International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

# Credit Card Fraud Detection Using Machine Learning

**Jitendra Kumar[1], Pankaj Kumar Goswami[2]**

[1]Student (M.Tech., ML & DS), Teerthankar Mahaveer University
[2]Associate Professor, FoE, Teerthankar Mahaveer University

**ABSTRACT**

Unprecedented advancement of e-commerce soars the frequency of online and offline financial transactions of Credit Card as a popular means of payment for public. With the tremendous frequency of transactions per minute worldwide, the multi-fold risk of fraudulent transaction has increased significantly for both the parties either user or issuer.

This paper presents the comprehensive survey on multiple machine learning approaches to credit card fraud detection (CCFD). The existing approaches are eliciting good responses in terms of accuracy but the precocious Deep Learning algorithm (here, Convolutional Neural Network) was deployed in the anticipation of better accuracy.

In this paper, comparative analysis has been carried out among various Machine Learning algorithms. Analytical parameters such as counts of layers, epochs & models have been employed. Outlandish outcome found for various machine learning classifier algorithms such as Random Forest, Support Vector Machine, K-Nearest Neighbor, Gaussian Naïve Bayes, Decision Tree, Logistic Regression, moreover, the dataset was fed to Convolutional Neural Network (CNN). The performance metrics for aforesaid classifiers in accordance with standard criteria was recorded. The best outcome was found with Random Forest Classifier depicting F1-score as 85.71%, Precision as 97.40%, and Accuracy as 99.96%.

## INTRODUCTION

Credit card fraud indicates to the unauthorized use of someone else's credit card information to make purchases or carry out fraudulent transactions. This type of fraud can occur through various means, and perpetrators often aim to exploit vulnerabilities in the credit card system for financial gain.

Transaction involves several stages as shown in figure 1, from the initiation of a purchase to the authorization and settlement of the transaction. Customer Initiates a Purchase then the process begins when a customer decides to make a purchase using a credit card. This can occur at a physical point of sale (POS) terminal or online through an e-commerce website [1]-[4]. As a next step there is merchant submission, for in-person transactions, the merchant (business or service provider) swipes, inserts, or taps the customer's credit card using a card reader at the POS terminal. For online transactions, user enters their CC details on the website's payment page. It leads to authorization request, the merchant's payment system sends an authorization request to the credit card issuer (bank or financial institution that issued the credit card) to confirm whether the user has a balance of sufficient credits and the transaction request is valid. It needs an authorization approval; the credit card issuer reviews the authorization request, checks the customer's credit limit, and assesses the transaction's validity. If approved, the issuer sends an authorization code back to the merchant. It moves to transaction approval at merchant; upon

receiving the authorization code, the merchant's system approves the transaction, and the customer is informed that the purchase has been successful. A record of the transaction details, including the authorization code, is stored by both the merchant and the credit card issuer for future reference. Throughout the day, the merchant accumulates authorized transactions and submits them in a batch to the acquiring bank or payment processor[5]-[8]. This is known as batch processing. The acquiring bank or payment processor forwards the batch of authorized transactions to the credit card network (such as Visa or MasterCard) for clearing. Clearing involves the exchange of transaction data between the acquiring and issuing banks. The issuing bank receives the details of the authorized transactions from the credit card network. If the transaction is approved, the issuing bank transfers funds to the acquiring bank.

| 1. Initiation of Transaction | 2. Merchant Submission | 3. Authorization Request | 4. Authorization Approval | 5. Authorization Response |
|---|---|---|---|---|
| Customer provides credit card information to the merchant to make a purchase. | Merchant submits transaction details to its acquiring bank (merchant bank) for authorization. | Acquiring bank forwards authorization request to the credit card issuer. | Credit card issuer evaluates the request based on credit limit, history, and fraud indicators, sending an authorization code if | Acquiring bank receives the authorization response and communicates it to the merchant. |
| **6. Transaction Settlement** | **7. Clearing** | **8. Issuing Bank Settlement** | **9. Funds Transfer** | **10. Merchant Receives Payment** |
| Merchant captures transaction details, and the acquiring bank initiates the process of settling funds. | Acquiring bank sends transaction details to the credit card network (Visa, Mastercard, etc.) for clearing. | Credit card network forwards transaction details to the issuing bank for settlement. | Issuing bank transfers funds to the acquiring bank for settlement. | Acquiring bank deposits funds into the merchant's account, completing the credit card payment process. |

**Figure 1: Credit card process of financial transaction**

This assures that the merchant gets paid for the services or goods provided to the customer. It is to emphasize that the credit card transaction process involves multiple parties, including the cardholder, merchant, acquiring bank, credit card network, and issuing bank. Additionally, security measures such as encryption and tokenization are implemented to protect sensitive cardholder information during the transaction process. Here, in process there lies a high vulnerability of credit card frauds.
Figure 2 shows few key areas of credit card fraud.


**Figure 2: Major domains of credit card fraud**

Criminals use skimming devices to capture data from the magnetic stripe on credit cards. These are often implanted on ATMs, gas pumps, or point-of-sale terminals.

Fraudsters broadcast deceitful emails, messages, or websites that appear legitimate, tricking individuals into provisioning their credit card information.

Cybercriminals gain unauthorized accessibility to a user's account, often through phishing or hacking, and use the victim's credit card for fraudulent transactions.

Criminals may use lost or stolen credit cards to make illegitimate purchases before the cardholder can report the loss. Criminals use pilfered credit card details to make small online purchases to test whether the card is still active before making larger transactions. Fraudsters file applications for credit cards using stolen or fabricated personal information to open new accounts.

Similar to card skimming, criminals use software to generate fake credit card numbers, which may be used for online transactions.

Table 1 shows the various aspects associated with the credit card fraud and evaluating impacts on all three parties card holder, issuing bank and point of sale. These become very severe and vulnerable over the level of finance associated with the fraud.

**Table 1: Descriptive investigation of credit card fraud**

| ASPECTS | DESCRIPTION |
|---|---|
| Methods of Fraud | Skimming, phishing, lost or stolen cards, data breaches, carding, identity theft, account takeover, etc. |
| Stolen Information | Criminals use skimmers, phishing emails, or deceptive methods to obtain credit card details. |
| Lost or Stolen Cards | Physical cards that are lost or stolen can be used for unauthorized purchases until reported. |
| Data Breaches | Large-scale breaches expose credit card information, making it vulnerable to hackers. |
| Carding | Validation of stolen credit card details through small online purchases before larger transactions. |
| Identity Theft | Stolen personal information is used to generate new account of credit card in the victim's name. |
| Account Takeover | Criminals gain access to existing accounts, change information, and make unauthorized transactions. |
| Fraudulent Applications | Submission of fake credit card applications using stolen or fabricated personal information. |
| Unauthorized Transactions | Illegitimate use of credit card information for purchases, cash advances, or transfers without consent. |
| Detection and Prevention | Financial institutions use fraud detection algorithms, and cardholders are advised to monitor statements. |
| Chargebacks | Cardholders can dispute unauthorized transactions, leading to a reversal of the transaction and a refund. |
| Secure Transactions | Adoption of technologies like EMV cards and tokenization enhances transaction security. |

| | |
|---|---|
| Legal Consequences | Credit card fraud is a criminal offense, and perpetrators can face fines and imprisonment if caught. |
| Protection Measures | Regular statement monitoring, prompt reporting of lost/stolen cards, use of secure passwords, cautious sharing of personal information. |

Cybercriminals gain unauthorized access to a user's account, often through phishing or hacking, and use the victim's credit card for fraudulent transactions. Lost or Criminals may use lost or stolen credit cards to make unauthorized purchases before the cardholder can report the loss. Criminals use stolen credit card information to make small online purchases to test whether the card is still active before making larger transactions. Fraudsters place the applications for credit cards using stolen or fabricated personal information to open new accounts. Similar to card skimming, criminals may install skimming devices on point-of-sale terminals to capture credit card data during transactions. Criminals use software to generate fake credit card numbers, which may be utilized for online transactions. Fraudsters manipulate individuals into revealing their credit card information through psychological tactics or impersonation. Fraudsters create a fake online store and use stolen credit card data to purchase goods from a legitimate store, having the goods shipped to the victim's address [15]-[18]. Legitimate cardholders dispute valid transactions with their credit card issuer, often claiming the transaction was unauthorized, resulting in chargebacks. Criminals steal credit card statements or new credit cards from the victim's mailbox. Fraudsters create a fabricated scenario to receive personal data, including credit card details, from the victim. Cybercriminals intercept and alter interaction between the user and a legitimate website, capturing credit card info. Malicious software installed on a user's device can capture such details inputted during online transactions. Figure 3 shows relative analysis of various frauds occurred in credit card financial systems.

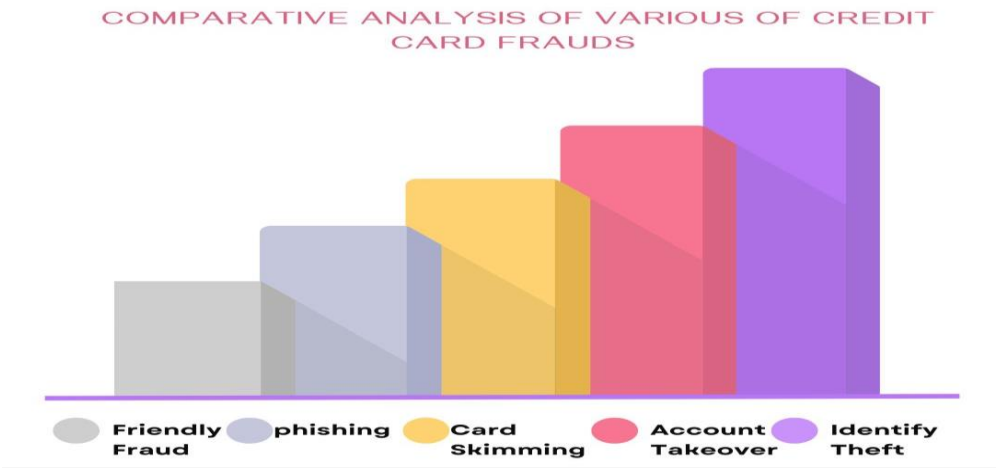


**Figure 3: Relative analysis of various types of fraud detection**

## LITERARY REVIEW

### CREDIT CARD FRAUD DETCTION CHALLENGES

Credit Card fraudulent transaction is unauthorized, unwanted & unlawful usage of Credit Card or account by someone other than the owner of that account. In other words, Credit Card Fraud may be described as a case where a person uses someone else's credit card for personal reasons while the owner and the card issuing authorities are unaware of the fact that the card is being used. A stolen, lost or fake Credit Card could lead to fraud [20]-[24].

**Table 2: challenges pertaining to credit card fraud detection**

| Challenge | Description |
|---|---|
| Sophisticated Techniques | Fraudsters employ advanced methods, including AI, machine learning, and advanced malware, to mimic legitimate transactions and avoid detection. |
| False Positives | Balancing accurate fraud detection with minimizing false positives to avoid inconveniencing legitimate cardholders and impacting user experience. |
| Imbalanced Data | Dealing with imbalanced datasets where the count of legitimate transactions far exceeds fraudulent ones, affecting the criterion of machine learning models. |
| Emerging Fraud Patterns | Staying ahead of fraudsters who continually innovate and adapt their tactics, challenging traditional rule-based systems to keep up with evolving fraud patterns. |
| Cross-Channel Fraud | Detecting fraud across diverse channels such as online, mobile, and in-person transactions, and integrating data from different sources. |
| Synthetic Identities | Identifying and preventing fraud involving synthetic identities created by combining real and fake information in credit card applications. |
| Insider Threats | Mitigating the risk of collusion or insider involvement in fraudulent activities, as insiders may have access to sensitive information. |
| Global Transactions | Monitoring and identifying fraudulent activities across borders, dealing with variations in regulations, transaction patterns, and data sources in different regions. |
| Data Privacy Concerns | Balancing effective fraud detection with user privacy considerations, implementing robust security measures, and complying with data protection regulations. |
| Dynamic Fraud Schemes | Adapting to fraudsters' quick changes in tactics and keeping in sync with technological advancements, requiring constant innovation in fraud detection methods and technologies. |
| Technological Challenges | Overcoming challenges associated with implementing and integrating new technologies, especially for organizations with legacy systems. |

As world is going toward cashless economy, usage of Credit or Debit Card in e-shopping is increasing day by day, and so associated frauds causing huge monetary loss. Credit card fraud can be divided into two main types: application fraud and behavioural fraud. Both types involve various tactics and techniques used by fraudsters to exploit vulnerabilities in the credit card system. Table 2 shows major challenges associated with the credit card fraud detection [25]-[30].

### A. Aspplication Fraud:

Identity Theft: Fraudsters may use stolen or fabricated personal information to file applcation for credit cards in someone else's name. Synthetic Identity Fraud; criminals create fictional identities by combining real and fake information to apply for credit. Over time, they build up the creditworthiness of these synthetic identities before exploiting them. Collusion; this involves an individual or group working with an insider, such as a corrupt bank employee, to submit fraudulent credit card applications.

Document forgery: Criminals may forge documents, such as pay stubs or utility bills, to support their credit card applications.

**B. Behavioural Fraud:**

Stolen Card Fraud: Criminals gain access to credit card information and use it for unauthorized transactions. Account Takeover; fraudsters gain control of a legitimate cardholder's account through various means, such as phishing or hacking, and make unauthorized transactions. Card Not Present (CNP) Fraud; this occurs when the physical card is not required for a transaction, such as online or over-the-phone purchases. Fraudsters may use stolen card details for these transactions. Skimming; criminals use devices called skimmers to collect credit card information from the magnetic stripe when a card is swiped at an ATM or point-of-sale terminal. Phishing and Social Engineering; fraudsters use deceptive emails, messages, or phone calls to trick individuals into providing their credit card information.

**C. Preventive Measures**:

Verification and Authentication: Implement robust identity verification processes during the application phase. This can include document verification, biometric authentication, and other advanced identity verification methods. Machine Learning and AI: Employ advanced technologies to analyse patterns of behaviour and to detect anomalies that may point out fraudulent activity. Encryption and Tokenization: Protect cardholder data by encrypting sensitive information and replacing it with tokens, making it harder for fraudsters to gain access to valuable data. Multi-Factor Authentication: Implement multi-factor authentication measures for online transactions to add an extra layer of security. Credit card issuers, financial institutions, and consumers all play a role in preventing and mitigating credit card fraud by staying vigilant and adopting best practices in security and fraud detection.

## CONVENTIONAL METHODS OF CREDIT CARD FRAUD DETECTION

Conventional methods of credit card fraud detection typically involve rule-based systems, heuristics, and predefined patterns to identify potentially fraudulent transactions. While these methods have been effective to some extent, they may struggle to keep pace with the evolving tactics of fraudsters. Figure 2 shows factors affecting the process effectiveness of CCFD. [31]-[36].

**A. Rule-Based Systems:**

Description: Rule-based systems employ predefined rules and conditions to flag transactions that match specific patterns associated with known fraud.

**Pros:** Simple to implement, easy to understand, and can quickly identify known fraud patterns.

**Cons:** Limited adaptability to new and emerging fraud schemes; may generate false positives or negatives.

**B. Transaction Monitoring:**

Description: Transaction-monitoring for uncommon patterns or deviations from a cardholder's normal spending behavior.

**Pros:** Can identify anomalies based on historical transaction data.

**Cons:** May trigger false alarms for legitimate but uncommon transactions; may not detect sophisticated fraud patterns.

**C. Address Verification System (AVS):**

Description: Verifying the billing address provided during a transaction against the one on file with the card issuer.

**Pros:** Adds an additional layer of verification.

**Cons:** Limited effectiveness in cases where billing information has been compromised; doesn't address online transactions without address verification.

**D. Card Verification Code (CVC) Check:**

Description: Verifying the three-digit code on the back of the credit card during transactions.

**Pros:** Adds an extra layer of security for online and card-not-present transactions.

**Cons:** Doesn't prevent fraud if the card information has been compromised; limited effectiveness in some scenarios.

**E. Velocity Checks:**

Description: Monitoring the frequency and volume of transactions within a specified timeframe.

**Pros:** Can identify unusual patterns such as a sudden increase in transaction frequency.

**Cons:** May initiate false positives for legitimate high-frequency transactions.

**F. Geolocation Checks:**

Description: Verifying the location of the transaction against the cardholder's usual locations.

**Pros:** Adds a layer of verification based on geographical information.

**Cons:** Limited effectiveness for individuals who frequently travel; may not detect remote or online transactions.

**G. Manual Review and Investigation:**

Description: Human review of flagged transactions for further investigation based on suspicion or predefined criteria.

**Pros:** Allows for nuanced judgment and investigation of complex cases.

**Cons:** Resource-intensive and time-consuming; may delay transaction approvals.

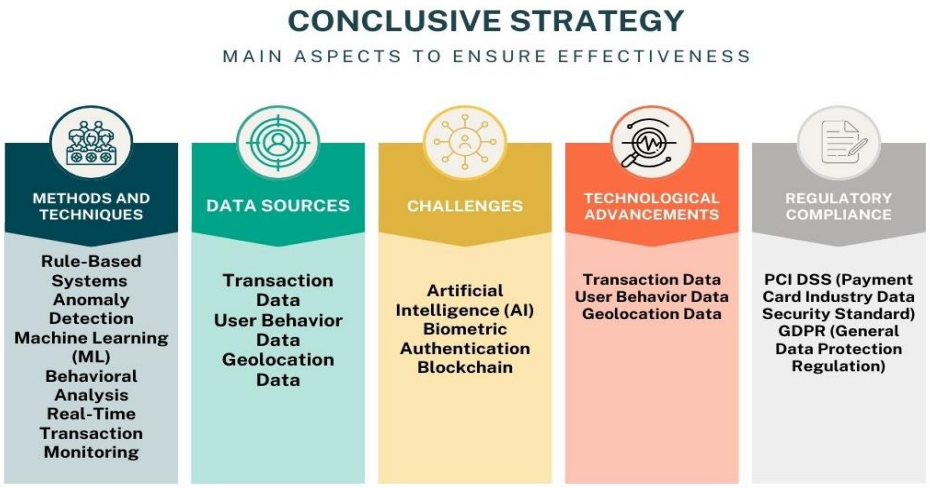


**Figure 4: Factors affecting the process effectiveness of credit card fraud detection**

While these conventional methods provide a baseline for credit card fraud detection, the industry is increasingly turning to advanced technologies, such as machine learning, artificial intelligence, and behavioural analytics, to enhance the efficiency and accuracy of fraud detection systems. These technologies can get adapted to evolving fraud patterns and deliver more accurate and timely identification of suspicious activities as depicted in figure 4. Further an important aspect is the considering main features to determine the fraud and its frequency. The threat may appear through any feature from the process of CC transaction as deliberated in table 3 [36]-[38].

**Table 3: Feature affecting credit card fraud detection**

| S.No. | Name of Feature | Description |
|---|---|---|
| 1 | Account Number | Related with account number |
| 2 | Open to buy | The availability of balance |
| 3 | Credit Limit | The maximum amount of credit of the associated account |
| 4 | Card Number | Number of Credit Card |
| 5 | Transaction Amount | The transaction amount submitted by the merchant |
| 6 | Transaction Time | Time of the transaction |
| 7 | Transaction Date | Date of the transaction |
| 8 | Transaction Type | Types of transaction, such as a cash withdrawal and purchase |
| 9 | Currency Code | The currency code |
| 10 | Merchant Category Code | The merchant business type code |
| 11 | Merchant Number | The merchant reference number |
| 12 | Transaction Country | The country where the transaction takes place |
| 13 | Transaction City | The city where the transaction takes place |
| 14 | Approval Code | The response to the authorization request, it means approve or reject |

## ADVANCED METHODS OF CREDIT CARD FRAUD DETECTION

As per record, there were 393207 instances of such frauds out of 1.4 million identity theft reported ones. Therefore, an Automated CCFD comes as a rescue and important tool for financial institutions. The CCFD using ML model is a supervised (classification) model to recognize fraudulent and non-fraudulent transactions. The ML model deals with various underlying problems such as system-reaction-time, cost-sensitivity, pre-processing of features (PCA etc.) to predict as per prior data patterns. Supervised ML algorithm like SVM is employed for solving linear and non-linear classification problem of dataset like Image Recognition, Credit Rating etc. SVM creates Hyper plane to separate input data in support vector. Although, several Deep Learning algorithms are available namely having application in computer vision, NLP, Heart disease detection, healthcare fraud detection, malware detection, intrusion detection, video surveillance detection, location tracking.. In this paper, we are using CNN (Convolutional Neural Networks) for identity theft pertaining to CCFD, whether the transaction is normal or fraudulent one. The process of CCFD as follows: Feature Selection algorithms are rendered to dataset to order the principal-features helping to make predictions based on classification. Feature-Extraction algorithms, using Deep Learning model, are deployed to extract/generate new features and solve clustering problem from the dataset of CCFD. The performance of CNN model is analyzed by adding layers. The comparative-assessment between ML & DL is carried out. The outcome shows that CNN model proposed outnumbers the ML model. To examine the accuracy of classifier in ML and clustering in DL, the Model Performance Evaluation benchmarks such as F1-scor, Precision, AUC curve and Accuracy are employed. The most recent dataset are used to carry out the experiments. The common procedure behind the process is shown in figure 5 [39]-[40].

a) Data preprocessing: Clean and preprocess the collected data. This involves eliminating duplicates, treating missing values, and normalizing or scaling the numerical features. It is also crucial to balance the dataset to ensure equal representation of both fraud and non-fraud cases.

b) Feature engineering: Extract relevant features from the dataset that can help in recognizing fraudulent and legitimate transactions. This may involve creating new attributes based on domain knowledge or using techniques such as dimensionality reduction.
c) Model selection: Choose an appropriate machine learning algorithm for fraud detection. Commonly used algorithms include logistic regression, decision trees, random forests, and neural networks. Consider the trade-offs between accuracy, interpretability, and computational complexity when selecting the model. Model training: Split the preprocessed dataset into training and testing sets. Train the selected machine learning model on the training set, using the features and the corresponding labels (fraud or non-fraud). Model evaluation: Evaluate the trained model on the testing set to measure its performance. Common evaluation metrics include accuracy, precision, recall, and F1 score. Adjust the model parameters or try different algorithms if the performance is not satisfactory.
d) Model deployment: If the model performance is satisfactory, deploy it into production. This involves integrating the model into the credit card payment system and continuously monitoring incoming transactions for potential fraud.
e) Model maintenance: Periodically retrain and update the model using new data to adapt to changing fraud patterns. Monitor the performance of model and make necessary adjustments to ensure its effectiveness.It's important to note that the above steps provide a general framework for credit card fraud detection using machine learning. The specific implementation details may vary depending on the dataset, algorithms, and tools used.

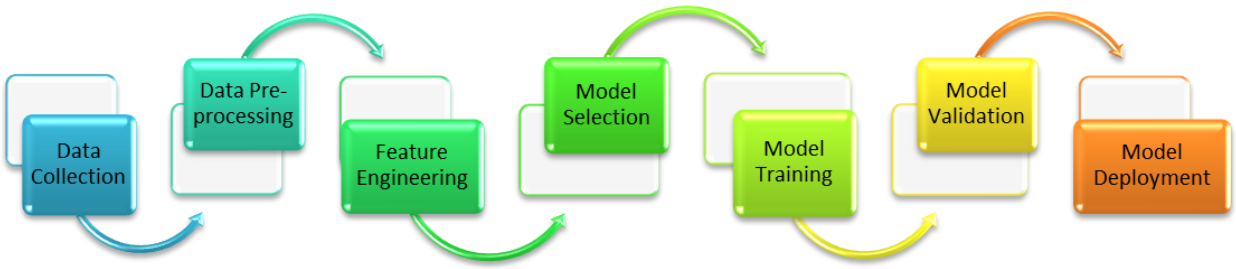


**Figure 5: Process flow of credit card fraud detection using machine learning**

As criminals continually find new methods to exploit weaknesses in the system, financial institutions and businesses are turning to machine learning algorithms as a solution. Few common methodologies are shown in figure 6.

Data preprocessing plays a crucial role in credit card fraud detection. It involves cleaning and transforming raw data to make it appropriate for analysis. Techniques for instance outlier removal, feature scaling, and dimensionality reduction are essential in improving the accuracy of model used for fraud detection. By identifying and removing outliers, scaling features to a common range, and reducing the dimensionality of the data, the models can better identify patterns and anomalies associated with fraudulent transactions.

Supervised learning algorithms are extensively employed in detection of fraud credit card transactions. These algorithms learn from labeled data to infer predictions on new, unseen data. Logistic regression, decision trees, random forests, and support vector machines are some prominent supervised learning algorithms applied in this context. By leveraging historical patterns and features, these algorithms

effectively classify transactions as either fraudulent or legitimate. Unsupervised Learning Techniques: Unsupervised learning algorithms come into play when labeled fraud data is limited or unavailable. These algorithms target to identify anomalous patterns in the supplied data without any previous knowledge of fraud instances. Clustering algorithms like k-means and DBSCAN besides outlier detection algorithms like Isolation Forest and Local Outlier Factor, are commonly employed in unsupervised fraud detection. By detecting patterns that deviate from the norm, these algorithms can discern potential fraudulent activities.

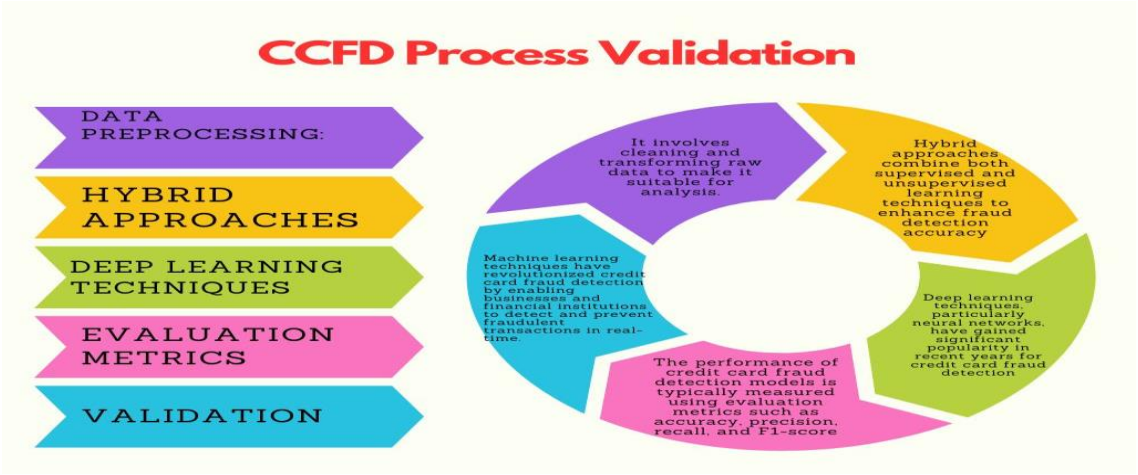


**Figure 6: Relative analysis of various types of fraud detections**

## RESEARCH GAP

**Table 4: Research Gap (Cons) of enhanced machine learning methods used for CCFD**

| Technique | Type | Key Features | Pros | Cons |
|---|---|---|---|---|
| Rule-Based Systems | Rule-Based | Set predefined rules to identify suspicious transactions | Simple, interpretable | Limited adaptability, may not capture new patterns |
| Anomaly Detection | Statistical | Identify deviations from normal behaviour | Detects novel fraud patterns | May initiate false positives for legitimate behaviour |
| Machine Learning | Supervised/ | Learn patterns from labeled data to predict fraud | Adaptable, effective with large datasets | Requires labelled training data |
| | Unsupervised | Identify patterns without labeled data | Effective for detecting unknown fraud patterns | May have greater false positive rates |
| Neural Networks | Deep Learning | Learn complex patterns in data using neural networks | Can handle intricate relationships in data | Requires substantial computational resources |

| | | | | |
|---|---|---|---|---|
| Ensemble Methods | Ensemble | Combine multiple models for improved performance | Robust, reduces over fitting | Increased complexity and computational cost |
| Isolation Forest | Anomaly Detection | Isolate anomalies using tree-based structures | Efficient with high-dimensional data, resistant to noise | May struggle with certain types of data distributions |
| One-Class SVM | Support Vector | Train on non-fraudulent instances only | Effective for novelty detection, suitable for imbalanced data | May struggle with highly complex data |
| Auto-encoders | Neural Network | Unsupervised learning for feature representation | Effective for detecting anomalies, learns data patterns | Requires careful tuning and may be computationally intensive |
| Feature Engineering | Pre-processing | Create new features or transform existing ones | Improves model performance by capturing relevant information | Requires domain expertise and understanding of data |

**PROPOSED METHODOLOGY**

Hybrid approaches combine supervised and unsupervised learning techniques to enhance fraud detection accuracy. By leveraging the strengths of both methods, these methodologies can generate more robust and accurate results. One example is the semi-supervised learning approach, where a small portion of labeled fraud data is used to train a supervised learning model. This model is then utilized to detect fraud in the unlabeled data. By combining the power of labeled data with the ability to identify anomalies in unlabeled data, hybrid approaches can improve the overall effectiveness of fraud detection systems. Deep learning techniques, especially neural networks have gained significant popularity in recent years for credit card fraud detection. These models can automatically learn complex patterns and relationships in the data, making them extremely effective in detecting fraudulent transactions. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have shown promising results in fraud detection tasks. By leveraging the hierarchical structure of neural networks, these models can capture intricate patterns and identify subtle indicators of fraud. The performance of credit card fraud detection models is typically measured using evaluation metrics such as accuracy, precision, recall, and F1-score. These metrics offer insights into the model's ability to accurately classify transactions as fraudulent or legitimate. Additionally, techniques like cross-validation and Receiver Operating Characteristic (ROC) curves help assess the model's generalization and robustness. By evaluating the models using these metrics, businesses and financial institutions can have informed decisions about the effectiveness of their fraud detection systems [41]-[42].

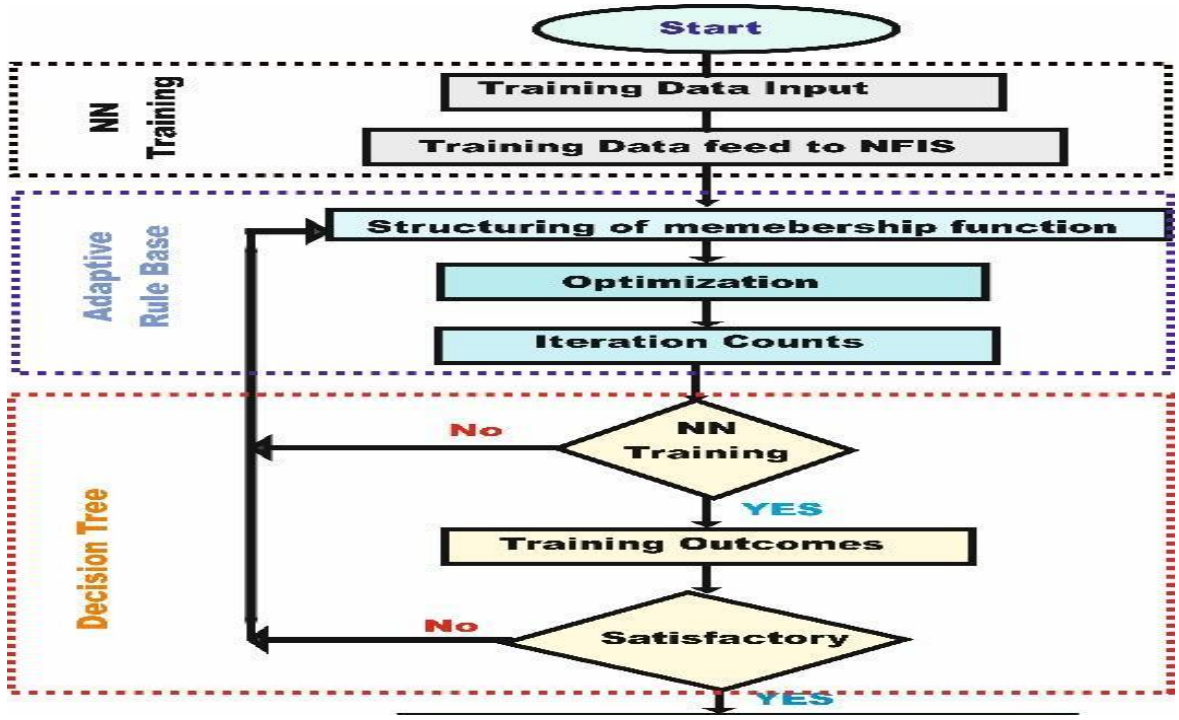


**Figure 7: Advanced soft computing process flow**

The standard operating procedure of implementation of advanced soft computing techniques is shown in figure 7. Advanced methods of credit card fraud detection go beyond traditional rule-based systems and leverage cutting-edge technologies to enhance accuracy and adaptability to evolving fraud tactics. There are numerous Machine Learning Algorithms are in place for Credit Card Fraud Detection (hereinafter CCFD) such as Random Forest, Decision Tree, Extreme Learning Method, SVM, Logistic Regression, & XG Boost. Here, European Card Benchmark dataset is put to use for CCFD. First, dataset is fed to Machine Learning algorithm and accuracy of fraud detection is recorded, then CNN is applied to the dataset and improved accuracy is noted down. Later, hidden layers are added to CNN to refine accuracy. The CCFD approaches and their impacts are shown in figure 8 & 9 respectively [43]-[44].

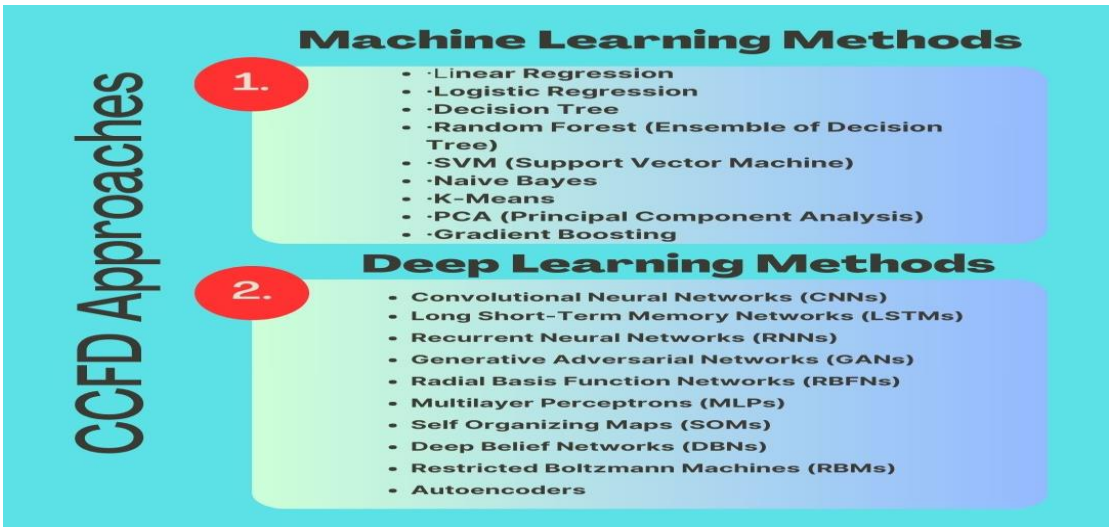


**Figure 8: ML and DL approaches for CCFD**

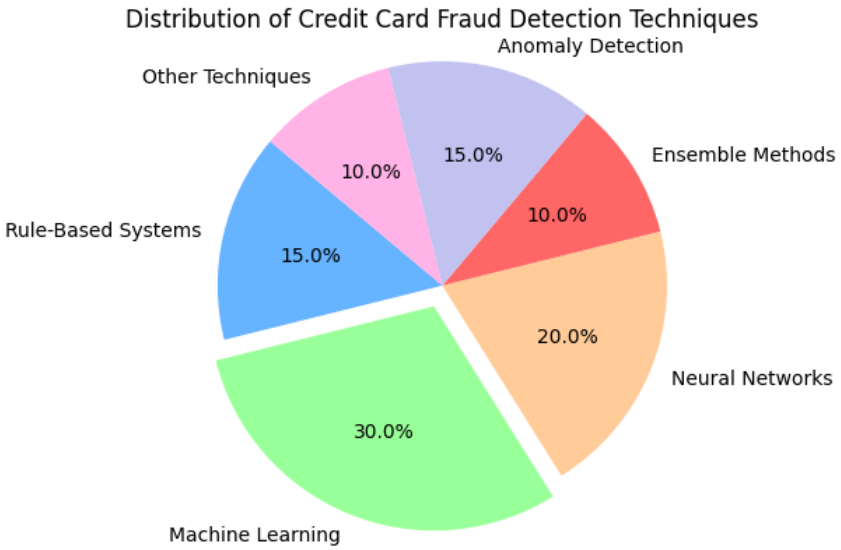


**Figure 9: Various advanced of CCFD techniques**

**Machine Learning and Predictive Analytics:**
Description: Using machine learning algorithms to analyze historical transaction data and learn patterns indicative of fraud. Predictive analytics can then identify potential fraudulent transactions based on these learned patterns.
Pros: Adaptable to new and emerging fraud patterns; can handle large datasets and complex relationships between variables.
Cons: Requires significant data for training; may be vulnerable to adversarial attacks.
**Behavioral Analytics:**
Description: Analyzing the behavioral patterns of cardholders to establish a baseline of normal activity. Deviations from this baseline, such as sudden changes in spending habits, may trigger fraud alerts.
Pros: Adapts to individual cardholder behavior; effective in detecting anomalies.
Cons: May initiate false positives during legitimate changes in spending behavior.
**Biometric Authentication:**
Description: Implementing biometric methods such as fingerprint, voice, or facial recognition for user authentication during transactions.
Pros: Adds a high level of security; difficult for fraudsters to replicate biometric information.
Cons: Implementation costs; potential privacy concerns.
**Device Fingerprinting:**
Description: Analyzing unique characteristics of devices used for transactions, including device type, IP address, and geolocation, to identify anomalies.
Pros: Enhances fraud detection for online transactions; adds an additional layer of verification.
Cons: May be counter effective for users who frequently change devices.
**Real-Time Transaction Monitoring:**
Description: Monitoring transactions in real-time and applying dynamic rules to detect anomalies as they occur.
Pros: Immediate response to suspicious activities; reduces false positives by considering current transaction context.

Cons: Requires robust real-time processing capabilities.

**Network Analysis:**

Description: Examining relationships and connections between entities, including cardholders, merchants, and devices, to identify patterns indicative of fraud networks.

Pros: Detects organized fraud schemes involving multiple entities; helps uncover hidden relationships.

Cons: Complexity in analyzing large networks; may require advanced algorithms.

**Deep Learning:**

Description: Utilizing deep neural networks to automatically learn and extract features from raw data, enabling more complex pattern recognition.

Pros: Powerful for handling unstructured data; can capture intricate relationships.

Cons: Requires substantial computational resources and large amounts of data for training.

**Blockchain Technology:**

Description: Implementing blockchain for secure and transparent transaction verification, reducing the risk of unauthorized modifications or tampering.

Pros: Enhances transaction integrity and security.

Cons: Limited adoption; challenges related to scalability and integration with existing systems.

**Customer Authentication Solutions:**

Description: Deploying advanced customer authentication methods, such as two-factor authentication (2FA) or tokenization, to secure transactions.

Pros: Adds extra layers of security; reduces the risk of unauthorized access.

Cons: Initial implementation costs; potential user inconvenience.

**Continuous Learning Systems:**

Description: Implementing systems that continuously learn and adapt to new fraud patterns, ensuring ongoing effectiveness in fraud detection.

Pros: Adaptable to changing fraud tactics; reduces the need for frequent rule updates.

Cons: Requires ongoing monitoring and refinement; initial implementation complexity.

Combining multiple advanced methods and technologies can create a robust and dynamic credit card fraud detection system that effectively addresses the challenges posed by sophisticated fraudsters. It's important for financial institutions to regularly update and improve their fraud detection systems to stay ahead of emerging threats.

**Table 5 : Summary of advanced methods used for CCFD**

| Advanced Method | Description |
|---|---|
| Machine Learning and Predictive Analytics | With machine learning algorithms to analyze historical transaction data and predict potential fraudulent transactions based on learned patterns. |
| Behavioral Analytics | Analyzing individual cardholder behavior to establish a baseline and detect anomalies, such as sudden changes in spending habits. |
| Biometric Authentication | Implementing biometric methods (fingerprint, voice, facial recognition) for user authentication during transactions. |
| Device Fingerprinting | Analyzing unique characteristics of devices used for transactions (e.g., device type, IP address) to identify anomalies and enhance |

| | |
|---|---|
| | online fraud detection. |
| Real-Time Transaction Monitoring | Monitoring transactions in real-time and applying dynamic rules to detect anomalies as they occur, providing an immediate response to suspicious activities. |
| Network Analysis | Examining relationships and connections between entities (cardholders, merchants, devices) to identify patterns indicative of fraud networks. |
| Deep Learning | Utilizing deep neural networks to automatically learn and extract features from raw data, enabling more complex pattern recognition. |
| Blockchain Technology | Implementing blockchain for secure and transparent transaction verification, reducing the risk of unauthorized modifications or tampering. |
| Customer Authentication Solutions | Deploying advanced customer authentication methods (e.g., two-factor authentication, tokenization) to enhance transaction security. |
| Continuous Learning Systems | Implementing systems that continuously learn and adapt to new fraud patterns, reducing the need for frequent rule updates and staying ahead of emerging threats. |

**Table 6: Enhanced machine learning methods used for CCFD**

| Reference | Methodology and Key Points |
|---|---|
| [8] | Unsupervised feature learning using a stacked sparse autoencoder (SSAE) for fraud prediction. |
| [9] | Neural network ensemble classifier with a hybrid data resampling method. Base learner: Long short-term memory (LSTM) in adaptive boosting (AdaBoost) technique. |
| [7] | Credit card fraud detection using ML algorithms (RF, NB, MLP) and Synthetic Minority Oversampling Technique (SMOTE) for imbalanced data. RF algorithm showed the highest accuracy. |
| [10] | Feature selection to minimize data overlap using algorithms (RONS, ROS, ROA) built through sparse feature selection. Binary classification with good performance. |
| [11] | Combination of oversampling and feature selection methods to improve classification algorithms. Significant performance improvement demonstrated. |
| [7] | Intelligent payment card fraud detection system. Assessment of aggregated features identified by a genetic algorithm for improved fraud detection accuracy. |

| | |
|---|---|
| [12] | Hybrid approach using Recursive Features Elimination (RFE), Hyper-Parameters Optimization (HPO), and SMOTE. Excellent performance across different datasets. |
| [13] | Data-point machine learning with SMOTE-based oversampling. Various classifiers tested, showing increased accuracy for identifying fraudulent transactions. |
| [14] | Feature selection using Enhanced Neural Networks (ENN) and Artificial Bee Colonies (ABCs) for improved accuracy in credit card fraud classifications. Logical relationships explored through LGBPs. |
| [6] | Two-stage approach: selection of optimal ML algorithms (LR, KNN, DT, NB, RF, GBM, Light GBM, XG Boost, Cat Boost) and integration with different resampling techniques. AllKNN-Cat Boost outperformed. |
| [15] | SVM hyper parameter optimization (c and sigma) using Cuckoo Search Algorithm, Genetic Algorithms, and Particle Swarm Optimization. Recommendation for exploring new algorithms in future work. |

**Table 7: CCFD performance analysis**

| S.No. | Datasets | Algorithms | Accuracy (%) | Reference |
|---|---|---|---|---|
| 1 | The bankcard enrolment records | LR-based | 75 | [12] |
| | | RF-based | 73 | |
| | | GBDT-based | 74 | |
| 2 | Commercial banks in China | SVM | 97.10 | [4] |
| | | RF | 96.90 | |
| 3 | Records of credit card transactions | Light Gradient Boosting Machine algorithm | 99.91 | [13] |
| 4 | Cardholders Dataset of Europe | CS-SVM | 98.05 | [14] |
| | | GA-SVM | 98.05 | |
| | | PSO-SVM | 98.05 | |

**Table 8: CCFD data sets analytics**

| S.No. | Datasets | Algorithms | Accuracy (%) | Reference |
|---|---|---|---|---|
| 1 | European Cards Dataset | LSTM | 87.02 | [30] |
| | | GRU | 86.02 | |
| | | Ensemble model as baseline models | 83.37 | |
| 2 | The Brazilian Dataset | LSTM | 88.47 | [30] |
| | | GRU | 84.13 | |
| | | Ensemble model as baseline models | 79.05 | |
| 3 | Commercial banks in China | Deep belief networks (DBN) | 97.02 | [15] |
| | | CNN | 97.24 | |

| | | | | |
|---|---|---|---|---|
| | | RNN | 97.25 | |
| 4 | Cardholders Dataset of Europe | GAN | 99.95 | [31] |
| | | VAE | 99.96 | |

**Table 9: CCFD algorithms accuracy [30]**

| S.No. | Algorithm Name | Accuracy (%) | F1-score (%) |
|---|---|---|---|
| 1 | Decision Tree | 99.93 | 81.05 |
| 2 | KNN | 99.95 | 85.71 |
| 3 | Logistic Regression | 99.91 | 73.56 |
| 4 | SVM | 99.93 | 77.71 |
| 5 | Random Forest | 99.92 | 77.27 |
| 6 | XG Boost | 99.94 | 84.49 |

## RESULTS

| S. No. | Machine Learning Classifier Name | Accuracy % | F1-score | Precision | Recall |
|---|---|---|---|---|---|
| 1 | Random Forest | 99.96 | 0.8571 | .9740 | .7653 |
| 2 | SVC | 99.83 | 0.0 | 0.0 | 0.0 |
| 3 | KNN | 99.84 | 0.0971 | 1.000 | .0510 |
| 4 | Decision Tree | 99.92 | 0.7843 | .7547 | .8163 |
| 5 | Logistic Regression | 99.87 | 0.5957 | .62222 | .5714 |
| 6 | Gaussian Naive Bayes | 99.30 | 0.2375 | .1462 | .6327 |
| 7 | AdaBoost | 99.93 | 0.7845 | .8554 | .7245 |
| 8 | XGBoost | 99.96 | 0.8587 | .9620 | .7755 |
| 9 | CNN (epochs=12) [optimizer= adam] | 99.94 | 0.8214 | .8500 | .7947 |
| 10 | CNN (epochs=18) [optimizer= adam] | 99.94 | 0.8141 | .8581 | .7744 |
| 11 | CNN (epochs=20) [optimizer= adam] | 99.94 | 0.8157 | .8518 | .7825 |
| 12 | CNN (epochs=35) [optimizer= adam] | 99.94 | 0.8180 | .8643 | .7764 |
| 13 | CNN (epochs=40) [optimizer= adam] | 99.94 | 0.8100 | .8750 | .7541 |
| 14 | CNN (epochs=12) [optimizer= sgd] | 99.94 | 0.8165 | .8285 | .8049 |
| 15 | CNN (epochs=18) [optimizer= sgd] | 99.94 | 0.8187 | 0.8559 | .7846 |
| 16 | CNN (epochs=20) [optimizer= sgd] | 99.94 | 0.8290 | .8568 | .8028 |
| 17 | CNN (epochs=35) [optimizer= sgd] | 99.94 | 0.8296 | .8653 | .7967 |
| 18 | CNN (epochs=40) [optimizer= sgd] | 99.94 | 0.8170 | .8316 | .8028 |

## CONCLUSION

Machine learning techniques have revolutionized credit card fraud detection by enabling businesses and financial institutions to detect and prevent fraudulent transactions in real-time. This has provided a

comprehensive review of various machine learning approaches, including supervised learning, unsupervised learning, hybrid approaches, and deep learning techniques. It is crucial to consider the strengths and limitations of each technique and choose the most appropriate approach based on the available data and resources. With ongoing advancements in machine learning, credit card fraud detection will continue to evolve, making it increasingly challenging for fraudsters to exploit the system. By staying up-to-date with the latest techniques and continuously improving fraud detection systems, businesses and financial institutions can effectively combat credit card fraud and protect their customers' financial security. Credit Card fraud is without a doubt an act of criminal dishonesty. It is a financial threat to both credit card issuing companies & its holders. Out of millions of transactions in a fraction of time, a robust classifier is required to distinguish between fraud transactions and non-fraud ones. Overall, credit card fraud detection is an ongoing challenge that requires continuous research and innovation. By leveraging the power of machine learning and privacy-preserving techniques, we can enhance the security of financial transactions and protect customers from fraudulent activities.

**REFERENCES**

1. Singh Y, Hussain I, Mishra S, Singh B, “Adaptive neuron detection-based control of single-phase SPV grid integrated system with active filtering”, IET Power Electron.,Vol. 10 Iss. 6, pp. 657-666, 2017.
2. Y. Abakarim, M. Lahby, and A. Attioui, ``An efficient real time model for credit card fraud detection based on deep learning," in Proc. 12th Int. Conf. Intell. Systems: Theories Appl., Oct. 2018, pp. 1_7, doi: 10.1145/3289402.3289530.
3. H. Abdi and L. J. Williams, ``Principal component analysis," Wiley Inter-discipl. Rev., Comput. Statist., vol. 2, no. 4, pp. 433_459, Jul. 2010, doi:10.1002/wics.101.
4. V. Arora, R. S. Leekha, K. Lee, and A. Kataria, ``facilitating user authorization from imbalanced data logs of credit cards using artificial intelligence," Mobile Inf. Syst., vol. 2020, pp. 1_13, Oct. 2020, doi: 10.1155/2020/8885269.
5. A. O. Balogun, S. Basri, S. J. Abdulkadir, and A. S. Hashim, ``Performance analysis of feature selection methods in software defect prediction: A search method approach," Appl. Sci., vol. 9, no. 13, p. 2764, Jul. 2019,doi: 10.3390/app9132764.
6. B. Bandaranayake, ``Fraud and corruption control at education system level: A case study of the Victorian department of education and early childhood development in Australia," J. Cases Educ. Leadership, vol. 17,
7. no. 4, pp. 34_53, Dec. 2014, doi: 10.1177/1555458914549669.
8. J. Baszczyski, A. T. de Almeida Filho, A. Matuszyk, M. Szelg_, and R. Sowiski, ``Auto loan fraud detection using dominance-based rough set approach versus machine learning methods," Expert Syst. Appl., vol. 163,
9. Jan. 2021, Art. no. 113740, doi: 10.1016/j.eswa.2020.113740.
10. B. Branco, P. Abreu, A. S. Gomes, M. S. C. Almeida, J. T. Ascensão, and P. Bizarro, ``Interleaved sequence RNNs for fraud detection," in Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2020, pp. 3101_3109, doi: 10.1145/3394486.3403361.
11. F. Cartella, O. Anunciacao, Y. Funabiki, D. Yamaguchi, T. Akishita, and O. Elshocht, ``Adversarial attacks for tabular data: Application to fraud detection and imbalanced data," 2021, arXiv:2101.08030.

12. S. S. Lad, I. Dept. of CSE Rajarambapu Institute of Technology Rajaramnagar Sangli Maharashtra, and A. C. Adamuthe, ``Malware classi_cation with improved convolutional neural network model," Int.J. Comput. Netw. Inf. Secur., vol. 12, no. 6, pp. 30_43, Dec. 2021,doi: 10.5815/ijcnis.2020.06.03.
13. V. N. Dornadula and S. Geetha, ``Credit card fraud detection using machine learning algorithms," Proc. Comput. Sci., vol. 165, pp. 631_641, Jan. 2019, doi: 10.1016/j.procs.2020.01.057.
14. I. Benchaji, S. Douzi, and B. E. Ouahidi, ``Credit card fraud detection model based on LSTM recurrent neural networks," J. Adv. Inf. Technol., vol. 12, no. 2, pp. 113_118, 2021, doi: 10.12720/jait.12.2.113-118.
15. Y. Fang, Y. Zhang, and C. Huang, ``Credit card fraud detection based on machine learning," Comput., Mater. Continua, vol. 61, no. 1, pp. 185_195, 2019, doi: 10.32604/cmc.2019.06144.
16. J. Forough and S. Momtazi, ``Ensemble of deep sequential models for credit card fraud detection," Appl. Soft Comput., vol. 99, Feb. 2021, Art. no. 106883, doi: 10.1016/j.asoc.2020.106883.
17. K. He, X. Zhang, S. Ren, and J. Sun, ``Deep residual learning for image recognition," 2015, arXiv:1512.03385.
18. X. Hu, H. Chen, and R. Zhang, ``Short paper: Credit card fraud detection using LightGBM with asymmetric error control," in Proc. 2nd Int. Conf. Artif. Intell. for Industries (AII), Sep. 2019, pp. 91_94, doi: 10.1109/AI4I46381.2019.00030.
19. J. Kim, H.-J. Kim, and H. Kim, ``Fraud detection for job placement using hierarchical clusters-based deep neural networks," Int. J. Speech Technol., vol. 49, no. 8, pp. 2842_2861, Aug. 2019, doi: 10.1007/s10489-019-01419-2.
20. M.-J. Kim and T.-S. Kim, ``A neural classi_er with fraud density map for effective credit card fraud detection," in Intelligent Data Engineering and Automated Learning, vol. 2412, H. Yin, N. Allinson, R. Freeman, J. Keane, and S. Hubbard, Eds. Berlin, Germany: Springer, 2002, pp. 378_383, doi:10.1007/3-540-45675-9_56.
21. N. Kousika, G. Vishali, S. Sunandhana, and M. A. Vijay, ``Machine learning based fraud analysis and detection system," J. Phys., Conf., vol. 1916, no. 1, May 2021, Art. no. 012115, doi: 10.1088/1742-6596/1916/1/012115.
22. R. F. Lima and A. Pereira, ``Feature selection approaches to fraud detection in e-payment systems," in E-Commerce and Web Technologies, vol. 278, D. Bridge and H. Stuckenschmidt, Eds. Springer, 2017, pp. 111_126, doi: 10.1007/978-3-319-53676-7_9.
23. Y. Lucas and J. Jurgovsky, ``Credit card fraud detection using machine learning: A survey," 2020, arXiv:2010.06479.
24. H. Zhou, H.-F. Chai, and M.-L. Qiu, ``Fraud detection within bankcard enrollment on mobile device based payment using machine learning," Frontiers Inf. Technol. Electron. Eng., vol. 19, no. 12, pp. 1537_1545, Dec. 2018, doi: 10.1631/FITEE.1800580.
25. S. Makki, Z. Assaghir,Y. Taher, R. Haque, M.-S. Hacid, and H. Zeineddine, ``An experimental study with imbalanced classi_cation approaches for credit card fraud detection," IEEE Access, vol. 7, pp. 93010_93022, 2019, doi: 10.1109/ACCESS.2019.2927266.
26. I. Matloob, S. A. Khan, and H. U. Rahman, ``Sequence mining and prediction-based healthcare fraud detection methodology," IEEE Access, vol. 8, pp. 143256_143273, 2020, doi: 10.1109/ACCESS.2020.3013962.

27. I. Mekterovi¢, M. Karan, D. Pintar, and L. Brki¢, ``Credit card fraud detection in card-not-present transactions: Where to invest?" Appl. Sci., vol. 11, no. 15, p. 6766, Jul. 2021, doi: 10.3390/app11156766.
28. D. Molina, A. LaTorre, and F. Herrera, ``SHADE with iterative local search for large-scale global optimization," in Proc. IEEE Congr. Evol. Comput. (CEC), Jul. 2018, pp. 1_8, doi: 10.1109/CEC.2018.8477755.
29. M. Muhsin, M. Kardoyo, S. Arief, A. Nurkhin, and H. Pramusinto, ``An analyis of student's academic fraud behaviour," in Proc. Int. Conf. Learn. Innov. (ICLI), Malang, Indonesia, 2018, pp. 34_38, doi: 10.2991/icli-17.2018.7.
30. H. Najadat, O. Altiti, A. A. Aqouleh, and M. Younes, ``Credit card fraud detection based on machine and deep learning," in Proc. 11th Int. Conf. Inf. Commun. Syst. (ICICS), Apr. 2020, pp. 204_208, doi: 10.1109/ICICS49469.2020.239524.
31. A. Pumsirirat and L. Yan, ``Credit card fraud detection using deep learning based on auto-encoder and restricted Boltzmann machine," Int. J. Adv. Comput. Sci. Appl., vol. 9, no. 1, pp. 18_25, 2018, doi: 10.14569/IJACSA.2018.090103.
32. P. Raghavan and N. E. Gayar, ``Fraud detection using machine learning and deep learning," in Proc. Int. Conf. Comput. Intell. Knowl. Economy (ICCIKE), Dec. 2019, pp. 334_339, doi: 10.1109/ICCIKE47802.2019.9004231.
33. M. Ramzan, A. Abid, H. U. Khan, S. M. Awan, A. Ismail, M. Ahmed, M. Ilyas, and A. Mahmood, ``A review on State-of-the-Art violence detection techniques," IEEE Access, vol. 7, pp. 107560_107575, 2019, doi: 10.1109/ACCESS.2019.2932114.
34. M. Ramzan, H. U. Khan, S. M. Awan, A. Ismail, M. Ilyas, and A. Mahmood, ``A survey on state-of-the-art drowsiness detection techniques," IEEE Access, vol. 7, pp. 61904_61919, 2019, doi: 10.1109/ACCESS.2019.2914373.
35. A. Rb and S. K. Kr, ``Credit card fraud detection using arti_cial neural network," Global Transitions Proc., vol. 2, no. 1, pp. 35_41, Jun. 2021, doi: 10.1016/j.gltp.2021.01.006.
36. N. F. Ryman-Tubb, P. Krause, and W. Garn, ``How arti_cial intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark," Eng. Appl. Artif. Intell., vol. 76, pp. 130_157, Nov. 2018, doi: 10.1016/j.engappai.2018.07.008.
37. I. Sadgali, N. Sael, and F. Benabbou, ``Adaptive model for credit card fraud detection," Int. J. Interact. Mobile Technol., vol. 14, no. 3, p. 54, Feb. 2020, doi: 10.3991/ijim.v14i03.11763.
38. Y. Sahin and E. Duman, ``Detecting credit card fraud by ANN and logistic regression," in Proc. Int. Symp. Innov. Intell. Syst. Appl., Jun. 2011, pp. 315_319, doi: 10.1109/INISTA.2011.5946108.
39. I. Sohony, R. Pratap, and U. Nambiar, ``Ensemble learning for credit card fraud detection," in Proc. ACM India Joint Int. Conf. Data Sci. Manage. Data, Jan. 2018, pp. 289_294, doi: 10.1145/3152494.3156815.
40. B. Stojanovi¢, J. Bo$^{o}$i¢, K. Hofer-Schmitz, K. Nahrgang, A. Weber, A. Badii, M. Sundaram, E. Jordan, and J. Runevic, ``Follow the trail: Machine learning for fraud detection in _ntech applications," Sensors, vol. 21, no. 5, p. 1594, Feb. 2021, doi: 10.3390/s21051594.
41. C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, ``Inception-v4, inception-Res Net and the impact of residual connections on learning," 2016, arXiv:1602.07261.
42. H. Tingfei, C. Guangquan, and H. Kuihua, ``Using variational auto encoding in credit card fraud detection," IEEE Access, vol. 8, pp. 149841_149853, 2020, doi: 10.1109/ACCESS.2020.3015600.

43. D. Varmedja, M. Karanovic, S. Sladojevic, M. Arsenovic, and A. Anderla, ``Credit card fraud detection_machine learning methods," in Proc. 18th Int. Symp. INFOTEH-JAHORINA (INFOTEH), Mar. 2019, pp. 1_5, doi: 10.1109/INFOTEH.2019.8717766.
44. S. Warghade, S. Desai, and V. Patil, ``Credit card fraud detection from imbalanced dataset using machine learning algorithm," Int. J. Com-put. Trends Technol., vol. 68, no. 3, pp. 22_28, Mar. 2020, doi: 10.14445/22312803/IJCTT-V68I3P105.
45. N. Youse_, M. Alaghband, and I. Garibay, ``A comprehensive survey on machine learning techniques and user authentication approaches for credit card fraud detection," 2019, arXiv:1912.02629.
46. X. Zhang, Y. Han, W. Xu, and Q. Wang, ``HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture," Inf. Sci.
47. F. K. Alarfaj, I. Malik, H. U. Khan, N. A., M. Ramzan, and M. Ahmed “Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms” Apr.2022, doi: 10.1109/ACCESS.2022.31668