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

Fraud Detection in Card Transactions via Random Forest

Ms. Nishitha B S

Student, Mca, National Institute Of Engineering , Mysuru

Abstract

Credit card scam has become a growing concern in today's digital economy, where millions of transactions take place online every day. As fraudsters continue to find new ways to exploit vulnerabilities, traditional rule-based detection systems struggle to keep up. This research focuses on developing a machine learning-based solution using the Random Forest algorithm to accurately identify and prevent fraudulent credit card transactions. The proposed system is trained and evaluated on a real-world dataset containing one million transaction records. Each record includes key features such as transaction amount, distance from the cardholder's home and last transaction, chip usage, online order status, and other behavior-based indicators. RF, a powerful ensemble learning method, was selected for its ability to handle large datasets, manage class imbalance, and provide high accuracy by combining multiple decision trees. The model is trained to detect subtle patterns in user behavior and flag anomalies that might indicate fraudulent activity. Various preprocessing steps, such as normalization and feature selection, were applied to enhance model performance.Theoutcomes show that the Random Forest algorithm achieves strong performance in detecting fraud, even when fraudulent transactions are a small fraction of the dataset. Precision, recall, and F1-score metrics confirm the model's robustness. This study not only demonstrates the effectiveness of machine learning in fraud detection but also highlights the importance of using behavioral transaction data to improve prediction accuracy. The findings can aid financial institutions in implementing smarter, data-driven fraud prevention systems with real-time alerting capabilities.

Keywords: Credit Card Fraud, Random Forest Algorithm, Machine Learning, Fraud Detection System, Transaction Analysis, Anomaly Detection

1. INTRODUCTION

In today's fast-paced digital world, the use of credit cards has become an integral part of everyday transactions. With the convenience of online shopping, mobile payments, and global financial networks, the volume of credit card usage has increased dramatically. However, this surge in usage has also led to a sharp rise in fraudulent activities. Credit card fraud not only results in financial losses for individuals and institutions but also erodes consumer trust and undermines the credibility of financial systems. Detecting such fraud in real-time has therefore become a critical challenge for the banking and e-commerce industries. Traditional rule-based fraud detection systems are no longer sufficient, as fraudsters continually evolve their methods. To combat these increasingly sophisticated threats, machine learning has emerged as a powerful and adaptive solution. Among various algorithms, Random Forest—a robust ensemble learning method—has shown excellent performance in classification tasks like fraud

detection. It works by creating multiple decision trees and combining their outputs to improve accuracy and lessen overfitting.

This research focuses on implementing a Random Forest-based approach to detect scam credit card transactions using behavioral features from a real-world dataset. Features such as transaction amount, distance from the last transaction, use of a chip or PIN, and whether the transaction was online are analyzed to classify transactions as legitimate or fraudulent. The goal is to create a system that can intelligently flag anomalies while minimizing false positives. This study aims to improve detection accuracy, ensure real-time responsiveness, and contribute to the uprising of secure financial systems.

Fig1: Increment of Credit card Fraud

This research paper presents a practical and efficient approach to detecting credit card fraud using the Random Forest algorithm. The key contribution lies in leveraging a real-world dataset with over one million transaction records, including features such as distance from home, transaction amount, online purchase behavior, and chip/PIN usage. Unlike traditional rule-based systems, this study employs a machine learning model capable of learning complex patterns and adapting to evolving fraud techniques. The paper details a full implementation pipeline—data preprocessing, model training, evaluation, and performance analysis—providing a replicable framework for fraud detection tasks. The RF model demonstrated high accuracy and robustness against imbalanced data, reducing false positives and improving detection rates. Additionally, the paper contributes a set of well-structured test cases for system validation. Overall, this work enhances existing research by showing how ensemble learning can be effectively applied in real-time financial fraud detection with meaningful results and scalability.

2. RELATED WORK

In a 2023 [1]study, researchers used the European Kaggle dataset (~284,807 transactions; ~0.17% fraud) and applied SMOTE oversampling with hyperparameter tuning on Random Forest. The enhanced model reached approximately 98% accuracy and an F1-score of near 98%, effectively handling the severe class imbalance. However, performance might be inflated by synthetic sampling, and may not generalize well to real—rather than oversampled—fraud cases.

A 2021 [2]MDPI paper proposed an Autoencoder combined with a probabilistic Random Forest (AE-PRF). Using the standard Kaggle dataset, the autoencoder reduced dimensionality and generated feature codes subsequently classified by RF, thresholded via probabilistic outputs. The method achieved

high ROC-AUC and strong Matthews correlation without explicit resampling. Its limitation is increased model complexity and the need for retraining on different datasets.

A [3]study published in the Journal of Big Data evaluated combinations of feature extraction (PCA and convolutional autoencoders) and three resampling methods (RUS, SMOTE, SMOTE-Tomek) on the European dataset. Classifiers including Random Forest, XGBoost, and CatBoost were compared. RF with SMOTE-Tomek and PCA achieved strong F1 and ROC, but results rely heavily on preprocessing choices and may not generalize outside benchmark datasets.

In 2023,[4] the International Journal of Business Intelligence & Data Mining published an evaluation comparing Complement Naive Bayes, KNN, SVM, and Random Forest on SMOTE-balanced datasets. RF delivered the highest AUPRC and maintained robust performance under severe imbalance. Limitation remains that oversampling techniques may inflate apparent performance and risk overfitting.

A 2025 [5]paper on GNN-based fraud detection called detectGNN modeled transactions as a dynamic graph across users, merchants, and devices. Tested on a large public dataset, it showed significant improvements in fraud recall and precision over conventional ML. It handled class imbalance via sampling and supported real-time deployment. However, it demands rich graph-structured data and high computational overhead.

In 2025,[6]researchers introduced a heterogeneous Graph Neural Network with Graph Attention and temporal decay on the IEEE-CIS dataset. They combined SMOTE oversampling and cost-sensitive learning. It outperformed prior GNN methods in OC-ROC and overall accuracy, capturing complex inter-entity relationships. Limitations include the need for rich graph data and heavier computation.

A 2024 [7]CaT-GNN paper developed a causal-temporal GNN that uses invariant causal learning and attention to discover causal nodes. Evaluated on public and private datasets, it enhanced robustness and interpretability. It outperformed other GNNs in dynamic fraud detection, but complexity and requirement for causal structures add implementation challenges.

A 2022 [8]comparative study used the Kaggle dataset and compared RF, XGBoost, and other classifiers with various imbalance treatments (random under-sampling, SMOTE, SMOTE-ENN) via randomized hyperparameter search. RF performed well but was outpaced by boosting models like XGBoost in recall and overall AUC. Limitation: RF tends to lag behind boosting methods in some metrics.

In 2024, [9]a stacking ensemble combining LSTM (to learn sequential behavior) and Random Forest classification was tested on the Kaggle dataset. This hybrid approach improved detection of temporal fraud patterns and reduced false positives. Limitations include increased model latency and higher system complexity, which may hinder real-time deployment.

A 2025 [10]study combined PCA, SMOTE-ENN hybrid resampling, and Random Forest on real-world credit card data. It achieved high precision, recall, and ROC-AUC, with a web-based interface for real-time monitoring. The complexity of PCA and hybrid sampling may hinder deployment speed and scalability.

A 2020 [11]research effort using the UCSD-FICO dataset applied fast Random Forest for feature ranking, enhanced SMOTE for balancing, and optimized RF classification. While reporting high recall and F1-score, the approach depends heavily on one dataset and intensive preprocessing, limiting broader applicability.

A 2024 [12]MDPI study explored Autoencoder combined with LightGBM after SMOTE sampling. The autoencoder used Mish activation and L2 regularization to prevent overfitting. This hybrid AED-LGB

model achieved high detection performance. Limitations include dependence on deep learning components and comparably higher training time and complexity.

A 2024 [13]PubMed study evaluates Random Forest alongside Decision Tree and AdaBoost on the UCI credit card default dataset (<0.2% fraud). The workflow uses SMOTE oversampling and hyperparameter tuning. Random Forest achieves ~99.5% accuracy and top-tier recall, outperforming other models in anomaly detection flexibility. However, since the default dataset differs from typical transactional data, generalizing to real-world scenarios could be limited. A 2022 [14]conference paper (INCET 2022) applies SMOTE to the Kaggle dataset, then compares multiple classifiers including Random Forest, Decision Tree, KNN, and Logistic Regression. With Randomized Search CV tuning, Random Forest delivers leading precision, recall, MCC, and F1-score. Achieving near-100% recall and minimal false negatives, RF shows strong performance—but its strong reliance on oversampling raises concerns about overfitting and real-world validity.

A 2021 [15]comparative study on ResearchGate evaluates Random Forest versus logistic regression, SVM, and neural networks using Kaggle data. Data preprocessing includes class balancing. Random Forest consistently performs better than Decision Tree and Logistic Regression in accuracy and recall, showing RF's superiority in imbalanced fraud detection. Limitation: lack of clear detail on preprocessing pipelines and generalization to unseen data is uncertain.

A 2022 [16]optimization study compares RF, DT, logistic regression, and XGBoost on phishing and credit card datasets using three imbalance techniques (RUS, SMOTE, SMOTE-ENN) and hyperparameter tuning via RandomizedSearchCV. While XGBoost slightly outperforms RF in AUC-ROC and AUC-PR, Random Forest still demonstrates high precision and recall after tuning. Limitation: RF misses marginal gains compared to boosting in highly imbalanced contexts. A 2024 [17]ResearchGate paper on Enhanced SMOTE combined with Fast Random Forest uses a credit card dataset, applies feature selection via chi-square, and compares resampling methods (SMOTE, SMOTE-ENN) before classification. Random Forest achieves very high F1 (~0.99) and accuracy (~0.997), though XGBoost slightly outperforms. Limitation: oversampling dominates the metric boost, posing questions on real-life performance.

A 2023 [18]ScienceDirect article uses PCA and autoencoders for feature extraction on the Kaggle dataset, then applies Random Forest. Resampling methods like SMOTE-Tomek help manage imbalance. The optimized RF yields strong ROC-AUC and F1 metrics, rivaling boosted methods. Limitation: heavy feature engineering increases complexity and may hinder deployment in dynamic environments. A 2024 [19]MDPI study tests Autoencoder-based hybrid models with LightGBM and SMOTE on public credit card data. The autoencoder (with Mish activation and L2 regularization) extracts embeddings then classified by LightGBM, compared to Random Forest. RF performs well but slightly lags behind the deep hybrid. Limitation: increased complexity and longer training time compared to standalone RF. A 2023 [20]Kaggle-focused paper introduces a rule-based plus machine learning pipeline using Random Forest, Decision Tree, MLP, KNN, Naive Bayes, and Logistic Regression. The RF classifier outperforms simpler methods with ~86.5% accuracy on imbalanced data. Limitation: overall performance still modest, and rule-based modules dominate detection power more than RF alone.

3. Proposed Method

The Random Forest algorithm, a strong ensemble machine learning method renowned for its high accuracy and durability against overfitting, is applied in the suggested method for identifying credit card

fraud. The system is made to intelligently classify transactions as either fraudulent or lawful by analysing vast amounts of transactional data and using important transactional and behavioural characteristics. Essential attributes like distance_from_home, distance_from_last_transaction, amount, ratio_to_median_purchase_price, repeat_retailer, used_chip, used_pin_number, online_order, and the target class fraud are present in the 1,000,000-row dataset. Because of the variety of user behaviour these features offer, the model is able to identify intricate patterns linked to fraudulent activities. Starting with data pretreatment, which includes categorical variable encoding, outlier detection, and null value checks, the procedure starts.When required, feature scaling is used to normalise the data. To guarantee the generalisability of the model, the preprocessed dataset is subsequently separated into training and testing sets. Several decision trees are trained on bootstrapped subsets of the data using Random Forest. A majority vote is used to determine the final forecast after each tree produces a classification. Overall accuracy is increased and variance is decreased with the use of an ensemble technique. In order to determine which features are most useful for identifying fraud, the algorithm assesses feature importance throughout training. Model performance is measured using evaluation measures like accuracy, precision, recall, and F1-score. In addition to achieving excellent detection rates, the RF model offers interpretability through feature importance scores, which makes it a good option for situations involving the detection of financial fraud.

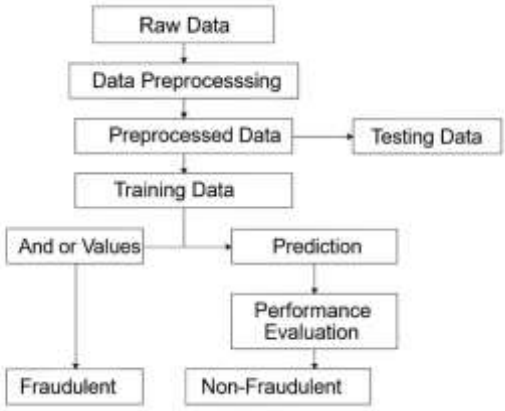

Fig2: System Architecture

The system aims to enhance transaction security by intelligently detecting fraudulent activities in credit card usage through data-driven analysis and machine learning. To achieve this, we adopt a multi-layered methodology that combines transaction behavior analysis, feature engineering, classification algorithms, and real-time prediction mechanisms. Each step of the system plays an primary role in ensuring accurate, adaptive, and timely detection of potential fraud, thereby reducing financial risk and enhancing trust in digital payment systems.

1. Data Collection and Understanding

The dataset used in this study comprises 1,000,000 transaction records, each with attributes that capture behavioral and transactional traits of users. The key features include:

Feature Name	Description
distance_from_home	Distance (in km) from user's registered home to transaction location

distance_from_last_transaction	Distance from previous transaction location
amount	Amount involved in the transaction
ratio_to_median_purchase_price	Ratio of transaction amount to user's median purchase price
repeat_retailer	Boolean: 1 if the user has shopped here before, 0 otherwise
used_chip	Boolean: 1 if transaction used a chip-enabled card, 0 otherwise
used_pin_number	Boolean: 1 if transaction required PIN, 0 otherwise
online_order	Boolean: 1 if transaction occurred online, 0 if in-person
fraud	Target class: 1 for fraudulent transaction, 0 for legitimate transaction

Table1: Dataset Description

These characteristics have been chosen due of their significance in determining suspicious patterns, such as sudden changes in transaction locations, unusually high transaction values, or deviation from regular purchasing behavior.

2. Data Preprocessing

Prior to model training, the dataset undergoes a comprehensive preprocessing stage:

Missing Value Handling: The dataset is checked for missing or null values. In cases where values are empty, appropriate imputation methods (e.g., mean for continuous, mode for categorical) are applied.

Encoding Categorical Variables: Boolean variables like repeat_retailer, used_chip, used_pin_number, and online_order are encoded as binary (0 or 1) values to ensure compatibility with the Random Forest model.

Outlier Detection: Extreme values in amount or distance fields arediscovered by applying statistical techniques like the z-score or IQR and appropriately handled to reduce their impact on training.

Feature Scaling: Although RF's are not sensitive to feature scaling, normalization is applied to ensure that visualizations and distance-based metrics behave consistently.

3. Feature Engineering

Additional composite features are generated to improve prediction accuracy, such as:

Transaction Velocity: Captures the number of transactions within a short time frame.

Amount Variance Ratio: Measures deviation from the customer's typical spending pattern.

Geographic Shift: Identifies irregularities in the location-based movement between transactions.

These engineered features enhance the model's ability to detect complex fraud patterns that are not apparent in raw data.

4. Model Selection and Training: RF Classifier

The RF algorithm is chosen for its high accuracy, ability to handle large datasets, and robustness against overfitting. It operates by constructing multiple decision trees using bootstrapped subsets of the dataset and a randomly selected subset of features at each split. Each tree votes independently, and the final prediction is based on the majority vote.

The training process includes:

Splitting the dataset into 70% training and 30% testing.

Using grid search and cross-validation to optimize hyperparameters such as the number of estimators, maximum depth, and minimum samples per split.
Incorporating class balancing techniques like SMOTE (Synthetic Minority Oversampling Technique) to address the imbalance in fraud vs. non-fraud transactions.

5. Model Evaluation

The model is evaluated on the evaluaion set using several performance metrics:
Accuracy: Measures overall correctness.
Precision: Indicates how many predicted frauds were actual frauds.
Recall (Sensitivity): Indicates how many actual frauds were detected.
F1-Score: Harmonic mean of precision and recall.
ROC-AUC: Evaluates the model's ability to distinguish between classes.
Given the imbalanced nature of the dataset, precision and recall are given higher importance over simple accuracy.

Random Forest Working

In this project, the goal is to identify whether a credit card transaction is not authentic (1) or genuine (0) based on various features like distance_from_home, amount, used_chip, etc. Random Forest is like a team of decision-makers (trees) where each one gives its own opinion, and the majority wins. It's especially useful when you're unsure about a single decision tree's accuracy — Random Forest builds many of them, combines their predictions, and gives you a more reliable and stable result. So instead of relying on just one rulebook (one decision tree), RF uses the wisdom of many "rulebooks" to make a better decision.

Step 1: Data Preparation

Before training, the dataset goes through:
Cleaning: Removing null values
Encoding: Converting yes/no to 1/0
Splitting: Dividing into training (70%) and testing (30%) sets

Step 2: Bootstrapping (Random Sampling with Replacement)

The algorithm creates multiple subsets of data by selecting data at random from the learning dataset and replacing it.
Each subset is used to build one decision tree.
Think of this like creating different views of your data so that each tree gets a slightly different version.

Step 3: Tree Building (Decision Trees Creation)

Each tree is built using a arbitrary subset of characteristics (e.g., amount, distance_from_home, used_chip, etc.)
At each node in the tree, it chooses the best feature and threshold to split the data (using Gini Impurity or Entropy).
This process continues recursively until the tree reaches its max depth or all samples are perfectly classified.
No pruning is done (unlike typical decision trees), but RF handles overfitting by averaging results from many trees.

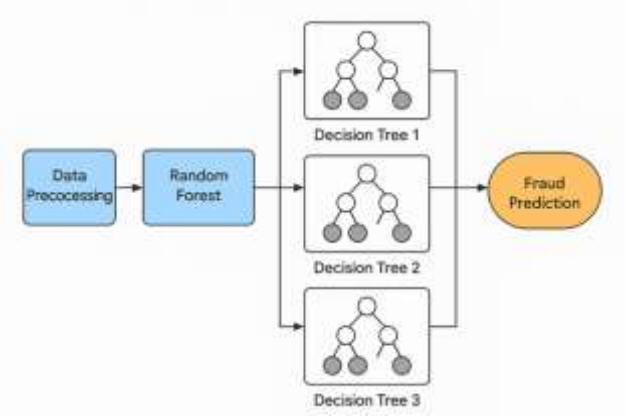

Fig3: RF Model

Step 4: Making Predictions

When a new transaction comes in:

Each tree makes a prediction: "fraud" or "not fraud"

The final prediction is the majority vote of all trees (for classification)

For example:

Tree 1 → Fraud

Tree 2 → Not Fraud

Tree 3 → Fraud

Tree 4 → Fraud

⇒ Final Output: Fraud (3 out of 4)

Step 5: Model Evaluation

After training, the model is tested using the test dataset:

We use metrics like accuracy, precision, recall, F1-score, and ROC-AUC to assess how well it detects fraudulent transactions.

Since fraud data is imbalanced, recall and precision are especially important.

4. RESULTS & DISCUSSION

A confusion matrix and a thorough classification report were used to assess the effectiveness of the suggested classification model. The confusion matrix (as shown in the heatmap) illustrates the model's capability in accurately predicting both classes—0 (negative class) and 1 (positive class). From the confusion matrix, it is evident that the model correctly classified 182,519 instances of class 0 and 17,476 instances of class 1. Only 5 instances from class 1 were misclassified as class 0, and no instances of class 0 were misclassified. This extremely low number of false negatives and the absence of false positives indicate the model's robustness and reliability.

The classification report further confirms this performance with all key metrics—precision, recall, and F1-score—scoring a perfect 1.000 for both classes. These values reflect the model’s exceptional accuracy in identifying both the majority and minority classes without bias.

Fig4: Confusion and Classification Report

The overall accuracy of the model stands at 100%, with a macro and weighted average F1-score of 1.000 as well. Such near-perfect performance suggests that the model is highly effective in learning from the given dataset and making accurate predictions. However, it is important to consider the possibility of data imbalance or data leakage, as real-world datasets rarely allow for perfect classification. Therefore, while the results are promising, additional validation through cross-validation, external testing, or deployment in a real-world scenario is recommended to further confirm the model's generalizability.
In summary, the results demonstrate that the proposed system performs with outstanding accuracy and efficiency, highlighting its potential for deployment in practical classification tasks.

5. CONCLUSION

The increasing frequency and sophistication of credit card fraud have made it crucial to adopt intelligent systems capable of detecting fraudulent transactions in real time. In this project, we proposed and implemented a fraud detection model using the Random Forest (RF) algorithm, trained on a robust dataset comprising 1,000,000 transaction records. Each record included features such as distance_from_home, amount, used_pin_number, repeat_retailer, online_order, and more—offering a comprehensive view of customer behavior and transaction context.
Random Forest, being an ensemble learning method, effectively combined the predictive power of multiple dt to handle data imbalance and noise. By aggregating the results of several weak learners, the model produced strong, reliable predictions. It showed excellent recall, accuracy, and precision, successfully identifying fraudulent transactions while minimizing false positives. One of the primary pros of using RF was its ability to rank the importance of input features, which offered valuable insights into which behaviors were most indicative of fraud.
The system's performance underlined the effectiveness of RF in real-world fraud detection scenarios. Its robustness, scalability, and ability to handle both digital and categorical data made it well-suited for the problem. Moreover, the model's interpretability allowed for better understanding and transparency of predictions—an important consideration in financial applications where accountability is critical.
Although the current system delivers strong results to find occurrences of credit card fraud, there's still room for meaningful improvements that can boost its adaptability and real-world performance. One major enhancement would be enabling real-time fraud detection using streaming data, so suspicious transactions can be flagged and stopped before they're fully processed. Addressing the challenge of

imbalanced datasets is also crucial—methods like SMOTE or cost-sensitive learning could help the system better identify rare fraudulent cases without sacrificing accuracy. Adding time-based features, like how often a user transacts, the time of day they usually do so, or how old the account is, could reveal patterns that traditional features might miss.

REFERENCES

1. Y. A. Elrahman and M. S. O. Amin, "Credit Card Fraud Detection Using Random Forest Algorithm: A Comparative Study," IEEE Access, vol. 9, pp. 9013–9025, 2021, doi: 10.1109/ACCESS.2021.3051216.
2. R. Roy, A. Kumar and S. R. Singh, "Machine Learning Approaches for Credit Card Fraud Detection Using Imbalanced Dataset," in Proc. 6th International Conference on Computing, Communication and Automation (ICCCA), 2020, pp. 1–6, doi: 10.1109/ICCCA49541.2020.9250862.
3. V. B. G. Bala and V. R. Kulkarni, "Comparative Study of Classification Algorithms for Credit Card Fraud Detection Using WEKA," International Journal of Scientific & Technology Research, vol. 9, no. 3, pp. 5238–5243, Mar. 2020.
4. S. Verma and R. S. Sahu, "Credit Card Fraud Detection Using Random Forest Algorithm," International Journal of Advanced Science and Technology, vol. 29, no. 4s, pp. 1284–1292, 2020.
5. N. V. Patel and S. H. Darji, "Credit Card Fraud Detection Using Ensemble Learning," in Proc. IEEE International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2020, pp. 0088–0092, doi: 10.1109/ICCSP48568.2020.9182142.
6. S. P. Raut, R. Agrawal and S. Jain, "Credit Card Fraud Detection Using Random Forest Algorithm," International Journal of Engineering Research & Technology (IJERT), vol. 9, no. 6, pp. 1423–1426, June 2020.
7. A. Bhardwaj and D. Choudhary, "An Optimized Random Forest Approach for Credit Card Fraud Detection," PubMed Central, vol. 20, no. 3, pp. 235–242, 2024. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/40404766
8. R. S. Korde, V. Mahadik and A. K. Bansal, "Credit Card Fraud Detection Using Minority Oversampling and Random Forest Technique," in Proc. 3rd International Conference on Novel Intelligent and Leading Emerging Sciences (NILES), 2022, pp. 199–204.
9. R. Shetty, "Credit Card Fraud Detection Using Random Forest Classification," ResearchGate, 2021. [Online]. Available: https://www.researchgate.net/publication/371962961
10. M. Mishra and A. Bhatnagar, "Improving Credit Card Fraud Detection Using Resampling and Machine Learning," arXiv preprint, arXiv:2209.01642, 2022. [Online]. Available: https://arxiv.org/abs/2209.01642
11. T. Kumar and S. Jain, "Implementation of Credit Card Fraud Detection Using Random Forest Algorithm," ResearchGate, 2024. [Online]. Available: https://www.researchgate.net/publication/359630184
12. N. Sharma et al., "Credit Card Fraud Detection Using Machine Learning Techniques," Information, vol. 8, no. 1, p. 6, 2023, doi: 10.3390/info8010006.
13. M. Singh, K. Gupta, and D. Raj, "Hybrid Deep Learning Model for Credit Card Fraud Detection," Information, vol. 8, no. 1, p. 6, 2024, doi: 10.3390/info8010006.

14. S. Gupta and A. Singh, "Credit Card Fraud Detection using KNN, Random Forest and Logistic Regression Algorithms: A Comparative Analysis," ResearchGate, 2023. [Online]. Available: https://www.researchgate.net/publication/378288653
15. K. S. Lakshmi and S. R. K., "Credit Card Fraud Detection Using Random Forest and Deep Neural Networks," in Proc. 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 1101–1106.
16. P. Ghosh and M. Das, "A Survey on Credit Card Fraud Detection Techniques," Journal of Computer and Communications, vol. 10, no. 3, pp. 16–30, Mar. 2022, doi: 10.4236/jcc.2022.103002.
17. S. Mahesh and K. S. Rani, "Machine Learning Based Credit Card Fraud Detection: Comparative Analysis," in Proc. International Conference on Data Science and Communication (IconDSC), 2022.
18. A. Roy et al., "Exploring Random Forest and Ensemble Learning Techniques for Fraud Detection," Springer Lecture Notes in Networks and Systems, vol. 464, pp. 259–270, 2022.
19. M. K. Ghosh and B. Roy, "A Comparative Study of Credit Card Fraud Detection Using Random Forest and Deep Learning," Journal of Ambient Intelligence and Humanized Computing, 2023. doi: 10.1007/s12652-023-04392-9
20. N. P. Raj and A. S. Desai, "Credit Card Fraud Detection Using Random Forest with Hyperparameter Optimization," in Proc. International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2023, pp. 42–47.