

Original Article

Analysis of Machine Learning Methods Performance for Fraud Recognition in Payroll ERP Systems

Rashmika Jalagam¹, Sridevi Allu², Venkata Phaneendra Reddi³, Pradeep Rao Vennamaneni⁴, Sri Prudhvi Adusumilli⁵, Khaja Moiezuddin Mohammed⁶

¹Matilda Cloud - Cloud Engineer/Python Developer, USA.

²American Airlines, AI/ML Engineer, USA.

³Atlantic Health systems - Oracle Integrations Developer, USA.

⁴Medtronic - Senior Software Engineer, USA.

⁵Charter Communications, Sr Software Developer, USA.

⁶Iqvia - Sr Data Engineer, USA.

Received: 08-05-2022

Revised: 08-28-2022

Accepted: 09-05-2022

Published: 09-07-2022

ABSTRACT

Financial institutions are increasingly worried about online payment fraud, which has increased the need for efficient and effective detection systems. In response, this work focuses on an innovative method for detecting online payment fraud by leveraging Random Forest (RF) and XGBoost (XGB) classifiers on the PaySim synthetic dataset. The methodology comprises the preprocessing of data, outlier identification, feature extraction, z-score normalization, and class balancing as a solution to severe imbalance between classes using the approach of Random Over-Sampling. Training of the models was done with the best hyperparameters and the evaluation was done in terms of accuracy, precision, recall, F1-score and ROC-AUC. The experimental findings prove that RF has almost perfect performance with 99.99% accuracy, 99.87% precision, 99.98% recall and 0.9998 ROC whereas XGB performs well too. The high effectiveness of proposed models is proved by the comparative analysis with traditional and DL models, such as Decision Tree (DT), Logistic Regression (LR), CNN, LSTM, BERT, etc. The analysis of feature importance and SHAP indicates that the account balance discrepancies and transaction amounts are the most prominent indicators of fraud. The main features such as balance gap and transaction volumes are the focus of detecting fraud, and hence the models are very precise, dependable, and can work on unbalanced time-sensitive financial data.

KEYWORDS

Fraud Detection, Payroll ERP Systems, Machine Learning, Financial Fraud, Random Forest, XGBoost, PaySim Data.

1. INTRODUCTION

Financial fraud is becoming more of a threat to organisations in today's fast-paced digital world [1]. Payroll systems are fundamental to their operations and one of the most important areas where fraud may have serious consequences. Payroll fraud is a severe issue that can manifest in a variety of ways, including the falsification of hours, the inflation of salaries, and the granting of unauthorised incentives. One kind of payroll fraud is the practice of using "ghost employees," or people who do not really exist but are recorded on the payroll. The organization's confidence might be eroded, its brand can be damaged, and long-term harm can result from these fraudulent operations beyond merely financial losses [2][3]. The increasingly complex strategies used by fraudsters make it impossible to identify and prevent payroll fraud using the conventional approaches, which often depend on human audits, rule-based systems, or recurring inspections [4][5]. As organisations develop and payroll data gets more complicated, it becomes more difficult to efficiently monitor and uncover fraudulent activity using traditional approaches. Fraudsters often discover methods to change payroll data unnoticed, exploiting flaws in outdated systems that are incapable of keeping up with the needs of current payroll operations.

Payroll fraud detection systems that are powered by AI are engaged. Through the utilisation of AI and ML, these sophisticated systems provide a revolutionary solution for organisations seeking to protect their financial operations [6]. A large amount of payroll data may be swiftly and effectively processed by AI and ML algorithms, which can spot trends and outliers that conventional systems miss [7][8]. AI-enabled systems can identify suspicious transactions automatically, eliminating the need for human interaction and continual manual inspection, unlike conventional fraud detection approaches. Payroll functions within a modern-day enterprise are being progressively handled as part of an integrated Payroll Enterprise Resource Planning (ERP) systems integration that collects employee, attendance, compensation and compliance data into one platform [9]. Though the payroll ERP systems enhance effectiveness and scalability, the system increases the attack surface to fraud and thus complex analytical detection systems are required. Thus, it is important to consider performance of various machine learning approaches in payroll ERP settings to determine models that can be accurate and implemented in practice.

This research is inspired by the fact that payroll ERP systems have become increasingly susceptible to more complex fraudulent activities that can no longer be contained using conventional rule-based or manual detection techniques. The sheer volumes and complexity of the payroll data are creating an urgent requirement for smart, automated, and precise fraud detection solutions. Machine learning will provide an opportunity to determine the hidden patterns and anomalies in payroll transactions, which will encourage a systematic performance analysis of ML models to identify the valid, efficient, and real-time deployable solutions to securing payroll ERP systems. Some of the contributions made in studies are:

- Developed highly accurate and reliable fraud detection models using RF and XGBoost (XGB) for real-time financial transactions.
- Applied feature importance and SHAP analysis to identify critical predictors such as balance discrepancies and transaction amounts, enhancing model interpretability.

- Addressed extreme class imbalance in the PaySim dataset using Random Over Sampling, improving detection of minority class (fraudulent transactions).
- Validated models in a controlled, deployable environment, demonstrating practical applicability for real-time financial fraud detection systems.
- Provided comparative analysis with existing methods, highlighting the superiority and robustness of the proposed models in handling imbalanced datasets.

The novelty of the work lies in the high quality of fraud detection when using Random Forest and XGBoost, which is higher than the quality of the traditional models (DT, LR) and deep learning (CNN, LSTM, BERT). In contrast to the past literature, it is not only highly accurate, precise, and recall, but also can give interpretable insights, including feature importance and SHAP analysis, which can be used to identify important predictors, including balance differences and transaction values. It proves to be a strong, dependable, and viable remedy to real-time financial fraud detection, which has crucial loopholes in performance and explain ability in the available methods.

1.1. Organization of the Paper

The paper has the following structure: Section II provides a review of relevant work, Section III addresses methods, Section IV examines findings and comparisons, and Section V concludes the research with future directions.

2. LITERATURE REVIEW

Several significant research studies on payroll fraud detection and prevention were reviewed and analysed to guide and strengthen the development of this work. Yapp and Yeh (2026) use the modified Friedman test to find out whether the approaches' performance differs significantly across all datasets for a certain assessment measure. Subsequently, a score is provided to every method based on the crucial difference, which is the minimum difference in average rankings needed for performance to be considered significant, and the Nemenyi post-hoc test. To determine the method that performs best overall, the scores are averaged across all evaluation criteria. XGBoost and RF outperformed all evaluation criteria, followed by AutoKeras. However, when it comes to the business-imposed constraint of 5% FPR, the most sophisticated convolutional spiking neural networks (CSNN) obtained the best ratio of young and old customers' FPRs, and remember [10].

Cedric and Ould-Slimane (2025) generated additional synthetic data by employing two autoencoders. They used five techniques to test our models' performance on both datasets at this point, with RF and XGBoost proving to be the most efficient. The final stage involved the implementation of an adversarial approach, which resulted in an average reduction of 15.77% in the recal performance measure for the first dataset and 20.22% for the second. In reaction, we used an adversarial training technique, which resulted in an average recall score improvement of 14.02% on the first dataset and 17.45% on the second dataset [11].

Tayebi and El Kafhali (2025) suggest improving the XGBoost algorithm to identify fraudulent transactions by adjusting the algorithm's hyperparameters using Bayesian optimisation, a smart

method. To test the effectiveness of our technique, various experiments are run on two credit card datasets including of both genuine and fraudulent transactions. They used random undersampling, SMOTE, and cross-validation to avoid overfitting on unbalanced datasets. The optimal results for Data 1 while utilising SMOTE were as follows: acc 0.9996, prec 0.9406, rec 0.8740, Fmeasure 0.8740, and AUC 0.9879. With an acc of 0.8325, prec of 0.8294, rec of 0.8378, Fmeasure of 0.8336, and AUC of 0.9088, the Random Under-sampling strategy produced the best results for Data 2. These experimental results demonstrate that our recommended strategy outperforms other ML techniques [12].

Jain *et al.* (2024) provide ML models to detect fraudulent financial transactions using the PaySim dataset. Label encoding, min-max scaling, and SMOTE-based balancing are utilized to overcome the intrinsic class imbalance in fraud detection tasks. Classification models like SVM, NB, RF, and LR are assessed using F1score, recall, accuracy, and prec. The outcomes indicate that Logistic Regression excels in accuracy with 98.99%, whereas Random Forest excels in recall and Naïve Bayes excels in precision. The study emphasizes the need for strong fraud detection systems and the trade-offs involved in optimizing evaluation measures to combat evolving financial crimes [13].

Tang and Liang (2024) suggest using convolutional feedforward generative (CFG) networks as the mechanism for graph expansion. This approach strengthens links between financial institution data by extending the transaction graph and federating the training of a CFNN and a GAN. It also suggests a cosine similarity-based adaptive model aggregation approach to improve efficiency post-FL model aggregation. Finally, the simulation results demonstrate that our proposed model achieved better recall and AUC than the baseline model across all test datasets. Our suggested model outperformed the baseline on the IEEE-CISdataset, which shows an improvement in recall of 11.87% to 33.9% and an increase in AUC of almost 3% [14].

Arslan and Güneş (2023) build a database of malicious and safe JavaScript scripts from a secret repository of web-based code. The contents of the program and its control flow diagram are reviewed first. As a second point, the OCSVM method is used to find outliers, and CNN is used as a classifier. Results. The results of the experiments demonstrate that the suggested OCSVM-CNN method identifies harmful scripts with a higher accuracy of 96.78%, as opposed to using only CNN (94.8% accuracy). Conclusion. The study's contribution to the creation of a multi-layered ERP software architecture with AI decision assistance will improve ERP systems' ability to identify fraud [15].

There are still numerous unanswered concerns in the field of fraud detection, despite the fact that the current literature suggests that ML and DL models are performing well. Most works are more concerned with high accuracy and recall with the help of complex models like XGBoost, RF, and deep neural networks, but provide little attention to interpretability, robustness, and real-time applicability in payroll ERP systems. Moreover, certain methods use synthetic/credit card data, as opposed to payroll-specific data, which makes them less practical. The problem of model stability on adversarial attacks, individual employee group fairness, and low false positives under business implications has not been sufficiently investigated yet, which is why efficient and explainable payroll fraud detection solutions are needed.

3. METHODOLOGY

The methodology section includes pre-processing and balancing the PaySim dataset, training XGBoost and Random Forest models, and evaluating performance using standard fraud-detection metrics. Fig 1 presents the proposed flowchart.

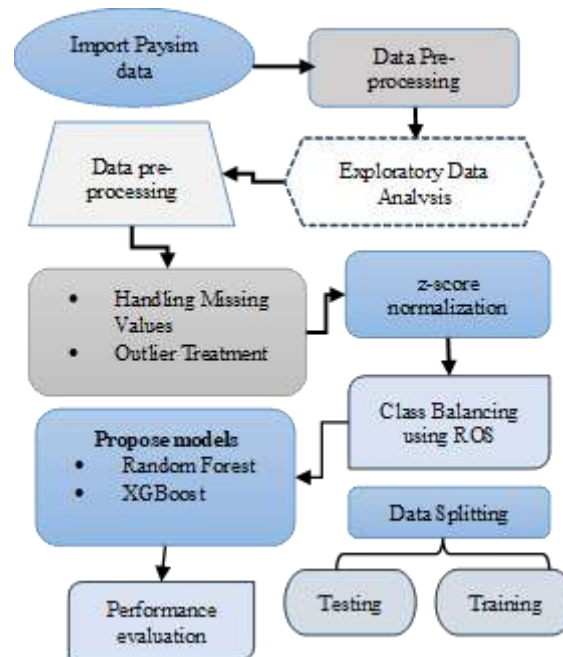


Fig 1 - Fraud detection in payroll system flowchart

The detail steps of the proposed methodology are explained below:

3.1. Data Gathering

A collection of fraudulent financial transactions, the PaySim Synthetic dataset is frequently employed by researchers in the fields of financial analytics and fraud detection. This dataset, which was created especially to simulate mobile money transactions, has characteristics that capture different aspects of these transactions. The dataset was acquired from Kaggle (<https://www.kaggle.com>) and comprises more than 6.3 million transaction records, which include attributes such as account balances, transaction types, and amounts.

3.2. Exploratory Data Analysis (EDA)

A deep and thorough process of EDA was carried out to uncover the internal patterns and distributional characteristics of the transaction data. Dataset transaction types (e.g., payments, transfers, cash out, debit, and cash in) are displayed in Figure 2.

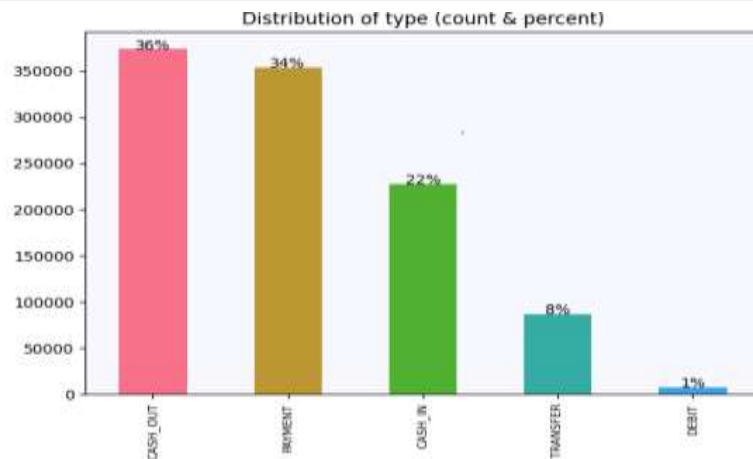


Fig 2 - Bar Graph for Transaction Type Distribution

Fig 2 displays the bar graph of the distribution of the types of transactions where CASH_OUT is the most common (36), followed by PAYMENT (34), CASH_IN (22), TRANSFER (8) and DEBIT (1). This reveals how withdrawal and payment transactions are more dominant than other categories.

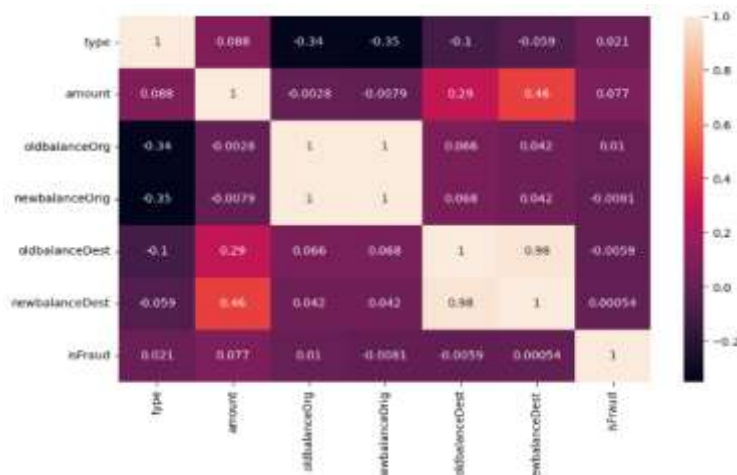


Fig 3 - Correlation Heatmap of Features

Fig 3 illustrates the correlation heatmap of features of the transactions, which shows the relationships between variables in the detection of fraud. There are strong positive correlations between oldbalanceDest and newbalanceDest whereas there is moderate positive correlation between amount and destination balances. Conversely, oldbalanceOrg and newbalanceOrg among other features exhibit weak or negative relationships with other variables. Individual features have very small correlations with the target variableFraud, which indicates that it is a complex phenomenon that cannot be based on individual features.



Fig 4 - Density Plot of Customer Balance Change

Figure 4 displays the density map of the change in customer balance, broken down by fraudulent and non-fraudulent transactions. The distribution indicates a sharp peak of non-fraud cases at zero, which means that the balance variation is insignificant in most of the legitimate transactions. Conversely, fraudulent transactions show a more dispersed and flatter pattern and indicate irregular changes in balance. This difference accentuates the fact that the changing balance can be a helpful tool in identifying fraud.

3.3. Data pre-processing

- Data cleaning, a critical component of the data pre-processing phase, is designed to eliminate any anomalies that could potentially detract from the efficacy of subsequent analyses. Ensure the integrity and reliability of the dataset by managing outliers and missing values.
- Handling Missing Values: Impute missing data using methods like mode, median, or mean imputation, or think about more sophisticated approaches like interpolation.
- Outlier Treatment: Identifying outliers and deciding whether to eliminate, transform, or handle them independently is done using statistical approaches or ML algorithms.

3.4. Feature Extraction

Classification models use the transaction characteristics as input. Various characteristics are linked to each kind of transaction. The value and frequency of transactions are two characteristics of transactions. There are a number of other information that may be included in a money transfer, such as the destination account, time, amount, geographical location (both the origin and the destination), entity (typical person or corporation performing the transfer), and kind of transaction (cash, money transfer). All are acknowledged as essential for the identification and classification of fraud. The semantic features are introduced into the classifier after they have been extracted from the original transaction data.

3.5. Data Normalization

The z-score normalisation process involves transforming the attribute values according to the standard deviation and mean. Equation (1), is as follows:

$$v'_i = \frac{v_i - A}{\sigma A}$$

(1) v_i stands for the value of an attribute A , whereas A and σA denote its mean and standard deviation, respectively. The mean will be normalised to zero, so if the value is less than the mean, it will be negative, and if it is more than the mean, it will be positive.

3.6. Class Balancing using ROS

The effect of using the Random Over Sampling (ROS) to correct the extreme imbalance in terms of the fraud data. First of all, the data is very skewed, with non-fraudulent transactions taking over with almost 99.87% versus 0.13 fraud cases. With the help of ROS the dataset is balanced and is represented in the 50:50 distribution as in Fig 5 which guarantees the equal representation of the two classes. This balance strategy has the effect of making the model more effective at learning patterns of fraudulent transactions, and reducing the bias it has to the majority class and overall detection.

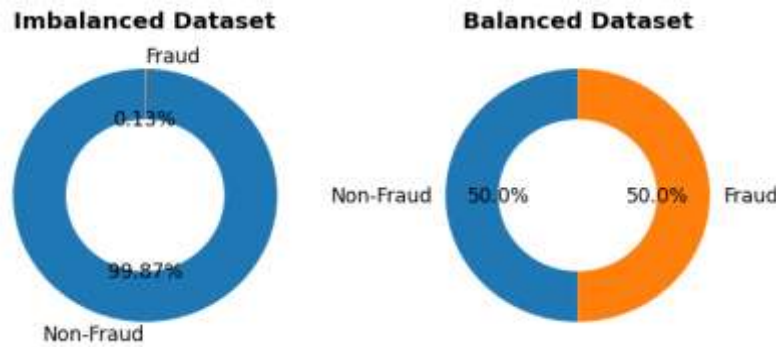


Fig 5 - Donut Chart for Imbalance and Balance Data Distribution

3.7. Train-Test Split

The standard procedure for evaluating a model's efficacy after pre-processing is to divide the dataset into a training set and a test set. The splitting ratio is 75:25.

3.8. Proposed Model

Two ML algorithms were devised and implemented for the purpose of real-time identification of fraud.

3.8.1. Proposed XGBoost Model

The XGBoost ensemble is a DT ensemble that is extremely scalable and is based on gradient boosting. The objective function is expanded additively by XGBoost, which minimises a loss function, in a manner similar to gradient boosting [16]. XGBoost manages the complexity of the trees by employing a variant of the lossfunction, as it exclusively employs DT as base classifiers, which is provided in Equations (2, 3).

$$L_{xgb} = \sum_{i=1}^N L(y_i, F(X_i)) + \sum_{m=1}^M \Omega(h_m) \quad (2)$$

$$\Omega(h) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (3)$$

where T represents the leaf count of the tree and w represents the leaf ratings for output. Pre-pruning techniques may be implemented by incorporating this loss function into the division criteria of decision trees. Trees get simpler as the value of γ increases. The value of γ controls the minimum gain in improvement of loss required to segment an internal node. XGBoost model was fitted with well-chosen hyperparameters to improve generalization and eliminate overfitting. In particular, 500 trees ($n_estimators$) were employed to obtain enough patterns, a maximum depth (max_depth) was set to regulate the complexity of the trees and a learning rate was set to 0.1 to balance speed and stability during convergence. The subsample ratio ($subsample$) of 0.8 and column sample by tree ($colsample_bytree$) of 0.8 were used to reduce overfitting, and regularization parameters $\gamma = 1$ and 0.8 were used to penalize complex trees and leaf weight. The objective function was binary:logistic, which is suitable for binary fraud. The training was done in an iterative manner to reduce the XGBoost loss function, which is a combination of prediction error, tree complexity, and early stopping of the training process, with a patience parameter of 50 rounds, using a validation set to optimize performance.

3.8.2. Proposed Random Forest

Breiman et al. were the first to create the Random Forest algorithm [88]. A network of linked learners known as an ensemble of classifiers is used to build decision trees. A training subset is utilised to create N decision trees using a greedy technique. Equation (4) displays how a forecast based on majority vote is subsequently generated by combining these decision trees.

$$R(y_i) = \frac{1}{N} \sum_{n=1}^N p_n(y_i) \quad (4)$$

where y_i is the class and $p_n(y_i)$ is the probabilities for each class. Random Forest classifier was trained based on a collection of decision trees with majority voting and hyperparameters were set to improve the performance. In order to assure stable prediction the model employed 300 trees ($n_estimators$) so as to capture the complex patterns, the maximum depth (max_depth) was 15, 10 samples per split ($min_samples_split$) and 5 samples per leaf ($min_samples_leaf$) to avoid overfitting and to ensure the stability of the node. $\sqrt{\text{total features}}$ ($max_features$) was used to decorrelate the trees and enhance diversity. Bootstrap sampling was also facilitated to train every tree on random samples of data, and this improved ensemble strength. The balanced dataset was trained, and the number of fraudulent and non-fraudulent transactions was equal, which enhances the capability of the model with regards to correctly classifying minority class cases.

3.9. Evaluation Parameters

Several confusion-matrix-based measures, including accuracy, precision, recall, F1score, and FNR, have been used to summarise the prediction outcomes on the test data after the suggested classifier has been repeatedly trained on the training data. Mathematically, the metrics can be represented by Equations (5)–(8), respectively.

$$ACC = TP + TN / TP + TN + FP + FN \quad (5)$$

$$Pre = TP / TP + FP \quad (6)$$

$$Rec = TP / TP + FN \quad (7)$$

$$F1 = 2 * Precision * Recall / Precision + Recall \quad (8)$$

where TP stands for true positives, FP for false positives, TN for true negatives, and FN for false negatives. The measure of accuracy examines the proportion of properly categorised cases. While recall value (also known as true positive rate or sensitivity) and FNR relate to the trained model's capacity to correctly identify every positive instance, precision (also known as positive prediction value) is the percentage of accurate positive predictions among all positive predictions made by the model. The number of times the model produced a successful prediction is shown by the F1 score, which combines recall and precision.

4. RESULTS AND DISCUSSION

This part presents a summary of the experimental assessment of the suggested models to detect fraud. The experiments were done under controlled environment on a high-performance workstation with Python and model performance was evaluated according to accuracy, precision, recall, F1-score and ROC-AUC. The outcomes show that both models are highly effective, with nearly perfect accuracy and reliability when using a Random Forest to detect fraud in real time.

4.1. Experimental Setup

All studies were performed in a confined context to confirm real-world deploy ability. Important system parameters:

- Hardware: Supercomputer with Windows 10, 12GB RAM, NVIDIA GTX 1660 Ti GPU.
- Software: Python, Jupyter Notebook, Intel Core i7 processor,
- Evaluation Metrics: Precision, Accuracy, Recall, F1-Score, ROC-AUC.

4.2. Model Performance Analysis

An iterative method of fine-tuning several hyperparameters and meticulously evaluating the model's performance was used in the proposal. The test set was the only one used to calculate any intermediate performance metrics throughout this procedure. Table I presents the performance of the proposed models in detecting fraud where Random Forest (RF) results in almost perfect performance with 99.99% accuracy, and ROC of 0.9998. The XGBoost (XGB) also performs quite well with an accuracy of 99.11% and 0.9897 ROC which indicates that both models are quite useful in the detection of fraudulent transactions.

Table 1: Proposed Models Performance for Fraud Detection

Models	Accuracy	Precision	Recall	F1-score	ROC
RF	99.99	99.87	99.98	96.86	0.9998
XGB	99.11	98.67	99.24	98.76	0.9897

The confusion matrices of the XGBoost and the Random Forest demonstrated in Figure. 6 indicate their ability to detect the fraud. The RF model shows almost perfect classification performance, which correctly recognizes and generates 103 FP and 3 FN, which shows excellent discrimination and little misclassification. Conversely, the XGBoost model also has high performance with relatively higher

error rates, whereby it recognized 550,368 non-fraud and 1,504 fraud cases with 2,071 false positives and 139 false negatives.

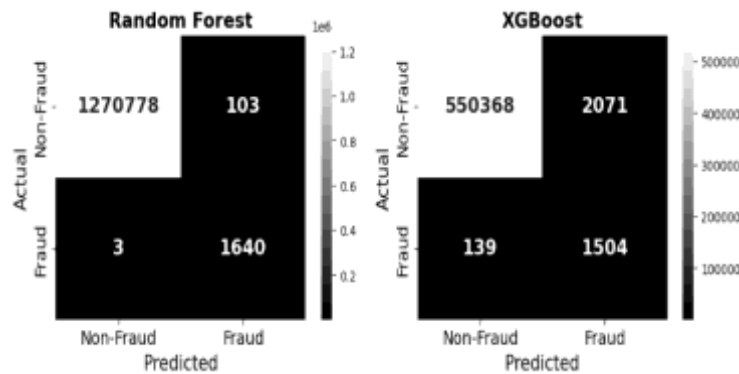


Fig 6 - Confusion Matrix for RF and XGBoost Model

Altogether, RF model is more robust in highly imbalanced fraud detection situations since it has better precision and recall especially when attempting to detect fraudulent transactions.

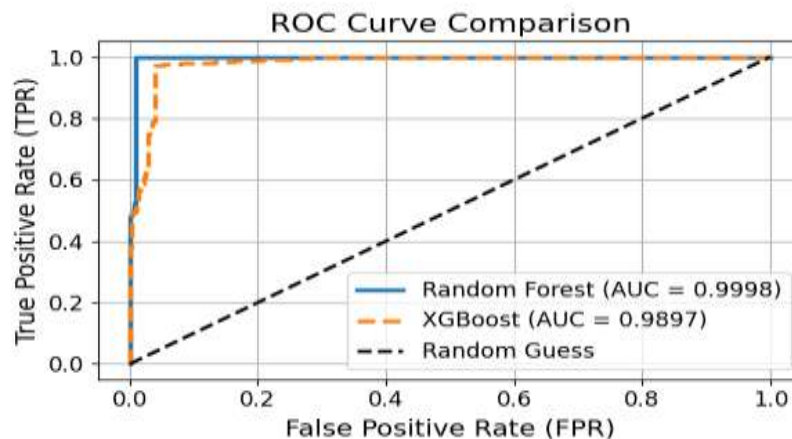


Fig 7 - ROC Curve Comparison between Propose Models

Fig 7 presents a comparison of the ROC curves between the XGBoost and the proposed Random Forest models, which can separate the fraud and non-fraud cases. Random Forest curve is almost on the top-left corner indicating the almost-perfect accuracy with an AUC of 0.9998 whereas the XGBoost also does well with an AUC of 0.9897 but is slightly lower. Both models can be seen as a great improvement over the random baseline, which is an important aspect to be considered, although the discriminatory power of the Random Forest can be seen as even greater.

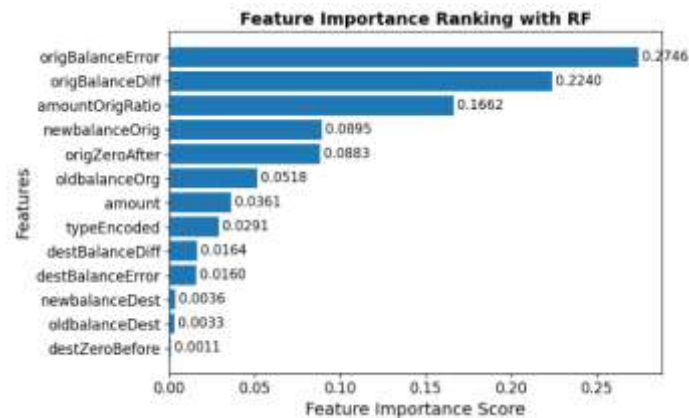


Fig 8 - Feature importance score using RF model

The importance of features of the variables according to the RF model is presented in Fig 8 and indicates the variables that have the most significant impact on fraud detection. The most influential ones are origbalanceError and origbalanceDiff, and the amountOrigRatio variable, whereas the variables like destZeroBefore and oldbalanceDest have very little influence. This distribution implies that account balance discrepancy and transaction ratios are the main features to be considered in decision-making of the model, and attributes related to destinations have a minor role in the model.

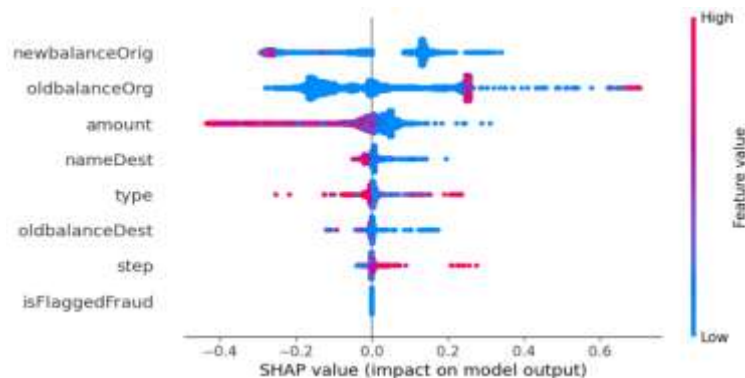


Fig 9 - SHAP Summary Plot for RF Model

Fig 9 depicts the SHAP summary plot to reveal the effect of significant features on predicting fraud, where the key features are newbalanceOrig, oldbalanceOrg and amount, and the more significant the value, the higher the risk of committing fraud. Other variables like nameDest and type play a moderate role with oldbalanceDest, step and isFlaggedFraud playing a small role. In sum, the plot affirms that balance- and amount-related aspects prevail in the decision process of RF model.

Table 2 compares the fraud detection models as decomposed by datasets, whereby the traditional methods, which include DT and LR, have an accuracy of 93-95 percent, whereas the deep learning models of LSTM and CNN have higher accuracy. The best models for identifying fraud are the ones suggested by Random Forest and XGBoost.

Table 2: Comparative Analysis for Fraud Detection on Two Different Datasets

Ref	Algorithms	Datasets	Evaluation Parameters				
			Accuracy	Precision	Recall	F1-score	ROC
[17]	DT	Credit Card Fraud Detection dataset	95.5	88.3	73.4	79.9	86.2
[18]	LR		93.23	91.23	95.72	93.42	93.22
[19]	CNN		84.6	84	85.7	84.8	84.6
[20]	XGB	Paysim	-	94	95	94	97
[21]	LSTM		97.9	95.6	93.4	94.5	98.2
[22]	BERT		89.6	89.6	91.2	90.4	-
Our	RF		99.99	99.87	99.98	96.86	0.9998
	XGB		99.11	98.67	99.24	98.76	0.9897

Some of the benefits of the proposed approach are that it has almost perfect performance in the detection of fraud, and a high level of interpretability with feature importance and SHAP analysis that aids in determining the factors that affect fraud. It is very effective in dealing with extreme class disparity and computationally efficient, so it can be deployed in real time. The research is also limited, as it is tested on an artificial dataset that may not reflect actual transaction behavior under real-world conditions. The use of Random Over Sampling can also lead to repetitive samples and overfitting. Moreover, the model can also be more or less effective in various financial fields, and the task is restricted to binary fraud detection and not to multi-class or evolving multi-class fraud.

5. CONCLUSION AND FUTURE WORK

Flow Payment and similar payment gateways perform millions of transactions every month and must have strong security measures to prevent financial losses due to fraudulent transactions. This paper suggested a good fraud detection model based on the PaySim dataset with RF and XGBoost (XGB) models, RF was found to perform very well with 99.99% accuracy and ROC-AUC of 0.9998 and XGB also performed well with 99.11-percent accuracy and 0.9897 ROC-AUC. These findings are evidently better than other conventional models like Decision Tree and Logistic Regression, and deep learning models like CNN, LSTM, or BERT, which underscores the strength of the suggested models when working with highly skewed transaction data. The importance of features and SHAP models also showed that transaction amount and balance discrepancies are the strongest predictors of fraud, which facilitated the interpretation of the model and its reliability. Future work will focus on testing the method on actual transaction data, investigating more sophisticated imbalance-handling strategies, and generalizing the framework to identify changing and multi-class fraud trends.

REFERENCES

- [1] Roy, A., Sun, J., Mahoney, R., Alonzi, L. P., Adams, S., & Beling, P. (2018). Deep learning detecting fraud in credit card transactions. In *2018 Systems and Information Engineering Design Symposium (SIEDS)* (pp. 129–134). IEEE. <https://doi.org/10.1109/SIEDS.2018.8374722>
- [2] Sinanc, D., Demirezen, U., & Sağiroğlu, Ş. (2021). Explainable credit card fraud detection with image conversion. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 10(1), 63–76. <https://doi.org/10.14201/ADCAIJ20211016376>
- [3] Tan, R., Tan, Q., Zhang, P., & Li, Z. (2021). Graph neural network for Ethereum fraud detection. In *2021 IEEE International Conference on Big Knowledge (ICBK)* (pp. 78–83). IEEE. <https://doi.org/10.1109/ICKG52313.2021.00020>

- [4] Jullum, M., Løland, A., Huseby, R. B., Ånonsen, K. H., & Lorentzen, J. (2020). Detecting money laundering transactions with machine learning. *Journal of Money Laundering Control*, 23(1), 173–186. <https://doi.org/10.1108/JMLC-07-2019-0055>
- [5] A hybrid deep learning model for online fraud detection. (2021). In *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)* (pp. 1–5). IEEE. <https://doi.org/10.1109/ICCECE51280.2021.9342110>
- [6] Dhieb, N., Ghazzai, H., Besbes, H., & Massoud, Y. (2020). A secure AI-driven architecture for automated insurance systems: Fraud detection and risk measurement. *IEEE Access*, 8, 58546–58558. <https://doi.org/10.1109/ACCESS.2020.2983300>
- [7] Gee, S. (2015). Payroll fraud. In *Fraud and fraud detection*. Wiley. <https://doi.org/10.1002/9781118936764.ch10>
- [8] Nguyen, G., Dlugolinsky, S., Bobák, M., Tran, V., López García, Á., Heredia, I., Malík, P., & Hluchý, L. (2019). Machine learning and deep learning frameworks and libraries for large-scale data mining: A survey. *Artificial Intelligence Review*, 52, 77–124. <https://doi.org/10.1007/s10462-018-09679-z>
- [9] Goes, A. S. de O., & de Oliveira, R. C. L. (2020). A process for human resource performance evaluation using computational intelligence: An approach using a combination of rule-based classifiers and supervised learning algorithms. *IEEE Access*, 8, 39403–39419. <https://doi.org/10.1109/ACCESS.2020.2975485>
- [10] Lucas, Y., & Jurgovsky, J. (2020). Credit card fraud detection using machine learning: A survey. *arXiv*. <https://arxiv.org/abs/2010.06479>
- [11] West, J., & Bhattacharya, M. (2019). Performance of machine learning techniques in the detection of financial frauds. *Procedia Computer Science*, 148, 45–54. <https://doi.org/10.1016/j.procs.2019.01.007>
- [12] Li, C., Ding, N., Zhai, Y., & Dong, H. (2021). Comparative study on credit card fraud detection based on different support vector machines. *Intelligent Data Analysis*, 25(1), 105–119. <https://doi.org/10.3233/IDA-195011>
- [13] Al-Hashedi, K. G., & Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. *Computer Science Review*, 40, Article 100402. <https://doi.org/10.1016/j.cosrev.2021.100402>
- [14] Liu, G., Tang, J., Tian, Y., & Wang, J. (2021). Graph neural network for credit card fraud detection. In *2021 IEEE International Conference on Big Data (Big Data)*. IEEE.
- [15] Liu, X., & Chen, J. (2021). Anomaly detection of JavaScript codes using one-class support vector machine and convolutional neural networks. *Journal of Information Security and Applications*, 61, Article 102857.
- [16] Nawrocki, P., & Osypanka, P. (2021). Cloud resource demand prediction using machine learning in the context of QoS parameters. *Journal of Grid Computing*, 19, Article 20. <https://doi.org/10.1007/s10723-021-09561-3>
- [17] Lucas, Y., & Jurgovsky, J. (2020). Credit card fraud detection using machine learning: A survey. *arXiv*. <https://doi.org/10.48550/arXiv.2010.06479>
- [18] Rajora, S., Li, D. L., Jha, C., Bharill, N., Patel, O. P., Joshi, S., Puthal, D., & Prasad, M. (2019). A comparative study of machine learning techniques for credit card fraud detection based on time variance. In S. Sundaram (Ed.), *Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1958–1963). IEEE. <https://doi.org/10.1109/SSCI.2018.8628930>
- [19] A. H. Almuteer, A. A. Aloufi, W. O. Alrashidi, J. F. Alshobaili, and D. M. Ibrahim, “Detecting Credit Card Fraud using Machine Learning,” *Int. J. Interact. Mob. Technol.*, vol. 15, no. 24, pp. 108–122, Dec. 2021, doi: 10.3991/ijim.v15i24.27355.
- [20] Psychoula, I., Gutmann, A., Mainali, P., Lee, S. H., Dunphy, P., & Petitcolas, F. A. P. (2021). Explainable machine learning for fraud detection. *arXiv*. <https://arxiv.org/abs/2105.06314>
- [21] Carminati, B., Ferrari, E., & Viviani, M. (2018). A fraud detection decision support system via human on-line behavior characterization and machine learning. In *Proceedings of the 2018 IEEE International Conference on Big Data (Big Data)*. IEEE.
- [22] Owolabi, O. O., & Olatunji, S. O. (2020). A framework for detecting credit card fraud with cost-sensitive meta-learning ensemble approach. *Scientific African*, 8, Article e00464. <https://doi.org/10.1016/j.sciaf.2020.e00464>.
- [23] Taluri, R. (2021). Cloud-Native Architectures for Enterprise Financial Data Management, Analytics, and Regulatory Reporting Compliance. *International Journal of Emerging Trends in Computer Science and Information Technology*, 2(2), 101–111. <https://doi.org/10.63282/3050-9246.IJETCSIT-V2I2P112>