. 02, 2024, Springer Nature. doi: 10.1186/s42400-024-00221-z.

[6] V. Kumar and R. Pahwa, "Credit Card Fraud Detection Using Machine Learning," Computational Intelligence and Machine Learning, vol. 4, no. 1, pp. 39–45, Apr. 2023. doi: 10.36647/CIML/04.01.A009.

[7] V. N. Dornadula and S. Geetha, "Credit Card Fraud Detection using Machine Learning Algorithms," Jan. 01, 2019, Elsevier BV. doi: 10.1016/j.procs.2020.01.057.

[8] A. S. Rawat and S. Tiwari, "A comprehensive review on credit card fraud detection using machine learning techniques," International Journal of Innovative Research and Growth, vol. 12, no. 2. May 31, 2023. doi: 10.26671/ijirg.2023.2.12.103.

[9] N. T. Ali, S. J. Hasan, A. Ghandour, and Z. S. Al-Hchimy, "Improving credit card fraud detection using machine learning and GAN technology," Jan. 01, 2024, EDP Sciences. doi: 10.1051/bioconf/20249700076.

[10] Y. Lucas and J. Jurgovsky, "Credit card fraud detection using machine learning: A survey," Jan. 01, 2020, Cornell University. doi: 10.48550/arxiv.2010.06479.

[11] N. Uchhana, R. Ranjan, S. Sharma, D. Agrawal, and A. Punde, "Literature Review of Different Machine Learning Algorithms for Credit Card Fraud Detection," Apr. 30, 2021, Blue Eyes Intelligence Engineering and Sciences Publication. doi: 10.35940/ijitee.c8400.0410621.

[12] G. I. Allen, L. Gan, and L. Zheng, "Interpretable Machine Learning for Discovery: Statistical Challenges & Opportunities." Aug. 04, 2023.

[13] S. Banabilah, M. Aloqaily, E. Alsayed, N. Malik, and Y. Jararweh, "Federated learning review: Fundamentals, enabling technologies, and future applications," Aug. 26, 2022, Elsevier BV. doi: 10.1016/j.ipm.2022.103061.

[14] X. Li et al., "Transaction Fraud Detection Using GRU-centered Sandwich-structured Model," May 01, 2018. doi: 10.1109/cscwd.2018.8465147.

[15] N. Rodríguez-Barroso, D. Jiménez-López, M. V. Luzón, F. Herrera, and E. Martínez-Cámara, "Survey on federated learning threats: Concepts, taxonomy on attacks and defences, experimental study and challenges," Sep. 15, 2022, Elsevier BV. doi: 10.1016/j.inffus.2022.09.011.

[16] P. Grover et al., "Fraud Dataset Benchmark and Applications," Jan. 01, 2022, Cornell University. doi: 10.48550/arxiv.2208.14417.

[17] X. Niu, L. Wang, and X. Yang, "A Comparison Study of Credit Card Fraud Detection: Supervised versus Unsupervised," Jan. 01, 2019, Cornell University. doi: 10.48550/arxiv.1904.10604.

[18] D. H. M. de Souza and C. J. Bordin, "Ensemble and Mixed Learning Techniques for Credit Card Fraud Detection," Jan. 01, 2021, Cornell University. doi: 10.48550/arxiv.2112.02627.

[19] D. H. M. de Souza and C. J. Bordin, "Ensemble and Mixed Learning Techniques for Credit Card Fraud Detection," Jan. 01, 2021, Cornell University. doi: 10.48550/arXiv.2112.

[20] J. Nicholls, A. Kuppa, and N. Le-Khac, "Financial Cybercrime: A Comprehensive Survey of Deep Learning Approaches to Tackle the Evolving Financial Crime Landscape," Jan. 01, 2021, Institute of Electrical and Electronics Engineers. doi: 10.1109/access.2021.3134076.

[21] T. T. Nguyen, H. Tahir, M. Abdelrazek, and M. A. Babar, "Deep Learning Methods for Credit Card Fraud Detection," Jan. 01, 2020, Cornell University. doi: 10.48550/arxiv.2012.03754.

[22] Rawat, Arvind & Tiwari, Sandeep. (2023). A comprehensive review on credit card fraud detection using machine learning techniques. International Journal of Innovative Research and Growth. 12. 10.26671/IJIRG.2023.2.12.103.

[23] Y. Xia, C. Liu, Y. Li, and N. Liu, "A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring," Feb. 10, 2017, Elsevier BV. doi: 10.1016/j.eswa.2017.02.017.

[24] E. Tuv, A. Borisov, G. C. Runger, and K. Torkkola, "Feature Selection with Ensembles, Artificial Variables, and Redundancy Elimination," Dec. 01, 2009, The MIT Press. Accessed: Dec. 2024. [Online]. Available: http://jmlr.csail.mit.edu/papers/volume10/tuv09a/tuv09a.pdf

[25] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," May 13, 2019, Nature Portfolio. doi: 10.1038/s42256-019-0048-x.

[26] Phua, Clifton & Lee, Vincent & Smith-Miles, Kate & Gayler, Ross. (2010).

ISSN: 2961-5992

A Comprehensive Survey of Data Mining-based Fraud Detection Research. CoRR. abs/1009.6119.

[27] G. K. Kulatilleke and S. Samarakoon, "Empirical study of Machine Learning Classifier Evaluation Metrics behavior in Massively Imbalanced and Noisy data," Jan. 01, 2022, Cornell University. doi: 10.48550/arXiv.2208.

[28] A. Singh, R. K. Ranjan, and A. Tiwari, "Credit Card Fraud Detection under Extreme Imbalanced Data: A Comparative Study of Data-level Algorithms," Apr. 03, 2021, Taylor & Francis. doi: 10.1080/0952813x.2021.1907795.

[29] A. Kumar and R. K. Dwivedi, "Fraud Detection in Credit Card Data using Unsupervised Machine Learning Based Scheme," Jul. 01, 2020. doi: 10.1109/icesc48915.2020.9155615.

[30] Grover, Prince & Li, Zheng & Liu, Jianbo & Zablocki, Jakub & Zhou, Hao & Xu, Julia & Cheng, Anqi. (2022). FDB: Fraud Dataset Benchmark. 10.48550/arXiv.2208.14417.

[31] K. Kerwin and N. D. Bastian, "Stacked Generalizations in Imbalanced Fraud Data Sets using Resampling Methods," Jan. 01, 2020, Cornell University. doi: 10.48550/arxiv.2004.01764.

[32] S. Makki, Z. Assaghir, Y. Taher, R. Haque, M.-S. Hacid, and H. Zeineddine, "An Experimental Study With Imbalanced Classification Approaches for Credit Card Fraud Detection," Jan. 01, 2019, Institute of Electrical and Electronics Engineers. doi: 10.1109/access.2019.2927266.

[33] S. Vijayaraghavan, T. Guan, J. Jason, and Song, "GAN based Data Augmentation to Resolve Class Imbalance," Jan. 01, 2022, Cornell University. doi: 10.48550/arxiv.2206.05840.

[34] G. Velarde, A. Sudhir, S. Deshmane, A. Deshmunkh, K. Sharma, and V. Joshi, "Evaluating XGBoost for Balanced and Imbalanced Data: Application to Fraud Detection," Jan. 01, 2023, Cornell University. doi: 10.48550/arxiv.2303.15218.

[35] M. Buda, A. Maki, and M. A. Mazurowski, "A systematic study of the class imbalance problem in convolutional neural networks," Jul. 29, 2018, Elsevier BV. doi: 10.1016/j.neunet.2018.07.011.

[36] B. Quinn, "Explaining AI in Finance: Past, Present, Prospects," Jan. 01, 2023, Cornell University. doi: 10.48550/arXiv.2306.

[37] W. J. Yeo, W. van der Heever, R. Mao, E. Cambria, R. Satapathy, and G. Mengaldo, "A Comprehensive Review on Financial Explainable AI," arXiv (Cornell University). Cornell University, Jan. 01, 2023. doi: 10.48550/arxiv.2309.11960.

[38] V. B. Nguyen, K. G. Dastidar, M. Granitzer, and W. Siblini, "The Importance of Future Information in Credit Card Fraud Detection," Jan. 01, 2022, Cornell University. doi: 10.48550/arxiv.2204.05265.

[39] V. B. Nguyen, K. G. Dastidar, M. Granitzer, and W. Siblini, "The Importance of Future Information in Credit Card Fraud Detection," Jan. 01, 2022, Cornell University. doi: 10.48550/arXiv.2204.

[40] J. Hasan, "Security and Privacy Issues of Federated Learning," Jan. 01, 2023, Cornell University. doi: 10.48550/arxiv.2307.12181.

[41] N. Bouacida and P. Mohapatra, "Vulnerabilities in Federated Learning," Jan. 01, 2021, Institute of Electrical and Electronics Engineers. doi: 10.1109/access.2021.3075203.

[42] B. Mytnyk, O. Tkachyk, N. Shakhovska, C. Федушко, and Y. Syerov, "Application of Artificial Intelligence for Fraudulent Banking Operations Recognition," May 10, 2023, Multidisciplinary Digital Publishing Institute. doi: 10.3390/bdcc7020093.

[43] J. Sen, R. Sen, and A. Dutta, "Machine Learning in Finance-Emerging Trends and Challenges," Jan. 01, 2021, Cornell University. doi: 10.48550/arXiv.2110.

[44] Y. Vivek, V. Ravi, A. A. Mane, and L. R. Naidu, "Explainable Artificial Intelligence and Causal Inference based ATM Fraud Detection," Jan. 01, 2022, Cornell University. doi: 10.48550/arxiv.2211.10595.

[45] T. Awosika, R. M. Shukla, and B. Pranggono, "Transparency and Privacy: The Role of Explainable AI and Federated Learning in Financial Fraud Detection," Jan. 01, 2024, Institute of Electrical and Electronics Engineers. doi: 10.1109/access.2024.3394528.

[46] W. Saeed and C. W. Omlin, "Explainable AI (XAI): A Systematic Meta-Survey of Current Challenges and Future Opportunities," Jan. 01, 2021, Cornell University. doi: 10.48550/arXiv.2111.

[47] N. B. Truong, K. Sun, S. Wang, F. Guitton, and Y. Guo, "Privacy preservation in federated learning: An insightful survey from the GDPR perspective," Jul. 17, 2021, Elsevier BV. doi: 10.1016/j.cose.2021.102402.

[48] Z. Li et al., "APPFLx: Providing Privacy-Preserving Cross-Silo Federated Learning as a Service," Jan. 01, 2023, Cornell University. doi: 10.48550/arXiv.2308.

[49] J. Wen, Z. Zhang, Y. Lan, Z. Cui, J. Cai, and W. Zhang, "A survey on federated learning: challenges and applications," Nov. 11, 2022, Springer Science+Business Media. doi: 10.1007/s13042-022-01647-y.

ISSN: 2961-5992