Optimized Credit Card Fraud Detection Leveraging Ensemble Machine Learning Methods

Al-Anood Al-Maari
School of Technology, Asia Pacific University of Technology and Innovation, Kuala Lumpur, Malaysia
alanood.ah98@gmail.com

Mohamed Abdulnabi
School of Technology, Asia Pacific University of Technology and Innovation, Kuala Lumpur, Malaysia
mohamed.shabbir@apu.edu.my (corresponding author)

Yogeswaran Nathan
School of Technology, Asia Pacific University of Technology and Innovation, Kuala Lumpur, Malaysia
yogeswaran.nathan@apu.edu.my

Aitizaz Ali
School of Technology, Asia Pacific University of Technology and Innovation, Kuala Lumpur, Malaysia
aitizaz.ali@apu.edu.my

Uzair Ali
Department of Computer Science, Abdul Wali Khan University, Mardan KPK, Pakistan
uzair9599@gmail.com

Maqbool Khan
Pak-Austra Fachhochschule, Institute of Applied Sciences and Technology (PAF-IAST), Haripur, Pakistan
maqbool.khan@fecid.paf-iast.edu.pk

Received: 19 January 2025 | Revised: 13 February 2025 and 19 February 2025 | Accepted: 27 February 2025

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | *DOI: https://doi.org/10.48084/etasr.10287*

ABSTRACT

The increasing number of online financial transactions, particularly those involving credit cards, underscores the urgent need for robust security systems to mitigate financial losses due to fraud. This paper presents a novel machine learning-based approach to credit card fraud detection that addresses the growing demand for enhanced security. Unlike traditional single-model approaches, the proposed ensemble model uniquely combines random forest, logistic regression, and AdaBoost techniques to accurately distinguish between legitimate and fraudulent transactions. The novelty of this work lies in its innovative soft voting mechanism, which aggregates the strengths of these diverse algorithms to achieve superior classification accuracy and minimize false positives. By leveraging the complementary strengths of each model, the ensemble approach provides a robust and adaptive framework capable of detecting emerging fraud patterns in real-time. The results demonstrate that the proposed ensemble model outperforms individual models, achieving an accuracy of 99.96%, a precision of 99.53%, and a recall of 100%, with an F1 score of 0.99 and an Area Under the Curve (AUC) of 1.0. These findings highlight the model's ability to significantly reduce false positives and negatives, making it a highly reliable solution for fraud detection. Furthermore, the model's performance was validated on the PaySim dataset, a large-scale synthetic dataset, where it achieved a 99.97% accuracy, demonstrating its generalizability and robustness in handling complex, real-world scenarios. This research contributes to the field by introducing a highly extensible and adaptable fraud detection framework that improves current solutions and provides a

foundation for future advancements in ensemble learning for fraud detection. The proposed model's ability to integrate multiple classifiers while maintaining interpretability and computational efficiency sets it apart from previous studies, offering a promising direction for enhancing the security of online financial transactions.

Keywords-credit card fraud detection; machine learning; ensemble learning; logistic regression; AdaBoost; random forest

I. INTRODUCTION

Financial services have been rapidly digitized to the point where financial transactions can be conducted in ways that were previously impossible, making the process far more convenient and accessible for consumers and businesses. However, this has also increased the problem of credit card fraud, which now poses a significant challenge to the financial ecosystem. Recent reports identify 2021 as the year when payment card fraud losses reached USD 32 billion, potentially rising to USD 48 billion by 2025 [1]. This alarming trend underscores the urgent need for robust, adaptive, and efficient fraud detection mechanisms to mitigate financial losses for consumers, businesses, and financial institutions. Traditionally, rule-based approaches and systems have been used for fraud detection. While they are capable of detecting known fraud patterns, they rely on predefined rules and historical data. Unfortunately, their inflexibility has made them increasingly inadequate in today's rapidly evolving and changing financial world. According to the authors in [2], rule-based systems are still limited to identifying only pre-qualified fraud patterns that can serve as targets for new and unprecedented fraud strategies. Typically, this results in a high false negative rate, i.e., fraudulent transactions are not detected, and a high false positive rate, i.e., many false alarms are generated, requiring costly manual intervention [3, 4]. In addition, due to the static nature of these systems, they need to be updated regularly as fraudsters' tactics are constantly changing, making them less efficient [5, 6].

To address these challenges, Machine Learning (ML) has been considered as a promising alternative for fraud detection. With the potential of ML models to process huge amounts of data, identify complex patterns, and easily adapt to new deceptions and fraud strategies in real-time, they have made great progress in improving fraud detection systems [7]. In terms of fraud detection, single ML models such as Support Vector Machines (SVMs), decision trees, and neural networks have been widely used and provide some good insights from the transaction data. For example, SVMs are very proficient in high-dimensional datasets and binary classification, while decision trees have the interpretability that is key to identifying determinants in fraud detection outcomes [8]. Neural networks, on the other hand, excel at modeling non-linear patterns, making them effective at identifying complex fraud schemes [9].

Despite their strengths, single ML models have limitations. Financial transaction data are often dynamic and imbalanced, leading to high rates of false positives and false negatives. In addition, these models require extensive data reformatting and are prone to overfitting, especially in high-dimensional datasets [10]. To address these shortcomings, researchers have focused on ensemble learning methods. Ensemble techniques have shown superior performance in fraud detection, for example, in terms of improved accuracy, precision and robustness when compared to single-model approaches [11].

Recent studies have demonstrated the effectiveness of these methods for fraud detection. For example, random forest, Adaptive Boosting (AdaBoost), and gradient boosting, as well as their ensemble models, have achieved accuracies of up to 95% [12, 13]. This is because these models work very well with imbalanced datasets and also mitigate false positives and negatives as fraud patterns change. However, the success of the existing ensemble methods has some shortcomings. The aforementioned methods combine similar types of models, such as tree-based algorithms, so they may not take advantage of different patterns that exist in the data [14]. Authors in [15] mention that some ensemble models lack interpretability, which is essential for real-life applications, especially in the financial sector where transparency is crucial. Furthermore, many studies ignore the problems of real-time processing and class imbalance, which are critical for effective fraud detection in dynamic environments [16].

This study addresses existing gaps by proposing a novel ensemble model that combines random forest, logistic regression, and AdaBoost. Logistic regression adds interpretability and probabilistic outputs, while random forest and AdaBoost enhance accuracy and robustness. By integrating these algorithms, the model overcomes the limitations of single-model and existing ensemble approaches, offering a more effective solution for real-time fraud detection. It is designed to handle imbalanced datasets, reduce false positives/negatives, and adapt to evolving fraud patterns.

II. METHODOLOGY

The main steps of the proposed model are shown in Figure 1, including data preparation and implementation of the ensemble model. Before using preprocessing steps such as outlier removal, missing value handling, and one-hot encoding to classify the transactions as legitimate or fraudulent, an autoencoder is used to train these transactions. This preprocessing enhances the ensemble model to accurately classify data using various algorithms.

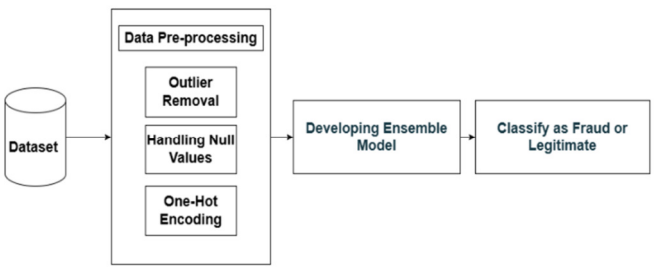

Fig. 1. Proposed credit card fraud transaction detection.

A. Data Collection

The data used in this research were obtained from Kaggle. The customer credit dataset [17, 18] was primarily used to train the ensemble model and evaluate its initial performance and efficiency. To further validate and confirm the model, the PaySim dataset [19, 20], a synthetic dataset of online payment transactions, was used. In this paper, the datasets were selected based on applicability and variability, and the models were trained and evaluated on the selected datasets.

1) Data Preparation

Data preparation is critical to correctly modeling the data. In this study, the dataset was cleaned to remove outliers and handle missing values to provide quality data for analysis.

a) Heatmap Analysis

Figure 2 shows a heatmap of the correlation matrix of financial behaviors and the transactions performed by the users. The rows are the transaction type and the columns are the customer ID or category. There is a strong positive relationship between 'balance' and 'balance_frequency', meaning that customers with high balances have more frequent transactions. Similarly, 'purchases' and 'oneoff_purchases' are highly correlated, indicating that one-time customers are also the highest spenders. However, 'credit_limit' and 'payments' have very low correlation coefficients. The heatmap summarizes which features to select and which to exclude when interpreting the model.

b) Outlier Detection

Z-scores identify anomalies by measuring deviations from the mean of the dataset. Data points that exceed a specified set (e.g., ±1.0) are flagged as outliers, improving data reliability [21].

c) Handling Missing Values

Missing entries were identified and removed to maintain dataset consistency. This simple yet essential step ensures that the model is accurate. However, care was taken to ensure the dataset's distribution [22].

d) One-Hot Encoding

Binary columns were created from categorical variables, preserving data integrity and using one-hot encoding. This contributes to efficient model training [23].

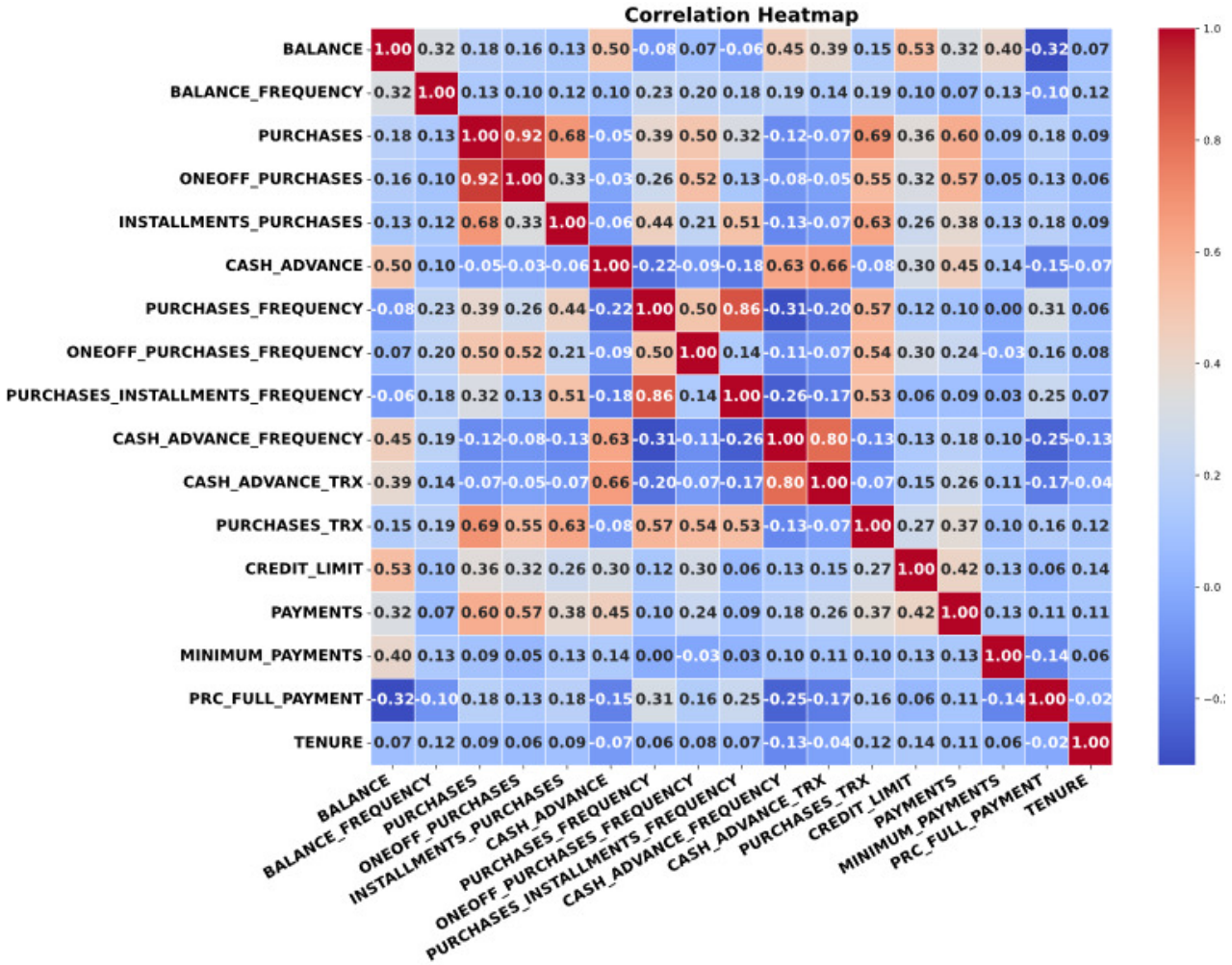

Fig. 2. Heatmap of feature correlations.

III. MODEL DEVELOPMENT

This paper presents an ensemble model to improve the predictive capabilities of fraud detection. Figure 3 illustrates the ensemble approach, which combines multiple machine learning techniques to improve precision and robustness.

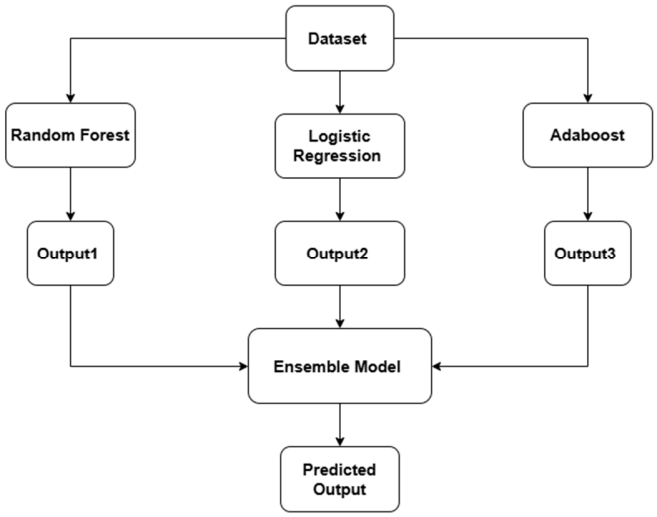

Fig. 3. Ensemble model.

A. Dataset

The success of a machine learning model depends on the quality of the dataset and rigorous preprocessing. In this study, the customer credit dataset [17] with 8,950 entries and 18 features was used, identifying key customer attributes as features and fraud indicators as the target variable.

1) Outlier Detection and Fraud Flag Creation

Outlier detection was conducted using the z-score method, which measures deviations from the mean. The z-scores for the 'purchases' feature were calculated and transactions greater than ±1.0 were flagged as outliers in order to identify anomalies without being too restrictive. Outliers were flagged as 1 for probable fraud and others were flagged as 0, based on a new binary 'fraud_flag' column for machine learning models. Figure 4 shows the added 'fraud_flag' column, which flags outliers as potential fraud. This improves the model's ability to distinguish normal from abnormal patterns, thereby improving fraud detection results.

FRAUD_FLAG
0
1

Fig. 4. Fraud_flag column.

2) Handling Missing Values

There were found 14 missing values: 1 in "credit_limit" and 13 in "minimum_payments." Due to the small number, the missing data were removed using Pandas' drop function, resulting in a dataset of 8,636 entries with 20 columns, including the added features 'z_score' and 'fraud_flag'. This approach improved the model's robustness and prediction accuracy by ensuring complete and reliable data.

3) Feature Encoding

The dataset has a categorical attribute, 'cust_id,' which represents different customer IDs. Since most machine learning algorithms typically require numerical inputs, a one-hot encoder was used to convert this categorical information into a numerical format. The encoding was done using the ColumnTransformer, which transformed the 'cust_id' attribute into a set of binary features, all related to unique customer IDs.

B. Ensemble Model Development

A voting classifier was used to combine several machine learning techniques to create the ensemble model. The ensemble method combines the most valuable features of several models, including logistic regression, random forest, and AdaBoost, to create a more robust, flexible support system for decision-making.

1) Aggregation of Model Outputs

Finally, soft voting aggregates the outputs of individual classifiers into a single, aggregated output. This ensemble learning technique makes the class prediction by averaging the predicted probabilities for each class across all classifiers and selecting the class with the highest averaged probability as the final prediction. Soft voting leverages the strengths of random forest, logistic regression, and AdaBoost classifiers to provide robustness and accuracy [24], allowing for more reliable decisions by combining multiple classifiers, especially in high-stakes scenarios such as fraud detection. For example, if random forest claims that there is no fraud, there is high confidence. If logistic regression claims there is fraud, the confidence is moderate. AdaBoost supports fraud with high confidence. This is the mechanism by which these probabilities are combined. The results show that this approach improves model performance by balancing the strengths of individual classifiers and considering the confidence levels in the final decision.

2) Classifiers

a) Random Forest

Random forest is a robust ensemble learning algorithm for classification and regression tasks. During training, it constructs multiple decision trees and outputs, and its predictions can belong to a mode (classification) or a mean (regression) [25]. Random forest was chosen for its ability to survive large datasets with many features. It reduces variance and thus increases stability and accuracy by averaging predictions across trees, even with noisy data.

b) Logistic Regression

Logistic regression is a basic statistical model that has been successfully used for binary classification. Fitting the data to an S-shaped logistic curve predicts the probability of class membership (e.g., fraud or non-fraud) [26]. Its simplicity and interpretability make it an attractive addition to the ensemble

model. Logistic regression provides linear decision boundaries and probabilistic outputs that are helpful in scenarios that require both classification and the probability of fraud. Logistic regression is incorporated into the proposed method to enhance more complex classifiers while maintaining interpretability in predictions.

c) AdaBoost

AdaBoost allows us to get more out of the model by focusing on the hard-to-classify cases. It combines different weak classifiers and iteratively adjusts the training data weights to give more weight to misclassified cases to form a strong classifier [27]. AdaBoost was included to improve the ensemble precision, especially when the ratio of fraudulent (minority) to non-fraudulent (majority) transactions is imbalanced. Ensemble's ability to correct for inaccuracies improves its performance in the fraud detection task.

d) Ensemble Method

Each classifier independently analyzes the data and assigns probabilities for predicting different patterns. These probabilities are aggregated from random forest, logistic regression, and AdaBoost, and the result is selected by the ensemble model based on the maximum combined probability. This method results in robust and reliable decision-making with the best performance in this study. Figure 5 shows the ensemble model pipeline, which consists of feature transformation and voting classifier integration.

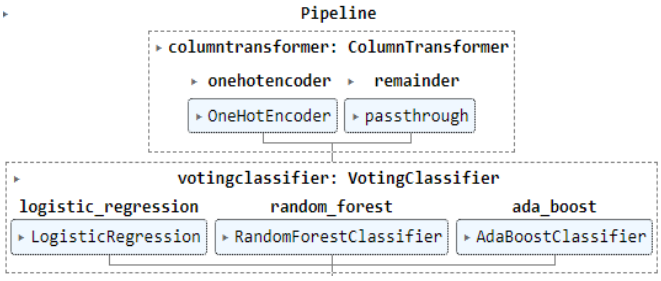

Fig. 5. Pipeline structure for model integration.

Random forest, logistic regression and AdaBoost were selected for the ensemble model because of their complementary strengths and suitability for the challenges of credit card fraud detection. Random forest was chosen for its high accuracy, robustness to overfitting, and ability to handle the high-dimensional data typical of transactional datasets. Logistic regression was chosen for its simplicity, interpretability and probabilistic outputs that are useful for understanding the likelihood of fraud. AdaBoost was chosen for its effectiveness in dealing with imbalanced datasets and its ability to focus on hard-to-classify cases, which are important for detecting rare fraudulent transactions. Other models, such as neural networks and SVMs, were ruled out based on their speed, interpretability, and ability to address issues like class imbalance and high dimensionality. The ensemble approach adds the strengths of the models and mitigates their weaknesses by using the combined model, resulting in a robust and efficient fraud detection system.

IV. MODEL EVALUATION

Accuracy, precision, recall, F1 score, confusion matrix and Area Under the Curve (AUC) were used to evaluate the performance of the proposed model [28].

Accuracy measures the total number of correct predictions using the total number of predictions and the proportion of correct predictions (true positive and true negative predictions):

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (1)$$

Precision assesses the accuracy of positive predictions, calculated as the ratio of true positives to the sum of true positives and false positives:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives+False Positives}} \quad (2)$$

Recall estimates how well the model detects all relevant instances, defined as the ratio of true positives to the sum of true positives and false negatives:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives+False Negatives}} \quad (3)$$

The F1 score combines precision and recall into a harmonic mean, providing a balanced evaluation of model performance:

$$\text{F1} = 2 * \frac{1}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}} \quad (4)$$

The confusion matrix captures the relationship between actual and predicted outcomes, including:

- True positives (TP): Correctly classified fraud cases.
- False negatives (FN): Missed fraud cases.
- False positives (FP): Misclassified non-fraud cases.
- True negatives (TN): Correctly classified non-fraud cases.

Finally, AUC measures the model's ability to distinguish between classes, with an AUC of 1 indicating perfect performance [29, 30].

A. Results and Discussion

The results on the customer credit dataset are compared between the individual classifiers and the ensemble model for their effectiveness and practical implications. A detailed summary of the metrics for logistic regression, random forest, AdaBoost, and the ensemble model is shown in Table I, where the ensemble model outperforms the others as follows:

- Accuracy: The ensemble model achieved an accuracy of 99.96%, only slightly lower than the best single model, AdaBoost. This high level of accuracy indicates that the model correctly diagnoses the majority of transactions as fraudulent or non-fraudulent.
- Precision and recall: The ensemble model achieved a perfect recall score of 1.0, detecting all fraudulent cases in the test dataset, which is necessary to avoid missing fraud cases. The score of 0.9953 is a high precision, meaning that almost all of the transactions flagged as fraudulent were correctly classified. The model is indeed reliable, as supported by the balanced F1 score.

- F1 score: The ensemble model achieved an F1 score of 0.99, indicating an excellent trade-off between precision and recall. Class imbalance is common in fraud detection, where the number of fraudulent cases is much smaller than the number of non-fraudulent cases, which is particularly critical.
- AUC and Receiver Operating Characteristic (ROC) curve: As shown in Figure 6, the ROC curve of the ensemble model resulted in an AUC of 1.0, which means perfect discrimination between fraudulent and non-fraudulent transactions. This indicates that the model is outstandingly reliable for fraud detection.
- Confusion matrix: The confusion matrix, shown in Figure 7, indicates that the ensemble model is highly accurate, with 2,376 true positives, 0 false positives, 0 false negatives, and 214 true negatives. Such a model supports error minimization and correct fraud detection.

TABLE I. COMPARISON OF SINGLE ML ALGORITHMS VS. THE ENSEMBLE MODEL

Model	Accuracy	Precision	F1 score	Recall
Logistic regression	0.9927	0.9756	0.95	0.9346
Random forest	0.9977	0.9940	0.97	0.9766
AdaBoost	0.9993	0.9953	0.98	0.99
Ensemble model	0.9996	0.9953	0.99	1

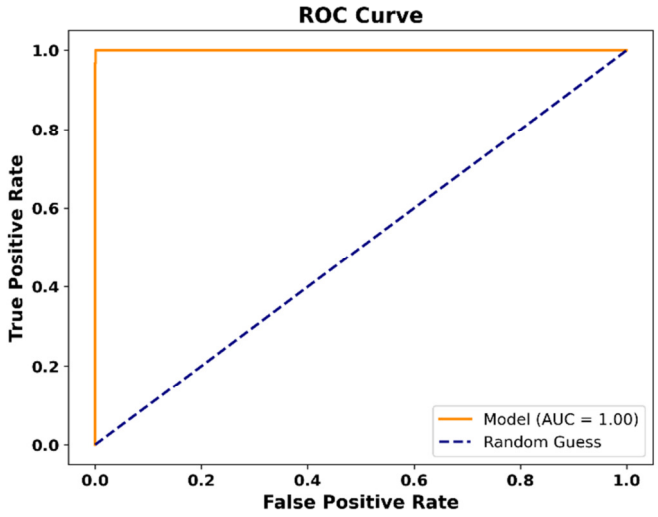

Fig. 6. AUC result of the fraud detection model.

Table II shows that the AdaBoost model was trained fastest with 0.50 s. Logistic regression took 2.29 s due to its iterative optimization. The ensemble model, which combines predictions from multiple classifiers, took 60 s, and the random forest, which builds different decision trees, took 25 s. These results suggest a trade-off between complexity and training time: simpler models, such as AdaBoost and logistic regression, are fast to compute but less accurate; more complex models, such as random forest and the ensemble model, are more accurate but computationally demanding.

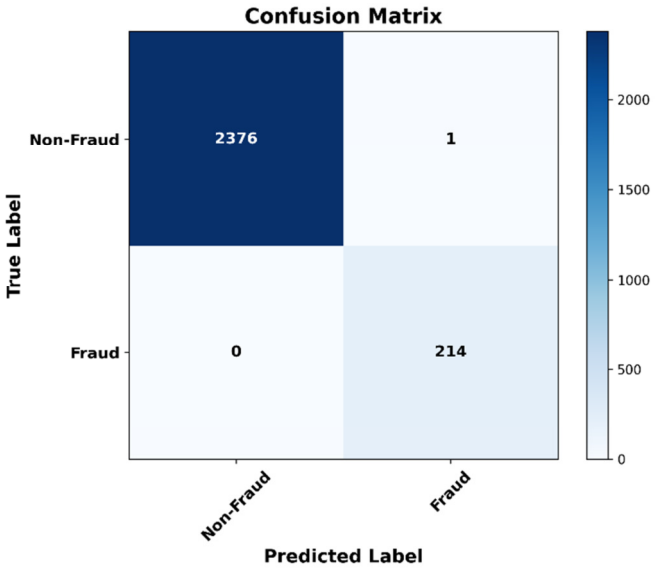

Fig. 7. Confusion matrix of the fraud detection model.

TABLE II. TRAINING TIME COMPARISON

Model	Training time (s)
AdaBoost	0.50
Logistic regression	2.29
Random forest	25
Ensemble model	60

B. Discussion of Ensemble Model Effectiveness

The results demonstrate the efficiency of integrating multiple classifiers using a soft voting strategy across all key metrics. Networked classifiers improve fraud classification accuracy, precision, and recall (key to minimizing the false positives and negatives). This method minimizes errors, by providing good recall of all fraudulent transactions while minimizing the misclassification of valid transactions as fraudulent with high precision and accuracy.

The unique advantages of each classifier are used to leverage the ensemble learning strategy, exceptionally soft voting. Random forest, AdaBoost, and logistic regression have strong baseline performance and interpretability. At the same time, logistic regression is easier to interpret. The ensemble model uses their outputs to produce a more accurate and reliable fraud detection system.

C. Application on PaySim Dataset

The research models were evaluated in experiments on the PaySim dataset [20] from Kaggle, a synthetic dataset of 6.4 million simulated online transactions across five categories. The dataset was used to test the generalizability and robustness of the ensemble model on a more complex structure. We compared the ensemble model, trained and tested on PaySim, to logistic regression, random forest, and AdaBoost in Table III to show its performance in fraud detection. Table III shows that the ensemble model achieved the highest accuracy of 99.97%. Its precision of 99.05% significantly reduced false positives, which is crucial to maintaining the integrity of fraud detection.

With a recall rate of 80.10%, it outperformed individual models in identifying fraudulent transactions. The F1 score of 85.64%, which balances precision and recall, confirmed the thorough and reliable performance of the ensemble model. The high AUC score (Figure 8) further validates the ensemble model's effectiveness in distinguishing between classes, even in large-scale datasets such as PaySim.

TABLE III. COMPARISON OF SINGLE ML ALGORITHMS VS. THE ENSEMBLE MODEL ON PAYSIM DATASET

Model	Accuracy	Precision	F1 score	Recall
Logistic regression	0.9976	0.3223	0.8000	0.4595
Random forest	0.9996	0.9905	0.7144	0.8301
AdaBoost	0.9993	0.8606	0.5193	0.6478
Ensemble model	0.9997	0.9905	0.8010	0.8564

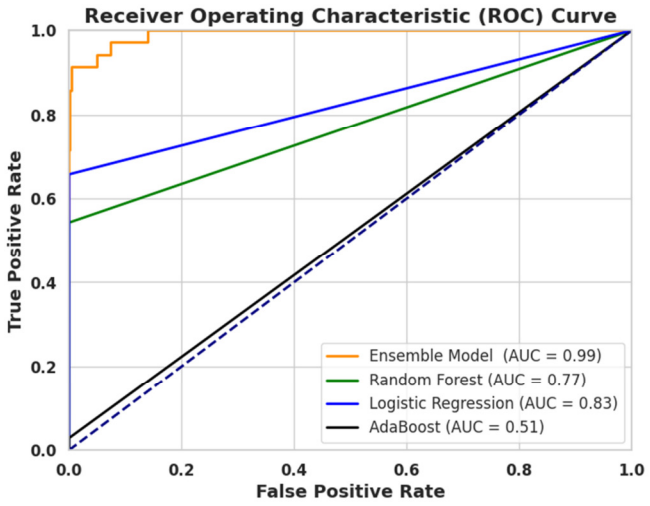

Fig. 8. AUC result on PaySim dataset.

Table IV shows the training times for logistic regression (3.39 s), AdaBoost (79.28 s), random forest (146.28 s), and the ensemble model (508.73 s), highlighting the trade-off between complexity and efficiency. Logistic regression is the fastest due to its simplicity. AdaBoost and random forest take longer due to iterative weak learner generation and decision tree construction. The ensemble model takes the longest because it combines individual classifiers and merges their outputs. While complex models like ensemble provide higher accuracy, they are less suitable for time-critical or resource-constrained applications.

TABLE IV. TRAINING TIME COMPARISON ON PAYSIM DATASET

Model	Training time (s)
AdaBoost	79.28
Logistic regression	3.39
Random forest	146.28
Ensemble model	508.73

V. CONCLUSION

This study proposes an ensemble machine learning model that combines random forest, logistic regression, and Adaptive Boosting (AdaBoost) to improve credit card fraud detection. The model was rigorously evaluated on two datasets, the customer credit dataset and the PaySim dataset, and demonstrated superior performance compared to individual classifiers. Key results include an accuracy of 99.96% on the customer credit dataset and 99.97% on the PaySim dataset, with precision and recall rates consistently outperforming standalone models. The ensemble model achieved a perfect Area Under the Curve (AUC) score of 1.0 on the customer credit dataset and a high AUC of 0.99 on the PaySim dataset, underscoring its robustness in distinguishing between fraudulent and legitimate transactions.

While the ensemble model demonstrates superior performance in terms of accuracy, precision, and recall, it comes at a higher computational cost. The training time for the ensemble model was significantly longer (60 s on the customer credit dataset and 508.73 s on the PaySim dataset) compared to individual models like logistic regression (2.29 s) and AdaBoost (0.50 s). This trade-off between accuracy and computational efficiency is an important consideration for real-time fraud detection systems, where resource constraints and processing speed are critical. Future work could explore optimization techniques, such as parallel processing or model pruning, to reduce the computational burden while maintaining high performance.

This study makes an important contribution to the field in several ways. Previous work has explored ensemble methods combining random forest with gradient boosting or combining a bagging and boosting approach, but the model of this paper offers an original combination of random forest, logistic regression, and AdaBoost. This combination improves the interpretability of the model, a property that is often discarded in the practice of ensembles. Furthermore, the proposed model is also more efficient in solving class imbalance problems than a single model solution, which cannot achieve a low false positive rate in imbalanced datasets.

The practical implications of this work are significant. Financial institutions and e-commerce platforms can deploy this ensemble model to detect fraudulent transactions in real-time, reducing financial losses and increasing customer trust. The model's ability to adapt to dynamic fraud patterns makes it particularly valuable in today's rapidly evolving digital economy. In addition, the inclusion of logistic regression provides a level of interpretability that is critical for regulatory compliance and stakeholder communication.

Future work could explore the integration of federated learning to enhance the model's scalability and privacy-preserving capabilities. Federated learning would allow multiple institutions to jointly train the model without sharing sensitive data, further enhancing its robustness and generalizability. This approach is consistent with the growing emphasis on data privacy and security in financial transactions. In summary, this study contributes to credit card fraud detection by proposing a novel ensemble model that

outperforms existing methods in terms of accuracy, precision, and recall. By addressing key challenges such as class imbalance and interpretability, this research provides a robust and practical solution for real-world fraud detection applications.

ACKNOWLEDGEMENT

The authors acknowledge and thank all parties, including the university and cybersecurity firms, for providing valuable support, knowledge, and insights for this research, and the editors and reviewers for their efforts in reviewing and providing helpful feedback for the review and publication of this paper. The authors acknowledge the grant that drives this paper under the grant number (RDIG/09/2022).

REFERENCES

[1] C. Mullen, "Card industry's fraud-fighting efforts pay off: Nilson Report," *Payments Dive*. https://www.paymentsdive.com/news/card-industry-fraud-fighting-efforts-pay-off-nilson-report-credit-debit/639675/.

[2] E. Frank and J. Oluwaseyi, "Challenges and limitations of fraud detection in NoSQL database systems." April, 2024.

[3] S. Rosset, U. Murad, E. Neumann, Y. Idan, and G. Pinkas, "Discovery of fraud rules for telecommunications—challenges and solutions," in *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, San Diego, CA, USA, 1999, pp. 409–413, https://doi.org/10.1145/312129.312303.

[4] A. A. Ojugo, A. O. Eboka, R. E. Yoro, M. O. Yerokun, and F. N. Efozia, "Framework Design for Statistical Fraud Detection," in *Mathematics and Computers in Sciences and Industry Series*, vol. 50, 2015, pp. 176–182.

[5] S. V. Suryanarayana, G. N. Balaji, and G. V. Rao, "Machine Learning Approaches for Credit Card Fraud Detection," *International Journal of Engineering and Technology*, vol. 7, no. 2, pp. 917–920, Mar. 2018, https://doi.org/10.14419/ijet.v7i2.9356.

[6] N. Ayub *et al.*, "Forecasting Multi-Level Deep Learning Autoencoder Architecture (MDLAA) for Parametric Prediction based on Convolutional Neural Networks," *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 21279–21283, Apr. 2025, https://doi.org/10.48084/etasr.9155.

[7] A. M. Fayyomi, D. Eleyan, and A. Eleyan, "A Survey Paper On Credit Card Fraud Detection Techniques," *International Journal of Scientific & Technology Research*, vol. 10, no. 09, pp. 72–79, Sep. 2021.

[8] P. Tiwari, S. Mehta, N. Sakhuja, J. Kumar, and A. K. Singh, "Credit Card Fraud Detection using Machine Learning: A Study." arXiv, Aug. 23, 2021, https://doi.org/10.48550/arXiv.2108.10005.

[9] N. K. Trivedi, S. Simaiya, U. K. Lilhore, and S. K. Sharma, "An Efficient Credit Card Fraud Detection Model Based on Machine Learning Methods," *International Journal of Advanced Science and Technology*, vol. 29, no. 05, pp. 3414–3424, Apr. 2020.

[10] S. K. Hashemi, S. L. Mirtaheri, and S. Greco, "Fraud Detection in Banking Data by Machine Learning Techniques," *IEEE Access*, vol. 11, pp. 3034–3043, 2023, https://doi.org/10.1109/ACCESS.2022.3232287.

[11] R. Sailusha, V. Gnaneswar, R. Ramesh, and G. R. Rao, "Credit Card Fraud Detection Using Machine Learning," in *2020 4th International Conference on Intelligent Computing and Control Systems*, Madurai, India, 2020, pp. 1264–1270, https://doi.org/10.1109/ICICCS48265.2020.9121114.

[12] J. Karthika and A. Senthilselvi, "Credit Card Fraud Detection based on Ensemble Machine Learning Classifiers," in *2022 3rd International Conference on Electronics and Sustainable Communication Systems*, Coimbatore, India, 2022, pp. 1604–1610, https://doi.org/10.1109/ICESC54411.2022.9885649.

[13] S. Saraf and A. Phakatkar, "Detection of Credit Card Fraud using a Hybrid Ensemble Model," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 9, pp. 464–474, Sep. 2022, https://doi.org/10.14569/IJACSA.2022.0130953.

[14] A.-A. Al-Maari and M. Abdulnabi, "Credit Card Fraud Transaction Detection Using a Hybrid Machine Learning Model," in *2023 IEEE 21st Student Conference on Research and Development*, Kuala Lumpur, Malaysia, 2023, pp. 119–123, https://doi.org/10.1109/SCOReD60679.2023.10563915.

[15] I. Sohony, R. Pratap, and U. Nambiar, "Ensemble learning for credit card fraud detection," in *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*, Goa, India, 2018, pp. 289–294, https://doi.org/10.1145/3152494.3156815.

[16] M. Devikar, A. Khadke, A. Lad, R. Sapkal, and S. Nikalje, "Credit Card Fraud Detection using Ensemble Learning," *International Research Journal of Engineering and Technology*, vol. 07, no. 05, pp. 7402–7406, May 2020.

[17] "Customer Credit Card Dataset." Kaggle, [Online]. Available: https://www.kaggle.com/datasets/parnianmalekian/customer-dataset.

[18] M. Jadhav Shradha, V. Chalwa Prabhavati, and S. Bakshi Asmita, "Online Credit Card Fraud Detection System," *International Research Journal of Engineering and Technology*, vol. 07, no. 07, pp. 573–577, Jul. 2020.

[19] B. Ravinder Reddy, N. Rajesh, K. V. Anand, and G. Srikanth, "Fraud Transaction Detection Approach Using Machine Learning Hybrid Techniques," *International Journal of Scientific Research in Science and Technology*, vol. 10, no. 2, pp. 90–96, Apr. 2023, https://doi.org/10.32628/IJSRST52310213.

[20] "Synthetic Financial Datasets For Fraud Detection." Kaggle, [Online]. Available: https://www.kaggle.com/datasets/ealaxi/paysim1.

[21] P. Venkataanusha, Ch. Anuradha, P. S. R. Chandra Murty, and C. Surya Kiran, "Detecting Outliers in High Dimensional Data Sets Using Z-Score Methodology," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 1, pp. 48–53, Nov. 2019, https://doi.org/10.35940/ijitee.A3910.119119.

[22] J. Brownlee, "How to Handle Missing Data with Python," *MachineLearningMastery.com*, Mar. 19, 2017. https://www.machinelearningmastery.com/handle-missing-data-python/.

[23] P. Cerda, G. Varoquaux, and B. Kégl, "Similarity encoding for learning with dirty categorical variables," *Machine Learning*, vol. 107, no. 8–10, pp. 1477–1494, Sep. 2018, https://doi.org/10.1007/s10994-018-5724-2.

[24] M. U. Salur and İ. Aydın, "A soft voting ensemble learning-based approach for multimodal sentiment analysis," *Neural Computing and Applications*, vol. 34, no. 21, pp. 18391–18406, Nov. 2022, https://doi.org/10.1007/s00521-022-07451-7.

[25] B. Devi Meenaksh, B. Janani, S. Gayathri, and N. Indira, "Credit Card Fraud Detection using Random Forest," *International Research Journal of Engineering and Technology*, vol. 06, no. 03, pp. 6662–6666, Mar. 2019.

[26] Y. Li, M. Fauß, and A. M. Zoubir, "Logistic Regression with Robust Bootstrapping," in *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Le Gosier, Guadeloupe, 2019, pp. 346–350, https://doi.org/10.1109/CAMSAP45676.2019.9022480.

[27] Y. Zhang *et al.*, "Research and Application of AdaBoost Algorithm Based on SVM," in *2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference*, Chongqing, China, 2019, pp. 662–666, https://doi.org/10.1109/ITAIC.2019.8785556.

[28] D. J. Hand, P. Christen, and N. Kirielle, "F*: an interpretable transformation of the F-measure," *Machine Learning*, vol. 110, no. 3, pp. 451–456, Mar. 2021, https://doi.org/10.1007/s10994-021-05964-1.

[29] R. Draelos, "Measuring Performance: The Confusion Matrix," *Glass Box*, Feb. 17, 2019. https://glassboxmedicine.com/2019/02/17/measuring-performance-the-confusion-matrix/.

[30] S. Hussain *et al.*, "An Enhanced Random Forest (ERF)-based Machine Learning Framework for Resampling, Prediction, and Classification of Mobile Applications using Textual Features," *Engineering, Technology & Applied Science Research*, vol. 15, no. 1, pp. 19776–19781, Feb. 2025, https://doi.org/10.48084/etasr.9148.