MAKÜ|MMF
MÜHENDİSLİK-MİMARLIK FAKÜLTESİ

Uluslararası Mühendislik, Tasarım ve Teknoloji Dergisi, 2025, 7(2), Sayfa: 70-77
International Journal of Engineering, Design and Technology, 2025, 7(2), Page: 70-77

Araştırma Makalesi
Research Article

Development of machine learning based fraud detection models for credit cards

Kredi kartları için makine öğrenimi tabanlı dolandırıcılık tespit modellerinin geliştirilmesi

Uygar Er[1], Ceren Ulus[2*], Mehmet Fatih Akay[2]

[1]Innovance, Department of Information Technologies, Istanbul, Türkiye
[2]Cukurova University, Faculty of Engineering, Department of Computer Engineering, Adana, Türkiye

Abstract: In today's global world, technology is rapidly developing and this can cause more risks, especially in sectors such as banking. Fraudsters create security vulnerabilities with many new techniques. Various approaches have emerged to prevent these vulnerabilities, but these approaches are generally inadequate due to reasons such as high data volume, multiple institutions, channels (mobile applications, websites, call centers) and fraudulent activities between locations. In this context, machine learning-based systems gain importance due to their dynamic structure. In this study, it is aimed to develop a model that provides fraudulent transaction detection using the Random Forest (RF) classifier. Docker and Kubernetes have been used for model distribution in the study. The performance of the developed model has been evaluated with Accuracy, Precision, Recall and F1 Score. With the developed fraud detection model, an Accuracy value of 0.771 has been achieved.

Keywords: Fraud Detection, Machine Learning, Random Forest

Özet: Günümüzün küresel dünyasında teknoloji hızla gelişmekte ve bu durum özellikle bankacılık gibi sektörlerde daha fazla risk oluşturabilmektedir. Dolandırıcılar birçok yeni teknikle güvenlik açıkları oluşturmaktadır. Bu açıkları önlemek için çeşitli yaklaşımlar ortaya çıkmıştır ancak bu yaklaşımlar yüksek veri hacmi, birden fazla kurum, kanal (mobil uygulamalar, web siteleri, çağrı merkezleri) ve lokasyonlar arası dolandırıcılık faaliyetleri gibi nedenlerden dolayı genellikle yetersiz kalmaktadır. Bu bağlamda dinamik yapıları nedeniyle makine öğrenmesi tabanlı sistemler önem kazanmaktadır. Bu çalışmada, Random Forest (RF) sınıflandırıcısı kullanılarak dolandırıcılık işlem tespiti sağlayan bir model geliştirilmesi hedeflenmiştir. Çalışmada model dağılımı için Docker ve Kubernetes kullanılmıştır. Geliştirilen modelin performansı Accuracy, Precision, Recall ve F1 Score ile değerlendirilmiştir. Geliştirilen dolandırıcılık tespit modeli ile 0,771'lik bir Accuracy değeri elde edilmiştir.

Anahtar Kelimeler: Dolandırıcılık Tespiti, Makine Öğrenmesi, Rastgele Orman

1. Introduction

Nowadays, rapidly developing technology has led to radical transformations in the banking sector. In this way, financial transactions can now be easily carried out through digital platforms without being limited to physical branches. These developments have enabled the diversification of customer profiles and the increase in the number of individuals benefiting from banking services.

Conducting banking transactions securely is of great importance in terms of brand image, customer satisfaction and loyalty. Establishing an infrastructure that provides secure payment, especially in credit card payments, facilitates banks' compliance with legal regulations and protects the integrity of the financial system. However, while the spread of digitalization has accelerated and facilitated banking transactions, it has also paved the way for an increase in cyber threats and security vulnerabilities. In addition, cyber attacks have become more complex, sophisticated and difficult to detect in parallel with technological developments. In this context, various types of attacks targeting banking systems, especially credit card fraud, stand out. For example, card information can be obtained with special devices placed on Automated Teller Machines (ATMs) or POS devices using the card skimming method; card data stolen using the card information stealing method can be used in unauthorized transactions over the internet. In phishing attacks, users are tricked into sharing their personal information via fake e-mails and links, and since social media has become an important socialization tool today, such attacks are spreading rapidly. In man-in-the-middle attacks, sensitive information is captured by targeting data communication between the user and the bank. In addition,

* İletişim Yazarı / *Corresponding author*. Eposta/*Email* : f.cerenulus@gmail.com
Geliş / *Received*: 15.05.2025, Revizyon / *Revised*: 27.05.2025
Kabul / *Accepted*: 13.06.2025

Burdur Mehmet Akif Ersoy University Press

https://doi.org/10.70669/ijedt.1700239

devices can be infiltrated with malware to record card information or provide remote access. Such attacks usually occur with fast and technically knowledge-requiring methods, and it is often not possible for customers to directly analyze these threats or intervene at the time of the attack. Therefore, these fraud methods lead to both financial losses for customers and serious reputational damage for banks (Shanshan Jiang et al., 2023). Therefore, early detection and prevention of possible attacks is of critical importance in terms of protecting the bank's brand image and sustaining customer loyalty.

Detecting large-scale frauds stands out as one of the biggest challenges in fraud investigations (Mahdi Rezapour, 2019). The main reason for this situation is that these fraudsters usually hide among millions of transaction data, causing security vulnerabilities. This data is high-volume and has a certain variety and noise. Therefore, it takes a long time to detect suspicious transactions. On the other hand, fraudulent activity is carried out through multiple institutions, channels (mobile application, website, call center) and locations. The location difference makes it difficult to collect and correlate data centrally. Therefore, different strategies are needed to detect credit card fraud.

Machine learning stands out as one of the most reliable and effective methods for credit card fraud detection today. Machine learning algorithms process these large data sets quickly and enable the detection of fraudulent transactions. Retrainable models provide a dynamic structure, so that as new fraud methods emerge, machine learning models can be updated against attacks by training them with new data. However, the reliability of these methods depends on their ability to correctly predict rights and wrongs. Fraudulent activities are designed to mimic normal user behavior, and a valid transaction performed by the user can lead the model to be mistaken and predicted as a false positive. This increases the false positive and negative rates, thus complicating the modeling and analysis processes. Therefore, these anomalies, which are considered fraudulent transactions, need to be correctly detected.

In this study, it is aimed to develop a model that provides fraudulent transaction detection using RF classifier in order to detect fraudulent transactions quickly and accurately.

This study is organized as follows: Section 2 includes relevant literature. Dataset generation is presented in Section 3. Development of fraud detection models is presented in Section 4. Section 5 provides model deployment process. Results and discussion are given in Section 6. Section 7 concludes the paper.

2. Literature Review

Mohamed Rusaam *et al.* (2025) presented a comparative analysis of machine learning algorithms for credit card fraud detection. A highly imbalanced dataset containing 284,807 transactions has been used. The performance of Logistic Regression (LR), Decision Trees (DT), RF, Support Vector Machine (SVM), Artificial Neural Network (ANN), and Gradient Boosting Machines (GBMs) has been analyzed. The results showed that ensemble methods, especially GBMs such as Extreme Gradient Boosting (XGBoost), effectively handle imbalanced data and outperform simpler models in terms of accuracy and recall.

Nazerke Baisholan *et al.* (2025) presented FraudX AI, an ensemble-based framework that addresses the challenges in fraud detection, including imbalanced datasets, interpretability, and scalability. The proposed model combines the results of RF and XGBoost, averaging probabilities, and optimizing thresholds to improve detection performance. The model has been evaluated on the European credit card dataset, preserving its inherent imbalance. It has been compared with eight baseline models, including LR and Gradient Boosting (GB). The results provided a recall value of 95% and an AUC-PR of 97%. Shapley additive explanations (SHAP) have been applied to interpret the model predictions. The proposed model achieved the best result in the comparisons.

Sreejith Sreekandan Nair *et al.* (2025) aimed to investigate how data balancing strategies and ensemble learning approaches can detect credit card fraud on imbalanced datasets. Data balancing strategies such as Synthetic Minority Oversampling Technique (SMOTE), Adaptive Synthetic Sampling Approach for Imbalanced Learning and Random Oversampling (RO) have been evaluated. K-Nearest Neighbors (KNN), DT, SVM and Long Short-Term Memory (LSTM) methods have been examined. Results have been provided with ROS on LSTM. When the obtained results have been analyzed, it has been observed that Random Oversampling ranked second on SVM with Recall 89.6%, F1 Score 91.5% and Precision 98.4%, Precision 97.6%, Recall 91.5% and F1 Score 93.9%, while Strengths, Weaknesses, Opportunities, Threats ranked third on SVM with Precision 91.7%, Recall 86.3% and F1 Score 89.9%. It has been observed that using SVM and LSTM classifiers along with oversampling strategies such as RO gives accurate and unbiased results for identifying credit card fraud in imbalanced datasets.

Kakaraparthi Hemanth *et al.* (2025) proposed a real-time, live detection approach for credit card fraud detection. A supervised machine learning approach has been used to cluster cardholders based on their transaction patterns and identify fraudulent behavior. The model has been iteratively updated by a feedback loop to adapt surprising responses to changes in fraud tactics. A LR has been used to detect credit card frauds through machine learning. Real-world credit card transactions with class imbalance have been used as the training dataset. Data preprocessing, normalization, coding, and sampling techniques have been applied. The resulting dataset has been divided into training and testing datasets.

Ronakbhai Bharatbhai Bhalala (2025) presented a review of the application of Python to deploy a fraud detection system for credit cards. First, a literature review of credit card fraud detection methods has been conducted. The challenges faced by traditional approaches have been highlighted. They proposed a model that required feature engineering, data preparation, and the application of various machine learning techniques. The results showed that the machine learning model achieved a high rate of success in detecting fraudulent credit card transactions. The effectiveness of the model has been evaluated using metrics such as accuracy, precision, recall, and F1 score.

Lossan Bonde and Abdoul Karim Bichanga (2025) proposed a novel ensemble deep learning model to detect credit card fraud. Convolutional Neural Networks (CNN) has been integrated for feature extraction, Gated Recurrent Units for sequential transaction analysis, and a Multilayer Perceptron (MLP) as a meta-learner within a stacking framework. To address the class imbalance issue and improve model accuracy, SMOTE and Edited Nearest Neighbors (SMOTE-ENN) technique has been applied. Experiments conducted on real-world credit card transaction data demonstrate that this hybrid approach outperformed traditional methods, enhancing accuracy and robustness, particularly in scenarios where fraudulent examples have been limited.

(Mahmoud Abdallah M. M. Mousa, 2025) proposed a machine learning based approach for credit card fraud detection. For this purpose, three different machine learning models including LR, RF, and XGBoost have been applied. These models have been trained on large-scale transaction data and evaluated using four performance metrics: Accuracy, Precision, Recall, and F1 score. The results showed that the RF model achieved an accuracy of 99.65%, the XGBoost model 99.963%, and the LR model 99.934%.

Irin Sultana *et al.* (2025) proposed a model that uses Graph Neural Networks (GNNs) for credit card fraud detection. The proposed model revealed hidden fraud and improved detection accuracy through time-based patterns and dynamic updates. Test results showed that GNNs detect complex and multi-layered fraud more effectively than traditional methods.

Yelleti Vivek *et al.* (2025) proposed a new profile-based One-Class Classification method. Clustering has been performed with K-Means. One rule has been generated from each cluster, thus creating a rule set consisting of K rules, each of which consists of conditions based on lower and upper bounds on all features. This rule set has been used to identify fraudulent transactions submitted in the testing phase. Sensitivity analysis has been performed by varying the number of conditions violated across the K rules. In the experimental results, the proposed approach exhibited successful performance in terms of classification rate.

Ge Yang (2024) aimed to effectively detect and prevent credit card fraud. For this purpose, he combined three machine learning algorithms, such as RF, SVM and LR. Through data preprocessing, feature engineering, and model optimization, the performance and stability of the models have been enhanced. Analysis using different levels of cross-validation showed that the RF model performed best in terms of accuracy and F1 score, SVM outperforms in Recall, and the LR model provided a high value of Area Under the Curve (AUC) in distinguishing between fraudulent and legitimate transactions.

Chang Yu *et al.* (2024) addressed the innovative applications of transformer models for fraud detection. To ensure the reliability of the data, data sources have been processed, and the issue of data imbalance has been resolved. Additionally, highly correlated vectors have been selected to strengthen the training process. Performance comparisons have been made with models such as SVM, RF, Neural Network, LR, XGBoost and TabNet to guarantee the reliability and practicality of the models. These models have been compared using metrics such as Precision, Recall, F1 Score, and Receiver Operating Characteristic (ROC)- AUC. The results showed that the transformer model achieved significant progress in the field of fraud detection, as well as in traditional applications.

Zengyi Huang *et al.* (2024 aimed to enhance the accuracy and efficiency of financial fraud detection to combat the increasing fraudulent activities. For this purpose, a machine learning based K-means clustering method has been proposed. As a result, compared to traditional rule-based detection methods, machine learning based approaches have been found to better adapt to the constantly evolving fraud techniques and patterns. Additionally, K-means clustering helped financial institutions to optimize resource allocation by focusing on high-risk areas.

Amal Al Ali *et al.* (2023) aimed to develop a financial statement fraud detection model using data from publicly accessible financial statements of businesses in the Middle East and North Africa region. LR, DT, SVM, Adaptive Boosting, RF, and XGBoost have been applied to develop the models. Additionally, the SMOTE algorithm has been used to address the issue of class imbalance in the dataset. According to the results, an Accuracy of 96.05% has been achieved by the XGBoost algorithm, which outperformed all other algorithms.

Yu-Teng Chang and Neng-Hsun Fan (2023) aimed to address the limitations of traditional methods in market targeting processes when faced with complex and constantly changing customer demands. To overcome this issue, an artificial intelligence-based market targeting approach has been proposed. Three different algorithms including Naive Bayes (NB), J48, and OneR have been used for model training and predictive analysis on test data. Accuracies of models have been measured as 100%,

91.7%, and 83.3% for NB, J48, and OneR, respectively while the F-measures for these algorithms have been 1, 0.909, and 0.8, respectively.

Shanshan Jiang *et al.* (2023) proposed a novel framework called Unsupervised Attentional Anomaly Detection Network-based to detect credit card fraud. Fraudulent transactions have been treated as anomalous samples and feature attention and Generative Adversarial Networks (GAN) and autoencoders have been used to effectively distinguish them from large volumes of transaction data. Experiments have been conducted on Kaggle and Institute of Electrical and Electronics Engineers - Computational Intelligence Society fraud detection datasets. The results showed that the proposed method outperforms existing fraud detection approaches.

Ketan Rathor *et al.* (2023) aimed to use automatic model generation techniques to learn normal behavior patterns from the state information of standard devices that constitute an ATMs. They performed data cleaning, integration, and data deduplication. Then, they performed feature selection and model training. Butterfly Optimization Algorithm has been used for feature selection and C-LSTM has been used for model training. Experimental results show that the proposed method performs well when compared with LSTM and CNN.

Yelleti Vivek *et al.* (2023) proposed ATMs fraud detection in static and streaming contexts. In static context, they investigated algorithms built on Spark and trained with various machine learning models, including NB, LR, SVM, DT, RF, Gradient Boosting Tree (GBT), and MLP. They balanced the data with SMOTE and its variants, GAN. In streaming context, sliding window based method is used. The method collected ATMs transactions and trained the models with these transactions. RF provided the best result, achieving the best average AUC of 0.975 in static context and the average AUC of 0.910 in streaming context.

Rong Zhang *et al.* (2023) proposed the use of machine learning algorithms to improve the detection of fraudulent transactions. The dataset obtained from Kaggle included features such as age, gender, payment area, and transaction amount. The dataset has been pre-processed to handle missing values and balance fraudulent samples. Then, KNN, NB, and SVM algorithms have been applied. The performance of these models has been evaluated using precision, recall, and F1-measure metrics. In the experimental results, it has been observed that the SVM model showed the highest performance with an Accuracy value of 99.23%.

Maram Alamri and Maram Alamri (2022) examined sampling techniques and their importance in solving the issue of imbalanced data. By reviewing previous studies, it has been concluded that hybrid sampling methods yield successful results and can enhance the effectiveness of fraud detection systems.

3. Dataset Generation

Data processing and environment design are given in ▶**Figure 1**.

Credit card and money transfer transactions, user information and transaction details have been collected (from

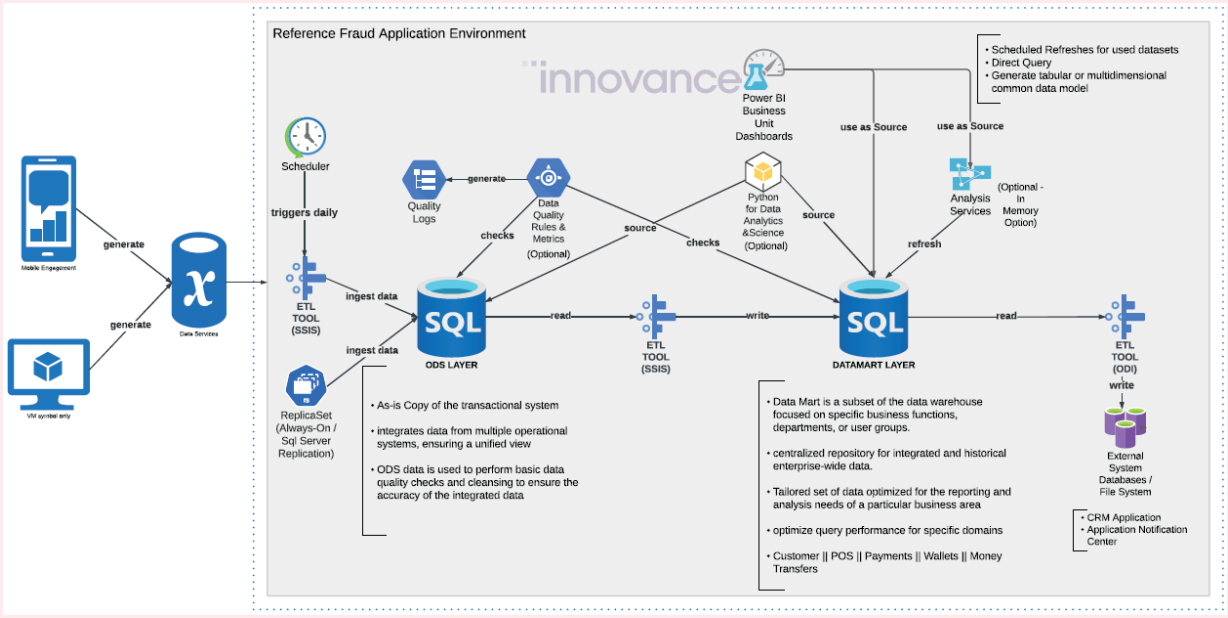

Figure 1. Data processing and environment design

banks/financial institutions). Data has been cleaned; missing or incorrect ones have been corrected, unnecessary features have been removed. For analysis, features such as transaction amount, date, type have been determined and derived. The basic attributes generated from credit card and money transfer transactions are given in ▶Table 1.

Table 1. The attributes and descriptions

Attribute	Description
Transaction Amount	The financial amount of the transaction.
Transaction Date and Time	The date and time the transaction was performed.
Transaction Type	The types of the transaction, such as shopping, cash withdrawal, online shopping.
Transaction Location and Geographic Location	The physical or virtual location where the transaction was performed.
Device Used	The type and features of the device performing the transaction (e.g., computer, mobile device).
Transaction Company or Store Information	The name and type of the company or store where the transaction was performed.
Transaction Approval Time	The time and process for confirming the transaction.
Payment Method	The payment method used in the transaction (e.g., credit card, debit card).
User Movements	The user's previous transactions and habits.
Shopping Frequency	The user's shopping frequency in a certain period of time.
Shopping Places	The places or types of shopping that the user prefers more frequently (e.g., grocery shopping, restaurant visits).
Shopping Time	The hours at which the user usually shops (e.g., morning, afternoon, evening).
Average Transaction Amount	The average amount the user usually spends.
Geographic Distribution of Shopping Locations	Different geographic regions or countries where the user shops.
Different Risk Parameters	Examples based on interpretation of the risk level of the transaction (specific region and time).

4. Fraud Detection Models

RF based predictive model has been developed for fraud detection using a dataset of 50,000 transactions including 1,000 confirmed fraud cases. RF is a popular ensemble learning technique in the literature that can be used for interaction detection, clustering, regression and classification. RF constructs trees using a large number of decision trees (DTs). A bootstrap sample and random selection of variables at each node are used to construct each tree. The accuracy of each tree is evaluated using Out-of-Bag (OOB) error rates. The final categorization is determined by the majority vote on all trees. In order to improve the performance of the developed RF model, it is necessary to adjust two important factors: the total number of trees and the amount of variables evaluated at each node (Soyoung Park and Jinsoo Kim, 2019). RF has been preferred in the model training process due to its high success in high-dimensional data, its ability to reduce overfitting thanks to its ensemble structure and its ability to capture non-linear relationships between features. To find the best value of hyperparameters, grid search has been used.

The hyperparameter values used for the development of forecasting model is given in ▶Table 2.

Table 2. Hyperparameter Values

Method	Hyperparameter Value
RF	"num_of_tree": [100] "Max_Depth": [10] "min_samples_leaf": [5]

5. Model Deployment

Docker is a containerization technology that enables software applications to run in an independent and isolated environment (Meenaxi M Raikar et al., 2024). As a first step, the Python application has been converted into a Docker container. For this, a Dockerfile has been created and the Python application has been copied into the Docker image.

The user sends a YAML file to Kubernetes. This YAML file contains information such as training time, data source, model type, and resource amount. Kubernetes starts the necessary pods based on the configuration. Orchestration is performed using tools such as Helm and Operator SDK. Information and commands related to the training process have been transmitted to worker pods via the created master pod. The necessary data from Google Cloud Storage is loaded via worker pods and training is performed. The obtained model output (artifact) data is transmitted to the master pod. The master pod combines the model output data to create the final model and saves the final model back to Google Cloud Storage.

Kubernetes automatically scales the application based on metrics such as CPU utilization. This allows the system to scale dynamically as data size increases.

6. Results and Discussion

The confusion matrix of the models is presented in ▶Table 3. Model performance metrics are presented in ▶Table 4.

Table 3. Confusion Matrix

	Predicted Fraud	Predicted Non-Fraud
Actual Fraud	685	315
Actual Non-Fraud	1140	47860

In the analysis of the confusion matrix,

- A total of 685 fake transactions have been correctly detected as fake (True Positive, TP).
- 1,140 non-fake transactions have been incorrectly classified as fake (False Positive, FP).
- 47,860 non-fake transactions have been correctly identified as not fake (True Negative, TN).
- 315 fake transactions have been incorrectly evaluated as not fake (False Negative, FN).
- This distribution provides an important basis for revealing the performance of the model in detecting fake transactions and the types of errors.
- The Positive Predictive Value (PPV) value is an important metric that shows how many of the examples the model predicts as "positive" are actually positive. When the PPV value is calculated, it is observed that it is approximately 37.53%. This situation shows that the model produces too many false positives when distinguishing fake transactions.
- The Negative Predictive Value (NPV) is a value that shows how many of the examples the model predicts as "negative" are actually negative. When the NPV value is calculated, it is seen that it is approximately 99.35%. Therefore, the model stands out as very successful in recognizing "clean" transactions.

Table 4. Model Performance Metrics

Metric	Value
Accuracy	0.771
Precision	0.543
Recall	0.685
F1-Score	0.605

In the results obtained with the developed model,

- The 77.1% Accuracy value shows that the model is generally successful.
- The Accuracy value showed that the model had a high rate of correct predictions for all classes.
- Approximately 2/3 of the positive classes were correctly captured and a 68.5% recall value has been obtained in this context. However, this value could be expected to be higher.
- The Precision value of 0.543 revealed that the model produced a significant number of false positives.
- The fact that recall is higher than precision shows that the model predicts more positives.
- With an F1 score of 0.605, the model provided a balanced performance against both false positive and false negative errors.

With the developed fraud detection models,

- It has been observed that transactions made at night have shown a higher correlation with fraud.
- It has been determined that the RF based model provides a high Accuracy value.
- Credit card and money transfer transactions can be monitored and analyzed instantly, and signs of fraud can be quickly detected.
- Signs of fraud are detected, financial losses are prevented, and financial damages are minimized.
- This innovative approach increases users' credit card security and strengthens customer trust, thus contributing to financial security and the protection of economic resources.

7. Conclusion

With today's developing technology, fraudsters can now cause faster and more dangerous vulnerabilities. Although various approaches have emerged to prevent security vulnerabilities, these approaches are not successful in big data and cannot adapt dynamically to developing technology. In this study, RF-based fraud detection models have been developed. Helm charts have been created for model distribution and Kubernetes has been used. With the developed models, signs of fraud could be detected, and financial losses have been prevented and financial losses have been minimized. The main contribution of this work is to provide a scalable and dynamic solution for detecting fraud in high-risk areas such as banking. The fact that the developed model can be run in a distributed manner with Docker and Kubernetes enables it to be integrated with real-time and multi-source data, which maximizes the practical applicability of the study. In the existing literature, traditional rule-based systems and statistics-based machine learning approaches are generally used. Moreover, these methods show limited success against real-world conditions. In this study, a machine learning model based on RF classifier is developed using real-world data. An important contribution to the literature is made by including model deployment and integration processes in the study. This is an important innovative feature that distinguishes the work from previous literature. This study provides contributions to technical teams such as data science, cyber security, IT operations operating in the banking sector and business operations such as risk management, customer

service, channel management. The model's suitability for real-time distribution provides a viable and value-added solution.

Research Ethics

Not applicable.

Author Contributions

Conceptualization: [Uygar Er, Ceren Ulus, Mehmet Fatih Akay], Methodology: [Uygar Er, Ceren Ulus, Mehmet Fatih Akay], Formal Analysis: [Uygar Er, Ceren Ulus, Mehmet Fatih Akay], Investigation: [Uygar Er, Ceren Ulus, Mehmet Fatih Akay], Resources: [Uygar Er], Data Curation: [Uygar Er], Writing - Original Draft Preparation: [Uygar Er, Ceren Ulus, Mehmet Fatih Akay], Writing - Review & Editing: [Uygar Er, Ceren Ulus, Mehmet Fatih Akay], Visualization: [Uygar Er, Ceren Ulus, Mehmet Fatih Akay], Supervision: [Mehmet Fatih Akay], Project Administration: [Uygar Er]

Competing Interests

The author(s) has declared no conflicts of interest.

Research Funding

None declared.

Data Availability

The raw data can be obtained at the request of the corresponding author.

Peer-review

Peer-reviewed by external referees.

Orcid

Uygar Er https://orcid.org/0009-0003-5659-1241
Ceren Ulus https://orcid.org/0000-0003-2086-6381
Mehmet Fatih Akay https://orcid.org/0000-0003-0780-0679

References

Baisholan, N., Dietz, J. E., Gnatyuk, S., Turdalyuly, M., Matson, E. T., & Baisholanova, K. (2025). FraudX AI: An Interpretable Machine Learning Framework for Credit Card Fraud Detection on Imbalanced Datasets. *Computers*, 14(4), 120.

Bhalala, R. B., & Patel, N. (2025). Machine Learning based Credit Card Fraud Detection Model. *IJFRI*, 1(1).

Bonde, L., & Bichanga, A. K. (2025). Improving Credit Card Fraud Detection with Ensemble Deep Learning-Based Models: A Hybrid Approach Using SMOTE-ENN. *Journal of Computing Theories and Applications*, 2(3), 384.

Hemanth, K., Virat, K. S., Rohith, M. D., Reddy, K. V. P., & Selv, A. S. (2025). Credit Card Fraud Detection using Machine Learning Methods. In 2025 Emerging Technologies for Intelligent Systems (ETIS), IEEE, pp. 1-6.

Mousa, M. A. M. (2025). Credit Card Fraud Detection in the Banking Sector: A Comprehensive Machine Learning Approach for Information Security. *Artificial Intelligence in Cybersecurity*, 2, pp. 1-13.

Nair, S. S., Lakshmikanthan, G., Belagalla, N., Belagalla, S., Ahmad, S. K., & Farooqi, S. A. (2025). Leveraging AI and Machine Learning for Enhanced Fraud Detection in Digital Banking System: A Comparative Study. In 2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT), IEEE, pp. 1278-1282.

Nijanthan, V., Muthukumaran, N., Pratheeshba, B., & Riyas Ahamed, M. (2025). The Impact of Machine Learning Algorithms on Credit Card Fraud Detection: A Comparative Study. In 2025 International Conference on Visual Analytics and Data Visualization (ICVADV), IEEE, pp. 1576-1580.

Sultana, I., Maheen, S. M., Kshetri, N., & Zim, M. N. F. (2025). detectGNN: Harnessing Graph Neural Networks for Enhanced Fraud Detection in Credit Card Transactions. arXiv preprint arXiv:2503.22681.

Vivek, Y., Ravi, V., Mane, A., & Naidu, L. R. (2025). Explainable One Class Classification for ATM Fraud Detection. In 2025 17th International Conference on COMmunication Systems and NETworks (COMSNETS), IEEE, pp. 114-119.

Huang, Z., Zheng, H., Li, C., & Che, C. (2024). Application of Machine Learning-Based K-Means Clustering for Financial Fraud Detection. *Academic Journal of Science and Technology*, 10(1), pp. 33-39.

Raikar, M. M., Patil, P., Guggari, S., Shavi, P., Mudavi, S., Patil, N., & Rangannavar, V. (2024). Leveraging Docker Containers for Deployment of Web Applications in Microservices Architecture. In 2024 First International Conference for Women in Computing (InCoWoCo), IEEE, pp. 1-6.

Yang, G. (2024). Credit Card Fraud Detection Based on Machine Learning Prediction. In 2024 2nd International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2024), Atlantis Press, pp. 35-45.

Yu, C., Xu, Y., Cao, J., Zhang, Y., Jin, Y., & Zhu, M. (2024). Credit card fraud detection using advanced transformer model. In 2024 IEEE International Conference on Metaverse Computing, Networking, and Applications (MetaCom), IEEE, pp. 343-350.

Ali, A. A., Khedr, A. M., El-Bannany, M., & Kanakkayil, S. (2023). A Powerful Predicting Model for Financial Statement Fraud Based on Optimized XGBoost Ensemble Learning Technique. *Applied Sciences*, 13(4), 2272.

Chang, Y. T., & Fan, N. H. (2023). A novel approach to market segmentation selection using artificial intelligence techniques. *The Journal of Supercomputing*, 79(2), pp. 1235-1262.

Jiang S, Dong R, Wang J, Xia M. (2023). Credit Card Fraud Detection Based on Unsupervised Attentional Anomaly Detection Network. *Systems*, 11(6):305.

Rathor, K., Vidya, S., Jeeva, M., Karthivel, M., Ghate, S. N., & Malathy, V. (2023). Intelligent System for ATM Fraud Detection System using C-LSTM Approach. In 2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC), IEEE, pp. 1439-1444.

Vivek, Y., Ravi, V., Mane, A. A., & Naidu, L. R. (2023). ATM fraud detection using streaming data analytics. arXiv preprint arXiv:2303.04946.

Zhang, R., Cheng, Y., Wang, L., Sang, N., & Xu, J. (2023). Efficient Bank Fraud Detection with Machine Learning. *Journal of Computati-*

onal Methods in Engineering Applications, 1-10.

Alamri M, Ykhlef M. (2022). Survey of Credit Card Anomaly and Fraud Detection Using Sampling Techniques. *Electronics*, 11(23):4003.

Mahdi Rezapour, (2019), Anomaly Detection using Unsupervised Methods: Credit Card Fraud Case Study, *International Journal of Advanced Computer Science and Applications(IJACSA)*, 10(11).

Park, S., & Kim, J. (2019). Landslide susceptibility mapping based on random forest and boosted regression tree models, and a comparison of their performance. *Applied Sciences*, 9(5), 942.