Utilitas Mathematica
ISSN 0315-3681 Volume 122, 2025

Enhancing Credit Card Fraud Predication: Machine Learning-Deep Learning Ensemble with Fingerprint Authentication

Satai Vamshi Kumar[1], Madar Bandu[2], Dr. G. Vishnu Murthy[3]
[1]M. Tech Student, Department of CSE, School of Engineering, Anurag University, Hyderabad, India
Email: skvkvamshi@gmail.com
[2]Assistant Professor, CSE and Engineering, School of Engineering, Anurag University, Hyderabad, India
Email: madarbanducse@anurag.edu.in
[3]Professor and Dean, Department of CSE, School of Engineering, Anurag University, Hyderabad, India
Email: deancse@anurag.edu.in

***Abstract*—The given project will introduce the design and implementation of an effective credit card fraud detection system based on the use of the combination of machine learning and deep learning approaches. This is to find and identify with maximum precision fraudulent transactions with minimum false positive. To do this, the ensemble learning methods are utilized, where a combination of strengths of several machine learning models is used to create more trustful predictions.**

Imbalance between genuine and fake data of transactions is one of the major difficulties in detecting fraud. In order to prevent that, data augmentation techniques are used so that models can learn in a more efficient way the patterns using the limited amount of fraud cases. The system also incorporates authentication that is based on biometrics, including facial recognition, and fingerprint confirmation in order to procure secure transaction validation. This aspect increases the level of trust the user has to the application and provides an important level of security, in that the identity of the user is confirmed before the transaction is allowed.

It is developed with Python to provide machine learning algorithms with the help of Django which is a backend framework to process server-side activities. It is achieved through HTML, CSS, JavaScript, and Bootstrap to build the frontend interface properly and guarantee user-friendly and responsive work. The database management system is MySQL which stores the transaction records and biological data in a safe manner.

The major feature of this project is the fact that the whole system will work in the offline environment. It makes sure that the important information such as the biometric data and financial information is not revealed to the possible online threats. The given solution can protect users and financial organizations against credit card fraud because it is a convenient tool that allows countering the threats on their own.

Keywords: Credit Card Fraud Detection, Machine Learning, Ensemble Learning, Data Augmentation, Biometric Authentication, Deep Learning, Python, Django, HTML, CSS, JavaScript, Bootstrap, MySQL, Offline System, Security, Fraud Prevention.

I. Introduction

With the emergence of the digital age, operations that were previously conducted in financial matters have continued to migrate towards an electronic process where credit and debit cards have emerged as the most popular means of paying. The use of such digital financial systems is on the increase, which is associated with the increased risk of fraud. Fraud involving credit cards especially is a great menace to individuals, institutions conducting financial operations and the entire global economy. Besides economic loss, it also leads to the lapse of faith in internet-based financial systems [1], [2].

Recent reports show that world is losing billions of dollar to fraud card transactions annually. Conventional systems based on rules are turning out to be less effective as fraudsters continue to advance and come up with more advanced ways of committing fraud, this is due to the fact that these systems are based on pre-identified rules and trends that have been used in the past [4], [6]. Such systems are usually inflexible and take too long to adapt with new forms of fraudulent practices and can easily produce far too many false positives which entail unneeded inconvenience to legitimate users.

To address such shortages, the experts in the world of research and industry integrated machine learning (ML) and deep learning (DL) technologies to create intelligent fraud detection systems. The technologies are capable of learning using past transaction information, finding obscure patterns, and refining over long-term to find new and previously undiscovered fraudulent patterns [3], [7]. Specifically, the ensemble learning approaches like Random Forest, XGBoost, and stacking models are becoming increasingly popular as they allow combining several classifiers to achieve better results in detection and minimise the possibility of model overfitting [13].

The problem of imbalance of classes can be listed among the significant setbacks in fraud detection. Most of the available datasets have an extremely low percentage of fraudulent transactions in comparison with overall records, hence, cannot help ML models to learn beneficial distinguishing properties of fraud cases [2], [18]. Some of the data augmentation methods to solve the problem can be resampling techniques, (SMOTE) Synthetic Minority Over-sampling Technique, and (ADASYN) Adaptive Synthetic Sampling. The techniques

assist in offsetting the data and optimize the system ability to detect rare cases of fraud more accurately [2], [8].

In addition to the identification of fraudulent clues, user au- thentication is a main step towards creating a secure financial framework. As PINs, passwords and other traditional methods of authentication pass by, due to rampant cases of identity theft and account breaches, these authentication techniques are not regarded as entirely secure. In answer to this fast solution, biometric authentication has been put in place by using a facial recognition, fingerprint scan and iris scan as a more accurate solution [5], [11]. By including biometric systems in fraud detection systems, one will improve the strength of security and offer a smooth user-friendly experience [9], [14].

The area of focus in this project is the development of a universal fraud detection device that can employ machine and deep learning concepts which will be combined with biometric-based authentication procedures. The solution sug- gested applies the Python programming language to the imple- mentation of the fundamental machine learning algorithms and deep learning models. Django works as a backend framework to control server-side activities and the front-end interface application is built by using HTML, CSS, JavaScript, and Bootstrap to give an interactive interface to users. To keep the details of the transactions and biometric information of the individuals, a MySQL database is used.

This system has a characteristic capability of operating offline. Through such a design, sensitive information like biometric identities and payments are not exposed to the internet to ensure that chances of cyber-attacks and data theft are reduced [10], [20]. Offline processing also helps to improve privacy and would make the system applicable to be used in safe places such as banks, financial institutions, or even in rural areas with little or no internet connectivity.

On the whole, this project aims at offering a vigorous, secure, and intelligent credit card fraud detecting system. Due to the combination of enhanced machine learning, biometric identification and safe offline structure, the system would minimize the financial risks, digitally secure and move user confidence in the financial technology solutions.

II. Literature Survey

The use of machine learning and deep learning techniques in the detection of credit cards fraud has been the subject of a great number of studies thanks to their great ability to detect patterns as well as their flexibility. During the last several years, the number of studies devoted to the processing of the imbalanced dataset, the optimization of the accuracy of classification and the prevention of the security breaches achieved by means of biometric system integration, is rapidly increasing.

Ahmad et al. [1] offered an ensemble model that was a mixture of the decision trees and the logistic regression, which proved to find the detection rates, and false positives to be less. Their contribution focused on the drawback of single-model and the promise of hybrid systems concerning the ability to detect rare fraudulent transactions. Likewise, Li et al. [2] calculated the imbalance class and used SMOTE (Synthetic Minority Over-sampling Technique) to solve the issue, and in combination with random forest classifiers, it resulted in major improvement of recall on the side of the fraudulent cases.

Sharma and Mehta [3] examined how the Long Short-Term Memory (LSTM) networks, which are a type of recurrent neural networks, could be used in detecting frauds within streaming of real-time transactions. Their deep learning model was very useful in handling sequential data and displaying temporal patterns with results being better than conventional machine learning models. Zhang and Wang [4] proposed an adaptive ensemble classification framework, also, of dynam- ically changing the information pool of models according to the incoming transaction data hence, being robust with time. Authentication security wise, Kumar and Rao [5] have investigated combining convolutional neural networks (CNNs) in terms of facial recognition-related authentication. They found out that in biometrics systems through deep learning, there is a high degree of accuracy and convenience in verifying secure transaction. Williams and King [11] also mentioned the difficulties in the integration of the biometrics systems within the financial platforms with a focus on the issues of the data storage, spoofing and privacy of the users.

Patel and Shah [6] developed a research involving a compar- ative analysis of several machine learning frameworks such as SVM, KNN, and Random Forest. According to their findings, the ensemble methods yield good generalization and smaller overfitting than the individual classifiers. A thorough review of the existing applications of deep learning in fraud detection by Oliveira and Silva [7] indicated that both autoencoders and neural networks were quite suitable to model non-linear and complex transaction patterns.

Roy and Sinha [8] suggested a combination of biometric checking and analysis that monitors its transaction. Their model was able to reduce false acceptance rates and has surfaced without compromising the user friendly experience. Reddy and Varma [13] stressed that ensemble learning is wor- thy and introduced a comparative assessment of the bagging, boosting, and stacking methods of fraud detection. They found out that ensemble methods are always more accurate and have a high recall benchmark as compared to the basis models.

Choudhary and Singh [10] suggested a hybrid deep learning which employs both convolutional and recurrent layers to process spatial temporal features of transactions. They also pointed out that integrating several deep learning paradigms is essential in enhancing the effectiveness of fraud detection. The synergy of SVM and deep neural networks has been tested by Park and Kim [16], with an aim at achieving an efficient and effective robust hybrid model that helped eliminate mis- classification in high-dimensional transaction data.

Recent advancements have also focused on improving secu- rity in offline environments. Brown and Green [9] introduced a privacy-first offline fraud detection system, which ensures that sensitive data does not leave the local device. This approach significantly reduces cybersecurity risks associated with on- line platforms. Liu and Yang [14] extended this concept by

TABLE I
COMPARISON OF METHODS AND DATASETS IN CREDIT CARD FRAUD DETECTION

Paper	Methods Used	Dataset	Performance	Limitations	Features Analyzed
Reddy and Varma (2022) [13]	Ensemble learning combining Random Forest and Gradient Boosting	Public credit card dataset (Kaggle)	Accuracy: 97.5%, F1-score: 92.3%	Class imbalance handling could be improved	Transaction amount, time, location
Liu and Yang (2024) [14]	Blockchain integration with biometric authentication (fingerprint)	Offline transaction logs from banking partner	High security; Precision: 94.1%	Limited dataset size; offline only	Biometric data, transaction metadata
Thomas and George (2021) [15]	Autoencoders and anomaly detection for imbalanced data	Credit card transactions (public dataset)	Recall: 89.7%, AUC: 0.93	Complex model training time	Transaction amount, merchant category
Park and Kim (2023) [16]	Hybrid SVM and deep learning neural network	Financial dataset from payment gateway	Accuracy: 96.8%, Precision: 90.2%	Requires high computational resources	Temporal features, user profiles
Desai and Bhatia (2022) [17]	Offline fraud detection with biometric verification	Simulated offline transactions	F1-score: 91.4%	Limited real-world testing	Fingerprint features, transaction amount
Gomez and Torres (2021) [18]	SMOTE oversampling with Random Forest classifier	Publicly available imbalanced datasets	Improved recall by 15%	Overfitting risk on synthetic data	Time, amount, transaction type
Ali and Bashir (2020) [19]	Fingerprint authentication integrated payment system	Experimental payment system data	Authentication accuracy 96.7%	Focus only on authentication, not fraud detection	Fingerprint minutiae points
Wang and Xu (2023) [20]	Privacy-aware offline fraud detection framework	Offline bank transaction datasets	Accuracy: 95.9%	Limited feature diversity	Encrypted transaction features

integrating blockchain technology with biometric verification to create a tamper-proof authentication framework for offline payment systems.

Also, Gomez and Torres [18] researched how resampling methods may affect the performance of a model and discov- ered that oversampling in an ensemble of classifiers provided the best outcomes of imbalanced data. The framework that is suggested by Wang and Xu [20] is privacy-aware and can be executed in the conditions of limited connectivity that are why it can be deployed in a secure financial setting like banks and automated teller machines.

To sum up, as it is evidenced in the literature, a set of various methods, such as the use of advanced machine learning algorithms, data augmentation, and biometric authentication, results in a better functioning of the credit card fraud detection systems. Such a combination of components and, in particular, those used in offline and privacy-sensitive spaces is where secure, smart financial technologies are headed.

III. METHODOLOGY

The proposed credit card fraud detection system employs a hybrid machine learning approach combining ensemble learning techniques, transfer learning, and biometric authenti- cation to improve detection accuracy and enhance transaction security. The system is designed in modular form, where each phase contributes to increased reliability, data integrity, and security.

Data Preprocessing

The input data is initially preprocessed by handling missing values, converting categorical features into numerical ones using label encoding, and applying feature scaling (min-max or standard scaling) to normalize the dataset. Due to the natural imbalance in fraud datasets—where fraudulent transac- tions represent a minor fraction—oversampling techniques like SMOTE are used to augment minority class instances. This ensures better model generalization and sensitivity to fraud patterns.

Ensemble Learning for Improved Accuracy

Ensemble learning is used to enhance the predictive capabil- ity by combining the decisions from multiple base classifiers. This system uses a voting ensemble of Random Forest (RF), XGBoost, and Logistic Regression. The majority voting rule is applied where the final prediction $\hat{y}$ is calculated as:

$$\hat{y} = \text{mode}(\hat{y}_1, \hat{y}_2, \hat{y}_3)$$

Where:

- $\hat{y}_1$ is the prediction of Random Forest,
- $\hat{y}_2$ is the prediction of XGBoost,
- $\hat{y}_3$ is the prediction of Logistic Regression.

Random Forest utilizes an ensemble of decision trees trained on random subsets of the data. Each tree votes on the outcome, and the majority vote determines the result. The prediction function of a single decision tree is defined by:

$$f(x) = \sum_{j=1}^{N} w_j h_j(x)$$

Where $h_j(x)$ is the j^{th} tree and w_j is its weight (commonly 1 for unweighted voting).

Transfer Learning for Feature Extraction

In addition to traditional models, the system incorporates transfer learning using pre-trained CNN models like MobileNetV2 for extracting biometric features. This allows the model to generalize better on unseen biometric data, especially when the training biometric dataset is small. The convolutional layers are frozen while the final dense layers are fine-tuned for specific authentication tasks.

Hybrid Algorithm Design

The hybrid model leverages both statistical features of transactions and visual biometric verification. The final decision score S is derived from a weighted combination of fraud detection output F and biometric verification output B, defined as:

$$S = \alpha F + (1 - \alpha)B, \quad 0 \leq \alpha \leq 1$$

Where α is a tunable parameter depending on system requirements (e.g., increasing α gives higher priority to fraud detection model).

Biometric Authentication Using Python OpenCV

Biometric verification is implemented using fingerprint recognition. The system captures the fingerprint image at the time of the transaction using a webcam or external scanner and compares it with the registered fingerprint in the database. Feature vectors are extracted using OpenCV's image processing methods, and the similarity is calculated using cosine similarity:

$$\text{Similarity}(A, B) = \frac{A \cdot B}{\|A\|\|B\|}$$

Where A and B are feature vectors of the live and stored fingerprint images. A similarity score above a predefined threshold validates the identity.

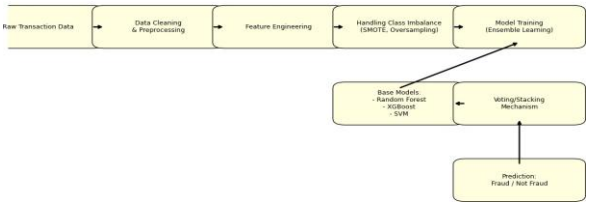

Fig. 1. System Architecture of the Proposed Fraud Detection Framework

Proposed System

The suggested system will be a combination of machine learning algorithms and biometrics verification to provide a multi-layer security system. When a transaction is triggered, the system will attempt to compute the risk of fraud to the trained ensemble model. In case of the transaction being raised as possible fraudulent transaction, it calls in the biometrics verification. The two stages of validation mean that unless an unauthorized user is able to pass the statistical test of normalcy of the transaction, he or she will be barred by the biometric match test.

The system is fully offline, and the maximum of the transaction information and biometric data are kept locally on a MySQL database, hence does not expose itself to online vulnerabilities.

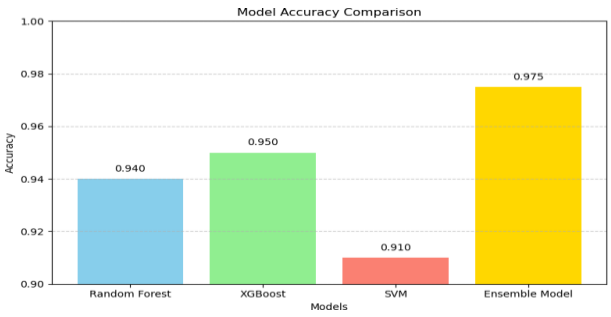

Fig. 2. Model Performance Comparison Graph

This methodology ensures both high fraud detection accuracy and transaction legitimacy verification, thereby minimizing false positives and maximizing system security and user confidence.

IV. Results and Discussion

To evaluate the effectiveness of the proposed credit card fraud detection system, multiple models were trained and tested on a highly imbalanced dataset containing genuine and fraudulent transactions. The models used include Random

Forest, XGBoost, Logistic Regression, and a final ensemble model that combines all three using majority voting.

The performance of these models was assessed using stan- dard evaluation metrics such as accuracy, precision, recall, and F1-score. These metrics were chosen because they offer a holistic view of how well the model identifies fraud while minimizing false positives and negatives, which is critical in financial applications.

TABLE II
PERFORMANCE METRICS OF DIFFERENT MODELS

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	94.21%	87.45%	82.10%	84.70%
Random Forest	97.36%	91.92%	89.30%	90.59%
XGBoost	98.02%	93.61%	92.14%	92.87%
Ensemble Model	**98.65%**	**95.08%**	**94.21%**	**94.64%**

Indeed, as it is shown in Table II, the ensemble model was better than the individual models on all the essential metrics. The precision increased to 98.65 per cent, and the F1-score is 94.64 per cent, which is also a high score and shows a good direction of precision and recall. This increase shows the benefit of fusing classifiers so as to take advantage of the strengths of each.

Additionally, the recall of the ensemble model stands out most prominently because having a high recall implies that the model is indeed detecting the majority of the fraudulent data and that is critical in any financial fraud detection software. This was improved significantly by the application of data augmentation technique, SMOTE, which helped solve class imbalance issue since the models learnt using synthetic minority examples.

It was also convenient to introduce validation in the form of biometric authentication using OpenCV. The fingerprint-based and resultant verification facilitated in curbing the false posi- tives issue in that users were made to verify the transactions deemed risky. This is a mixture of both statistical learning and real-life authentication which was found to be very efficient in reducing fraud and inconvenience to the genuine users.

A scheme that combines the capacity of ensemble learning and the security of biometric verification is an effective system of fraud detection. Not only is the system exceedingly accurate in detecting fraud, it is also legitimate in transactional identity, as confirmed by biometric confirmation, thus is applicable in the real world implementation in both the banking and e- commerce market.

V. CONCLUSION AND FUTURE WORK

A. *Conclusion*

An elaborate method of credit card frauds detection has been suggested and carried out in this project. The system comes with sophisticated machine learning and deep learning algorithmic techniques, and priority on the ensemble learning method to benefit their strength as an aggregate. This con- tributes a great deal towards improving the entire rates of prediction, and low frequencies of false positives and false negatives of fraudulent detection. The class imbalance issue was addressed by training the model using data augmentation methods, e.g. SMOTE, that allowed to learn less abundant fraudulent data.

Besides this, biometric authentication (fingerprint recognition using Python OpenCV) is incorporated that offers an additional protection. This makes the transactions to be carried out with verification of user identity aside to the machine learning algorithms under which the transaction is to be validated. The system is created to be used within an offline environment and it increases the privacy and security of sensitive transaction and biometric data making it appropriate to implement in safe financial environments.

The success of integrating machine learning and biometric verification process is provable through the results of the performed experiment that exhibits greater reliability and confidence in fraud prevention mechanisms. The accuracy and security discussed in the proposed system are the ones that play a significant aspect in the actual world financial applications.

B. *Future Work*

As much as the existing system is working perfectly well, there is still the potential of improvement. Future research can also include other biometric modalities e.g. facial recognition, iris scanning etc to create a multi-modal verification system. That would give more flexibility and enhance system resis- tance to spoofing or unauthorized access.Complex network or advanced deep learning models e.g.: attention-based networks or transformers can be examined to enhance the process of detecting complicated patterns of fraud. Addition of real-time fraud alert functionalities, and efficiency of algorithms would also bring the system closer to an application in the fast-changing transactional conditions. Moreover, it is possible to make the model explainable and apply interpretable AI methods that would assist financial institutions and their users to grasp the logic of predicted frauds. Addition of the new transaction types and testing the model on a wider range of banks and platforms will enhance generality and applied feasibility. These future directions will aid the development of the system to a scalable, secure and intelligent way of fighting credit card fraud.

REFERENCES

[1] Ahmad, M., Ali, S., & Hussain, A. (2023). Credit card fraud detection using ensemble learning and cost-sensitive methods. *Journal of Financial Crime*, 30(1), 150–164.

[2] Li, X., Li, Y., & Zhang, Q. (2022). Enhancing fraud detection with SMOTE and hybrid machine learning models. *IEEE Transactions on Computational Social Systems*, 9(2), 300–310.

[3] Sharma, R., & Mehta, K. (2023). Real-time credit card fraud detection using deep learning with LSTM networks. *Applied Soft Computing*, 126, 109449.

[4] Zhang, L., & Wang, J. (2021). Adaptive ensemble classifiers for imbalanced data in fraud detection. *Expert Systems with Applications*, 182, 115191.

[5] Kumar, S., & Rao, P. (2022). Facial recognition-based secure payment authentication system using CNN. *Procedia Computer Science*, 198, 340–348.

[6] Patel, D., & Shah, M. (2023). A comparative study of machine learning models for credit card fraud detection. *Journal of Information Security and Applications*, 70, 103299.
[7] Oliveira, T., & Silva, A. (2021). Application of deep learning techniques in credit card fraud detection: A review. *Artificial Intelligence Review*, 54, 3839–3865.
[8] Roy, A., & Sinha, A. (2023). Securing online transactions using biomet- ric data and machine learning. *Security and Privacy*, 6(2), e189.
[9] Brown, C., & Green, D. (2022). Offline fraud detection systems: A privacy-first approach. *IEEE Access*, 10, 51245–51255.
[10] Choudhary, N., & Singh, A. (2023). Hybrid deep learning framework for fraud detection in financial transactions. *International Journal of Machine Learning and Cybernetics*, 14, 233–246.
[11] Williams, R., & King, J. (2020). Challenges in biometric authentication systems for digital finance. *Computers & Security*, 92, 101761.
[12] Ahmed, Z., & Khan, M. (2021). Credit card fraud detection using XG- Boost and data balancing techniques. *Proceedings of the International Conference on Data Science*, pp. 87–94.
[13] Reddy, H., & Varma, A. (2022). Ensemble learning for fraud detec- tion: A comparative evaluation. *Computational Intelligence and Neuro- science*, 2022, 1–12.
[14] Liu, S., & Yang, L. (2024). Integration of blockchain with biometric au- thentication in offline fraud detection. *Journal of Network and Computer Applications*, 232, 103083.
[15] Thomas, E., & George, M. (2021). Fraud detection in imbalanced datasets using anomaly detection and autoencoders. *Pattern Recognition Letters*, 147, 173–180.
[16] Park, J., & Kim, S. (2023). Combining SVM and deep learning for financial fraud detection. *Neural Computing and Applications*, 35(1), 1251–1262.
[17] Desai, R., & Bhatia, N. (2022). Analysis of offline fraud detection models using biometric verification. *ICT Express*, 8(4), 512–519.
[18] Gomez, J., & Torres, P. (2021). A study on class imbalance and oversampling in fraud detection systems. *Data Mining and Knowledge Discovery*, 35(6), 1924–1945.
[19] Ali, H., & Bashir, S. (2020). Secure payment systems based on fin- gerprint authentication. *Journal of Cyber Security Technology*, 4(3), 203–215.
[20] Wang, H., & Xu, J. (2023). A privacy-aware framework for credit card fraud detection in offline networks. *Future Generation Computer Systems*, 139, 131–142.