# New Time-Efficient Fraud Detection System Using a Random Forest Model and AI Chatbot Layer

Shadmaan Akhand
Dept. of Computer Science
Georgia State University
Atlanta, Georgia
sakhand2@student.gsu.edu

Chirag Dwivedi
Dept. of Computer Science
Georgia State University
Atlanta, Georgia
cdwivedi1@student.gsu.edu

***Abstract*—Financial fraud continues to ramp up as modern digital payment systems scale in size and complexity, creating an environment where attackers can exploit automation and evolving criminal patterns. Traditional fraud detection methods (rule based, threshold driven, etc) struggle to keep pace with these novel threats. They frequently suffer from slow response times, limited adaptability, and high false positive rates. This results in an increased workload on security analysts, which consequently reduces operational efficiency, thus inadvertently deepening the issues rather than mitigating them as resources get spent and exhausted.**

**This paper presents a machine learning (ML) based fraud detection system that integrates with an artificial intelligence (AI) powered chatbot interface to further enhance accessibility and streamline the transaction process/review. Using a publicly available Kaggle dataset containing over 6.3 million real world financial transaction records, a Random Forest classifier was trained to distinguish between legitimate and fraudulent activities across things such as payments, transfers, cash outs, and deposits. This was possible by applying preprocessing steps such as one-hot encoding, numerical normalization, and feature selection to improve model performance. A chatbot layer allows users to input transaction attributes in natural language and numbers, then receive predictions instantly, reducing manual review time and improving user interaction. Results have shown a strong accuracy and robustness against overfitting, highlighting Random Forest's ability to tackle large scale and imbalanced datasets.**

**This research intends to provide a practical foundation for real-time fraud detection and an approachable interface suitable for analysts, security teams, businesses, end users, and provide as much peace of mind for all.**

***Keywords—Fraud Detection, Random Forest, Machine Learning, AI Chatbot, Financial Security***

## I. Introduction

The expansion of digital banking, mobile payments, and online money transfers has quickly led to more sophisticated forms of financial fraud, which security has not fully gotten a grasp on. Traditional fraud detection systems typically were rule-based and anomaly-based in their approaches, which encoded expert knowledge manually. Rule-based systems make their assessments on the notion that there are a set of rules that transactions should follow or pass compliance in order to be considered legitimate, and with anomaly-based systems essentially monitoring any deviations from normal network activity. These methods, however, fail to keep pace with unique fraud techniques. Furthermore, both systems mentioned have brought with them a high number of false positives, this is when the system flags legitimate transactions as malicious [10]. Issues like this, along with many others that arise from these systems, overwhelm security teams as it greatly reduces efficiency, wastes resources, lowers system trust, and lowers response capability. And as financial fraud rises, the need for automated, scalable, and adaptive detection mechanisms is more needed than ever.

Machine learning approaches and implementations have demonstrated significant potential in the realm of fraud detection, due in part to the ability of machine learning to pick apart complex, non-linear patterns from historical transaction data/activity [1]. Machine learning models such as logistical regression, Random Forest, neural network, and gradient boosting have shown success in recent studies [6]. These models, however, require substantial interpretation efforts to understand how the model reached its conclusion, essentially becoming a black box. At the same time, conversational AI chatbots and natural language assistants have been seeing an increase in adoption in many businesses [15]. Yet, an intuitive interface for security analyst staff or non-technical users, integrated directly with ML-based fraud detection, remains underexplored.

This research proposes a new time-effective fraud detection system that combines a Random Forest classifier with an AI chatbot layer to improve accessibility, reduce operational workload for security, and support near real time decision making. The Random Forest model is trained using a large Kaggle dataset with millions of labeled transactions, and the chatbot layer allows users to input transaction details and then receive immediate detection results along with confidence scores. This architecture aims to address the major limitations of fraud detections; slow decision making, limited interpretability, and lower operational workload. Simply put, this is less time spent manually inspecting low risk transactions, and customers can quickly verify suspicious activity.

## II. Related Works and Publications

Artificial intelligence and machine learning will only get better and play an even bigger role in the advancement of better fraud detection systems in today's environment and the future environment.

In a 2024 study, the study reported the effectiveness of an ensemble classifier (specifically of Random Forest and XGBoost) dealing with imbalanced and noisy financial datasets and it managed to achieve high accuracy readings, and reduced variance in predictions [6]. This shows that Random Forest remains a viable method for finance related situations as it has the capability to resist overfitting, have

high recall capability, and maintain a robustness in environments with a variety of transaction attributes. However, despite these findings, the machine learning models still struggled slightly against imbalanced data, especially where fraud transactions represented 1% or less of the whole dataset. This shows the need for more sampling and cost-sensitive learning techniques.

Another recent experiment used deep learning models to help detect novel or more unique detections in financial systems, with the report being there was strong performance at the price of significantly higher computational costs and an interpretability issue [1], making it not a viable candidate for deployment and scale especially in an high speed environment. Additionally, the knowledge, frameworks, and configurations required for deep neural networks are outside the reach of most security analysts from a practical standpoint.

AI has increasingly been implemented in the financial sector for customer support, account verification, etc [7]. Modern chatbot systems can help guide users through steps, gather transaction information, and reduce overall time spent, as noted in case reviews. However, there is a limited amount of studies currently available that study chatbot integration directly with machine learning based fraud detection models. This presents an opportunity to bridge machine learning with practical interfaces.

The system proposed in this paper builds on the advancements previously mentioned, by combining the accurate and predictive strength of Random Forest, with the AI chatbot layer, which helps create a more usable and time effective detection workflow.

## III. Description of New Method

### A. Overview

We are proposing a hybrid model that integrates an AI chatbot with a machine learning model in order to identify fraud and explain the reasoning behind its identification using natural language. This reduces investigation time and increases understandability for human investigators.
Our system consists of two components:

1. Classification model using random forest algorithms.
2. Explanation chatbot using OpenAI API.

### B. Description of Dataset

We used the Fraud Detection Dataset by Aman Ali Siddiqui from Kaggle [9], containing 6.3 million entries with the following columns:

1. Step – Time-Step in hours
2. Type of Transaction – Cashout, Transfer, etc
3. Amount – Transaction amount
4. OldBalanceOriginal – Sender's balance before transaction
5. NewBalanceOriginal – Sender's Balance after transaction
6. OldBalanceDest – Receiver's balance before transaction
7. NewBalanceDest – Receiver's balance before transaction
8. IsFraud – 1 = fraud and 0 = not fraud
9. IsFlaggedFraud – Attempts to Flag suspicious transaction

### C. Data Visualization

We performed the data visualization on Jupyter Notebook using the NumPy and Pandas Libraries.

The following are the most important visualizations that helps to better understand the underlying structure of the dataset:

1. Transaction type distribution:
   The bar chart (*Fig. 3.1*) shows the frequency for each type of transaction. According to the graph, Cashout and Payment dominate the dataset. The categorical distribution is highly informative and emphasizes the model's need to handle categorical imbalances efficiently.

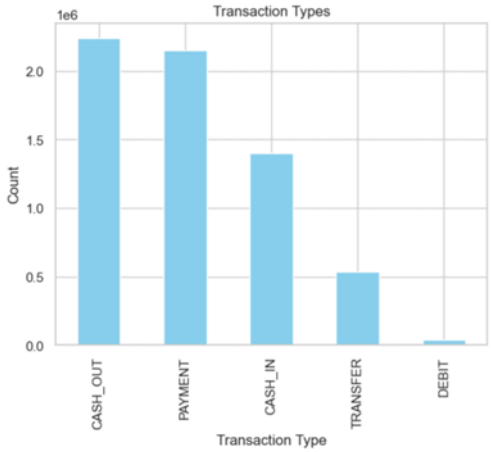


Fig. 3.1

2. Fraud Rate by Type:
   This fraud frequency by type chart (Fig. 3.2) shows that Cash Out and Transfer have the highest amount of fraud whereas types like Payment, Cash in and Debit have virtually no frauds. This reenforces to treat "Type" as a high impact predictive feature.

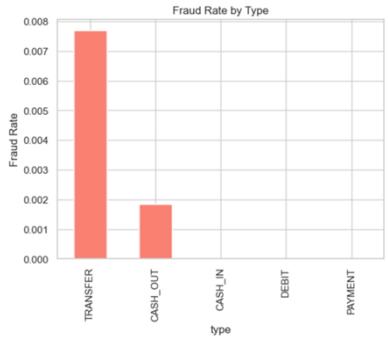


Fig 3.2

3. Transaction Amount Distribution:
   Fig 3.3 Shows the distribution of different transaction amounts. This gives an idea of what amount of transaction happens the most, which helps the model differentiates between normal transactions which are around the average amount and fraudulent transactions that happen at higher-than-average amounts.

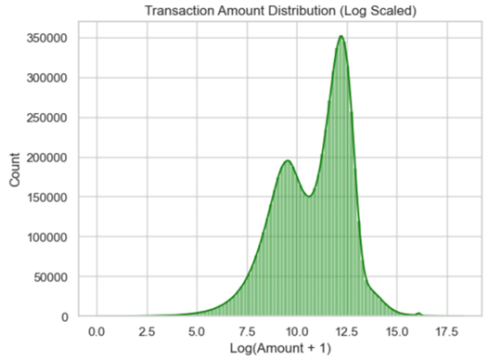


Fig 3.3

4. Boxplot of Amount vs IsFraud:
   Fig 3.4 shows two box plots. One for fraud transactions and the other for legitimate transactions. The boxplot shows that fraudulent transactions seem to have a much higher amount than legitimate transactions with the median and upper quartile significantly elevated. The plot shows that transaction amounts are a important factor.

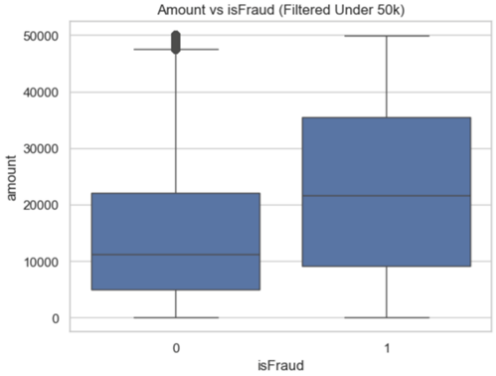


Fig 3.4

5. Correlation Matrix:
   The heatmap (Fig 3.5) reveals several important relationships:
   - Sender and Receiver balance change
   - Fraud has moderate relations with numerical features (amount: 0.08)

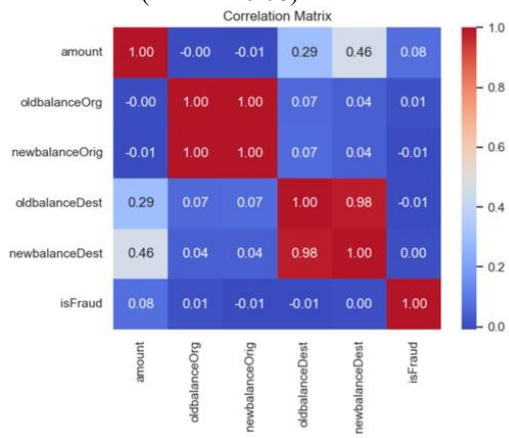


Fig 3.5

### *D. Original Method: Logistic Regression Model*

We first used the Logistic Regression Algorithm to build the prediction Model. It is a supervised classification algorithm used to predict whether an input belongs to class 0 or class 1 (Legitimate or Fraud).

Works by calculating a weighted combination of input features and passing it through a sigmoid function to output probability between 0 and 1.

$$\sigma(z) = 1 / (1 + e^{-z})$$

If Probability > Threshold (0.5), the model predicts Fraud = 1, Otherwise Legit = 0.

Since Our data was highly imbalanced with fraud being less than 1%, Logistic regression was not the best choice for the model as predictions defaulted to labeling majority of the transactions as fraud, giving a false sense of accuracy in catching frauds.

We needed to come up with a better solution, so we decided to go with a tree model as our final solution that can handle non-linear patterns as most of the data was not linear.

### *E.* Final Method: Random Forest

The final Algorithm that we decided to use for our predictive model was the Random Forest Algorithm. Due to our data being highly imbalanced, a linear algorithm would not be reliable and therefore we chose Random Forest which is a tree algorithm which uses nonlinear data structures.

Random Forest is a supervised machine learning algorithm that builds many decision trees (each tree learns a slightly different pattern) and combines their outputs. Each tree "votes," and the final prediction is decided by the majority vote. Works extremely well with complex, non-linear patterns, which is perfect for fraud detection.

This method proved to be much better than Linear regression due to its features like:

- Handling non-linear patterns
- Works better with categorical/Noisy data.
- More Robust to data imbalance – Doesn't default to predict only majority classes.
- Reduces overfitting – each tree is slightly different; combined, they generalize better than single model.
- Feature Importance – Shows which feature matter the most.

This model performed really well, which is discussed later in Section 5.

### *F. Chatbot Explainability Layer*

Along with the prediction model, we have also implemented a chatbot using OpenAI API that takes the following data as input: Predicted probability and transaction data. Using this data, the chatbot provides a brief explanation to the human interpreter about why a transaction was labeled as fraud.

This reduces fatigue of human interpreters and increases the understandability of a transaction.

## IV. System Implementation

### A. Model Training Environment

The model was trained in the following environment:

- Jupyter Notebook – For model training and visualization.
- Python Libraries – Pandas, NumPy and Scikit-learn.
- Streamlit – For user interface and transaction input.
- Joblib – to load trained data into app.

### B. Workflow

1. User inputs transaction details.
2. Model predicts fraud (0) or Legitimate (1).
3. Chatbot takes the prediction + transaction features and sends it to OpenAI server
4. OpenAI generates a summary and explanation of the transactions' nature and sends it back to the application.
5. Application prints the AI explanation so that the user is better able to understand why a transaction was flagged as fraud.

### C. Chatbot Implementation

We used the OpenAI API to implement the chatbot. Inputs include balance, amount, type, and prediction. The prompt is engineered to provide detailed and analyst-style explanations. After OpenAI generates the output, it is returned back to application through streamlit.

Fig. 4.1 shows the prompt that is sent to OpenAI API:

```
system_prompt = (
    "You are an assistant for a fraud detection system. "
    "Your job is to explain to a non-technical user why a transaction was "
    "predicted as potentially fraudulent or legitimate.\n\n"
    "Rules:\n"
    "- Always mention that this is a prediction, not a guarantee.\n"
    "- Focus on the main risk factors only (2–4 reasons).\n"
    "- Be concise: 2–5 sentences.\n"
    "- Use simple language.\n"
    "- Do not invent data that is not in the JSON.\n"
)
```

## V. Simulation Results and Performance Analysis

### A. Results for Logistic Regression Model

As discussed in prior sections that logical regression algorithm performed very poorly, due to which we had to move to Random Forest.

This section discusses the results that were observed from this model. The following Figs 5.1 shows all performance statistics that were observed:

```
print(classification_report(y_test, y_pred))
```

✓ 0.5s

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 1.00 | 0.95 | 0.97 | 1906322 |
| 1 | 0.02 | 0.95 | 0.04 | 2464 |
| accuracy | | | 0.95 | 1908786 |
| macro avg | 0.51 | 0.95 | 0.51 | 1908786 |
| weighted avg | 1.00 | 0.95 | 0.97 | 1908786 |

Fig 5.1.1

```
pipeline.score(X_test, y_test)
```

✓ 0.8s

```
0.9463926286131604
```

Fig 5.1.2

Figure 5.1.1 Shows the precision, recall and f1 score for the model and Figure 5.1.2 shows the accuracy.

These Statistics show that accuracy was very high (94%) which can be deceiving because on the other hand precision was found to be 0.02. This meant that the model was giving a lot of false positives, which is very dangerous for an actual fraud detection system. The recall was pretty good at 0.95 but due to weak precision, the f1 score also took a hit (f1 = 0.04).

### B. Results for Random Forest Model

The Random Forest model gave great results in terms of performance due to its ability to handle non-linear patterns well.

Figs 5.2.1 and 5.2.2 show the results that were observed by running this model.

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 1.000 | 1.000 | 1.000 | 1270881 |
| 1 | 0.965 | 0.792 | 0.870 | 1643 |
| accuracy | | | 1.000 | 1272524 |
| macro avg | 0.982 | 0.896 | 0.935 | 1272524 |
| weighted avg | 1.000 | 1.000 | 1.000 | 1272524 |

Fig 5.2.1

```
pipe.score(X_test, y_test)*100
```

```
99.96950941593244
```

Fig 5.2.2

Figure 5.2.1 shows the precision, recall, and f1 score for the model, and figure 5.2.2 shows the accuracy.

According to these statistics, we got a high accuracy of 99.96%. Our precision also bumped up to 0.965, which meant we got significantly less false positives. The recall did go down a little bit to 0.792, but it is still good enough to make the system reliable. As the precision and recall are now balanced, the f1 score was also observed to be higher than before (f1 score = 0.870).

### C. Performance Analysis

The simulations results highlight a clear progression from a weak baseline model to a highly dependable fraud detection approach. The Logistic Regression model demonstrated how misleading accuracy alone can be, as its extremely low precision produced an unacceptable number of false positives, an issue that already overwhelms operations. By transitioning to a Random Forest model, we obtain balanced performance with higher precision and strong recall.

## VI. Model Output

Output for potentially fraud transactions: (Fig. 6.1)

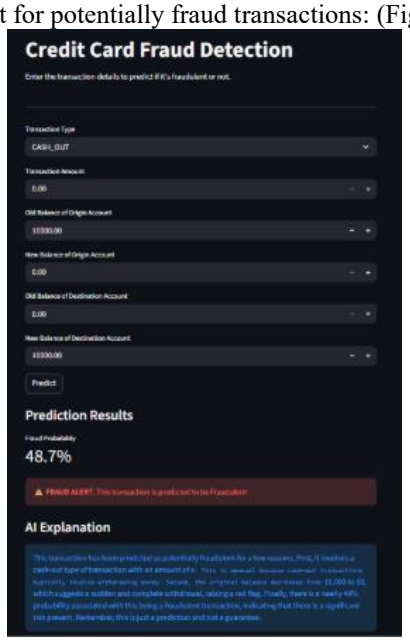


Fig 6.1

Output for Legitimate transactions: (Fig. 6.2)

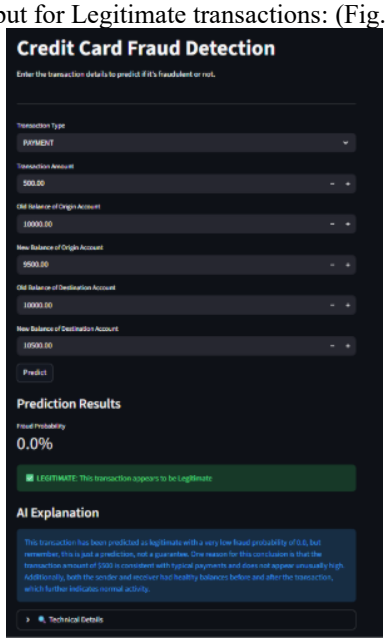


Fig. 6.1

As shown in the images, the following values are taken as inputs and sent over to the random forest predictive model: Transaction type, Transaction Amount, Old Balance Origin, New Balance Origin, Old Balance Receiver, and New Balance Receiver. Upon pressing the predict button, the model runs the decision tree algorithm to predict the probability of fraud. The threshold of fraud is set to 0.1 for this model, so if the probability goes over 10%, the transaction is concluded to be a possible fraud, which is highlighted in red in Fig. 6.1. Next, this prediction and the input values are sent to OpenAI where it generates and returns a reasoning behind the transaction being labeled as fraud or non-fraud as illustrated by Fig.6.1 and Fig.6.2 with text highlighted in blue.

## VII. Conclusion

This research introduced a new time-effective fraud detection system that integrated a Random Forest classifier with an AI powered Chatbot layer for interaction. This system aims to address major shortcomings of most fraud detection approaches, which include heavy manual workload placed on analysts, slow response times, high false positive rates, and transparency. By training the model on more than 6.3 million financial transaction records, the system was able to demonstrate a strong 99.96% accuracy and robustness in other metrics, while maintaining low prediction latency. The addition of the chatbot layer improves accessibility by allowing both technical and non-technical users to interact with the model.

Overall, the results show that combining machine learning based detection with a conversational chatbot layer offers a practical, scalable, and transparent foundation for a modern fraud detection system.

## VIII. Future Works

Several directions remain open for improving and evolving this system. First, additional techniques for handling extreme class imbalance (such as SMOTE, ADASYN, cost-sensitive learning, Focal Loss, etc.) may help increase the recall on rare and novel fraud cases. Second, integrating explainable AI mechanism, for example SHAP, which would provide further justification for predictions and further solidify the transparency factor that is often required in financial environments. Third, expanding the chatbot layer to better understand plain language and basic contextual details, as a way to improve user experience and reduce input constraints.

Future works can also branch into further hybridization or deep ensemble models, such as combining Random Forest with a transformer-based architecture to better capture things like sequential spending patterns. Lastly, connecting the system to live financial systems would train the detection system and provide real time monitoring.

## References

[1] Albalawi, Tahani, and Samia Dardouri. “Enhancing credit card fraud detection using traditional and deep learning models with class imbalance mitigation.” Frontiers in artificial intelligence vol. 8 1643292. 8 Oct. 2025, doi:10.3389/frai.2025.1643292

[2] Afriyie, J.K., Tawiah, K., Pels, W.A., Addai-Henne, S., Dwamena, H.A., Owiredu, E.O., et al. (2023) A Supervised Machine Learning Algorithm for Detecting and Predicting Fraud in Credit Card Transactions. Decision Analytics Journal, 6, Article ID: 100163. https://doi.org/10.1016/j.dajour.2023.100163

[3] Aburbeian, A.M., Ashqar, H.I. (2023). Credit Card Fraud Detection Using Enhanced Random Forest Classifier for Imbalanced Data. In: Daimi, K., Al Sadoon, A. (eds) Proceedings of the 2023 International Conference on Advances in Computing Research (ACR’23). ACR 2023.

Lecture Notes in Networks and Systems, vol 700. Springer, Cham. https://doi.org/10.1007/978-3-031-33743-7_48

[4] Yisong Chen, Chuqing Zhao, Yixin Xu, Chuanhao Nie, Yixin Zhang, Deep Learning in Financial Fraud Detection: Innovations, Challenges, and Applications, Data Science and Management, 2025, ISSN 2666-7649,https://doi.org/10.1016/j.dsm.2025.08.002.

[5] Devarakonda, Rahul Roy, Machine Learning Approach for Fraud Detection in a Financial Services Application (February 01, 2023). SSRN: https://ssrn.com/abstract=5234670 or http://dx.doi.org/10.2139/ssrn.5234670

[6] Hernandez Aros, L., Bustamante Molano, L.X., Gutierrez-Portela, F. et al. Financial fraud detection through the application of machine learning techniques: a literature review. Humanit Soc Sci Commun 11, 1130 (2024). https://doi.org/10.1057/s41599-024-03606-0

[7] Eustaquio-Jiménez, Roberto & Durand-Azurza, Mercedes & Gamboa Cruzado, Javier & León Morales, María & Ruiz, Nancy & Montoya, Reyna & Céliz, N.. (2024). Chatbots for customer service in financial entities—A comprehensive systematic review. Journal of Infrastructure, Policy and Development. 8. 10122. 10.24294/jipd10122.

[8] Compagnino AA, Maruccia Y, Cavuoti S, Riccio G, Tutone A, Crupi R, Pagliaro A. An Introduction to Machine Learning Methods for Fraud Detection. Applied Sciences. 2025; 15(21):11787. https://doi.org/10.3390/app152111787

[9] S. Aman, “Fraud Detection Dataset”, 2025, Kaggle dataset. [Online].Available: https://www.kaggle.com/datasets/amanalisiddiqui/fraud-detection-dataset?resource=download

[10] Olowu, Olawale & Adeleye, Ademilola & Omokanye, Abraham & Ajayi, Akintayo & Adepoju, Adebayo & Omole, Olayinka & Chianumba, Ernest. (2024). AI-driven fraud detection in banking: A systematic review of data science approaches to enhancing cybersecurity. GSC Advanced Research and Reviews. 21. 227-237. 10.30574/gscarr.2024.21.2.0418.

[11] Sundaravadivel, P., Isaac, R.A., Elangovan, D. et al. Optimizing credit card fraud detection with random forests and SMOTE. Sci Rep 15, 17851 (2025). https://doi.org/10.1038/s41598-025-00873-y

[12] Wang, Chao & Nie, Chuanhao & Liu, Yunbo. (2025). Evaluating Supervised Learning Models for Fraud Detection: A Comparative Study of Classical and Deep Architectures on Imbalanced Transaction Data. 10.48550/arXiv.2505.22521.

[13] Chen, Zhisheng. (2025). Revolutionizing finance with conversational AI: a focus on ChatGPT implementation and challenges. Humanities and Social Sciences Communications. 12. 10.1057/s41599-025-04725-y.

[14] Uddin, Nasir. (2025). Role of AI in Preventing Financial Crime: A Comprehensive Analytical Review. Journal of Economic Criminology. 100200. 10.1016/j.jeconc.2025.100200.

[15] Bhattacharjee, Rajat & Dev Rroy, Aruna. (2024). Artificial intelligence (AI) transforming the financial sector operations. ESG Studies Review. 7. e01624. 10.37497/esg.v7iesg.1624.