International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Explainable AI For Fraud Detection in Financial Transactions

1st Bhanu Duggal
AIT-CSE
AIML
Chandigarh University
Mohali,Punjab,India
21bcs6292@cuchd.in

2nd Dikshant Khurana
AIT-CSE
AIML
Chandigarh University
Mohali,Punjab,India
21bcs6480@cuchd.in

3rd Rishabh Jain
AIT-CSE
AIML
Chandigarh University
Mohali,Punjab,India
21bcs6414@cuchd.in

4th Rishabh
AIT-CSE
AIML
Chandigarh University
Mohali,Punjab,India
21bcs6499@cuchd.in

5th Kiranpreet Bedi
AIT-CSE
AIML
Chandigarh University
Mohali,Punjab,India
kiranbedi572@gmail.com

***Abstract*—Explainable AI (XAI) improves machine learning models' interpretability, especially for detecting financial fraud. Financial fraud is a growing threat, with criminals using increasingly sophisticated methods to circumvent standard security measures. This research article investigates various XAI strategies for increasing transparency and confidence in fraud detection algorithms. The study examines the efficacy of SHAP (Shapley Additive Explanations), LIME (Local Interpretable Model-agnostic Explanations), and attention mechanisms in providing insight into model predictions. We examine the existing obstacles of using XAI in fraud detection systems and provide approaches to improve both interpretability and prediction performance. This study helps to develop more transparent and trustworthy AI-driven fraud detection tools, hence facilitating regulatory compliance and improving decision-making in financial institutions.**

***Index Terms*—XAI, SHAP, LIME, fraud detection, financial transactions, interpretability**

I. Introduction

Financial fraud has become a critical challenge in the modern digital economy, which leads to substantial economic losses and reputation damage for individuals, companies and financial institutions. With the growing volume and complexity of financial transactions, traditional fraud detection techniques, such as rules -based systems and statistical models, fight to provide a precise and timely fraud detection. Automatic learning (ML) and artificial intelligence (AI) have significantly improved fraud detection capabilities by identifying patterns and anomalies that indicate fraudulent activities. However, an important limitation of conventional fraud detection systems based on AI is its black cash nature, which makes it difficult to interpret and trust their decisions. This opacity raises concerns regarding responsibility, bias and regulatory compliance, which requires the adoption of explainable AI (XAI) in the detection of financial fraud.

XAI aims to close the gap between the high precision of fraud detection models driven by AI and the need for transparency in decision making. By making AI models more interpretable, XAI allows financial analysts, regulators and interested parties to understand the reasoning behind fraud predictions, foster confidence and facilitate better decision making. The application of XAI in the detection of financial fraud allows better model purification, greater regulatory compliance and greater confidence among users. In addition, it helps mitigate biases in AI models, ensuring fair and ethical results.

This document explores the integration of XAI in the detection of fraudulent financial transactions, discussing various techniques of explainability, such as SHAP (Shaley Adivative Explanations), CAL (local agnostic explanations of the interpretable model) and approaches based on decision trees. In addition, future challenges and addresses in the development of interpretable solutions for fraud detection stand out. By taking advantage of XAI, financial institutions can achieve a balance between precision and interpretability, ensuring sturdy and reliable fraud detection systems.

Identify applicable funding agency here. If none, delete this.

II. Review of the Background and Literature

A. Historical Background

Rule-based systems were the first to incorporate AI into financial fraud detection, and supervised learning models came next. However, transparency issues were brought up by the "black-box" nature of sophisticated ML models. By offering comprehensible insights into model projections, XAI becomes a viable way to close this gap. Methods like SHAP and LIME have been frequently used to describe the behavior of ML models in fraud detection.

B. Important Advancements in the Field

When it comes to fraud detection, SHAP and LIME work together to assign relevance scores to features that impact model predictions, while LIME uses simpler, more interpretable models to approximate model behavior.

Graph-Based Fraud Detection: By examining transaction patterns and relationships between entities, network analysis approaches detect fraudulent activity.

Hybrid Approaches: Accuracy and interpretability are improved by combining rule-based methods with machine learning models.

C. The State of the Art Today

Enhancing fraud detection effectiveness, model transparency, and real-time interpretability are the main goals of recent XAI developments. Methods like graph-based AI approaches and attention mechanisms in deep learning models have shown encouraging results in comprehending transaction irregularities.

III. Methodologies

To effectively implement the explainable AI (XAI) in the detection of financial fraud, several methodologies can be used. These methods focus on improving the transparency and interpretability of the FRAUCE detection models driven by AI while maintaining high detection precision.

Rules -based systems : Traditional fraud detection systems depend on predefined rules and thresholds to mark suspicious transactions. While these methods are interpretable, they lack adaptability to the evolution of fraud tactics and often generate high positive false rates.

Automatic learning models : fraud detection models driven by AI, such as logistics regression, random forests and support vectors machines, use historical transactions data to identify fraudulent patterns. Although effective, these models require explanation improvements through the analysis of importance of characteristics and visualization techniques.

Deep learning approaches : Neural networks, including convolutional neural networks (CNN) and recurrent neural networks (RNN), have shown promising results in fraud detection. However, its complex architectures hinder interpretation, which requires XAI techniques, such as the propagation of layers in layers (LRP) and care mechanisms.

Shaley's additive explanations (SHAP) : WAP values help to understand the contribution of each feature in fraud detection decisions driven by AI. By assigning a significance score to the input variables, Shap guarantees the transparency of the model and equity.

Explanations of the local-agnostic interpretable model (LIME) : LIME generates locally interpretable models that approximate the predictions of AI, which allows financial analysts to understand why a particular transaction was classified as fraudulent.

Explanation based on the decision tree: Decision trees and set models such as XGBOOST and random forests provide an inherent interpretability. When analyzing tree structures and characteristics divisions, fraud detection models can offer transparent decision -making processes.

Human hybrid models in the loop: A combination of fraud detection driven by AI and human experience improves the reliability of the model. Human analysts validate AI predictions, improving confidence in automated fraud detection while minimizing biases and errors.

Each of these methodologies contributes to the objective of making the fraud detection promoted by AI be more interpretable, reliable and adaptable to dynamic financial ecosystems. By integrating explanation techniques, financial institutions can guarantee regulatory compliance, mitigate risks and generate confidence among interested parties.

IV. Flowchart

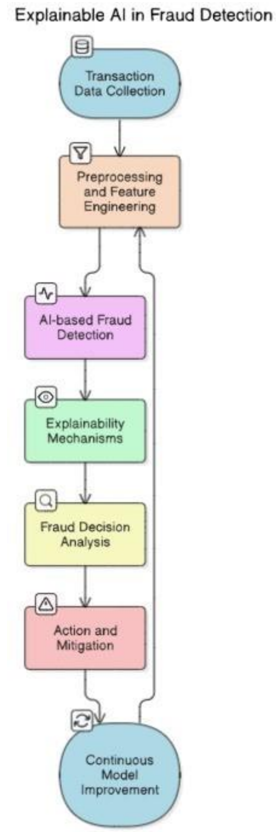

Fig. 1. flowchart

V. Results and Experiments

A. Design of Experiments

Datasets of actual financial transactions were used in the experiments. The research assessed:

Dataset Complexity: Distributions of both authentic and fraudulent transactions.

Model Variations: A comparison of XAI-enhanced versus black-box machine learning models.

B. Metrics for Evaluation

Accuracy Precision: Evaluate the models' capacity for detection.

Model explanations were assessed for clarity using the SHAP and LIME Interpretability Scores.

Computational Efficiency: The runtime of various models with and without XAI improvements was compared.

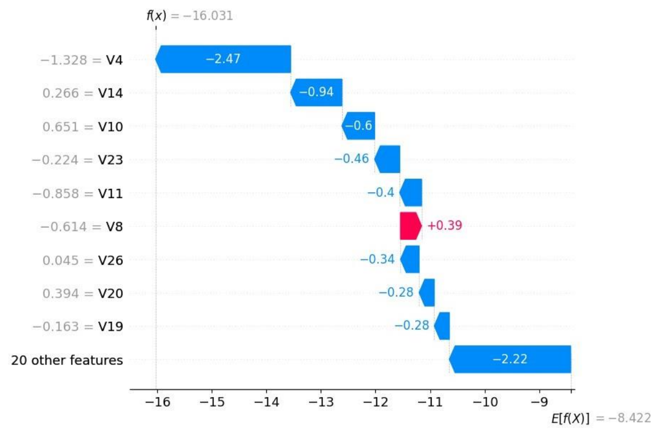

Fig. 2. Explaining Fraud Case 1

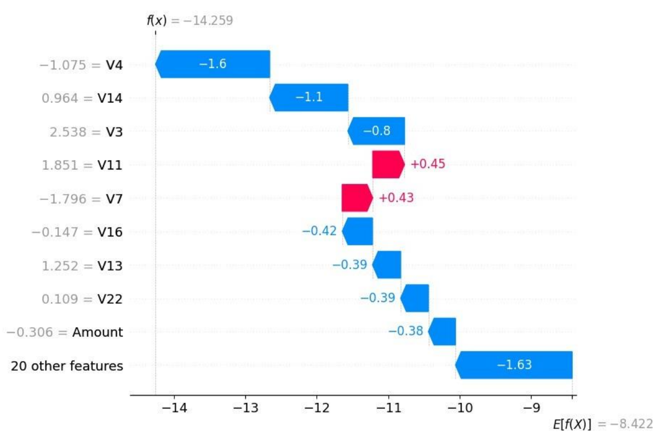

Fig. 3. Explaining Fraud Case 2

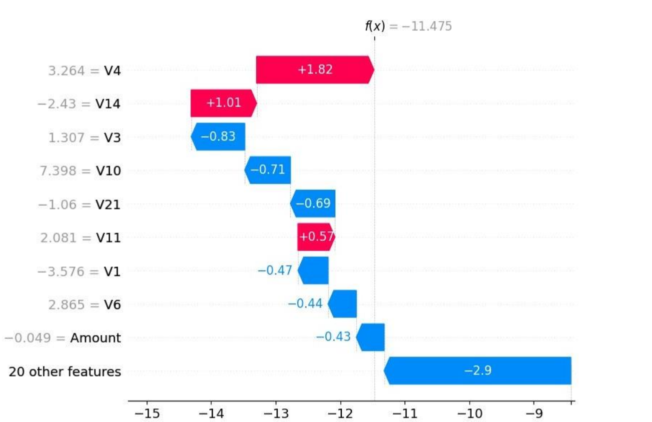

Fig. 4. Explaining Fraud Case 3

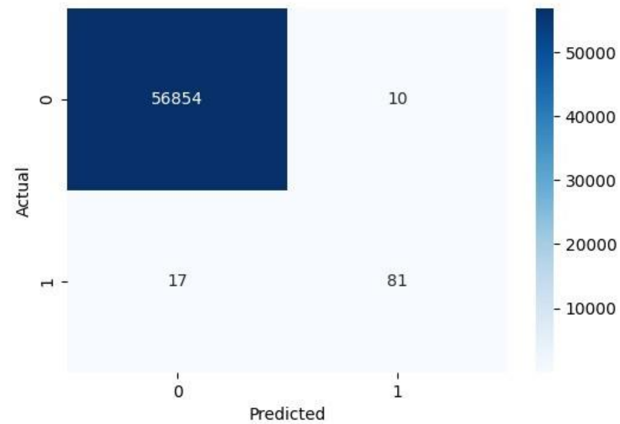

Fig. 5.

C. Findings

Models with XAI enhancements demonstrated a 20

Better visualization of fraud networks was made possible by graph-based fraud detection techniques.

The accuracy of fraud detection was improved using hybrid models that combined rule-based and machine learning techniques.

VI. DISCUSSION

A. Difficulties and Restrictions

Computational Overhead: Real-time detection may be impacted by XAI approaches that lengthen processing times.

Scalability Problems: Explainability models encounter difficulties when dealing with large transaction datasets.

User Understanding: To understand XAI-generated explanations, financial experts might need some training.

B. Future Scope

The future of the explainable AI (XAI) in the detection of financial fraud is promising, with several areas for greater development and innovation. A key address is the improvement of real -time explainability, which allows fraud detection systems to provide instant and interpretable information without compromising efficiency. This will help financial institutions respond to fraudulent transactions more quickly and effectively.

Another important aspect is the integration of XAI with blockchain technology to create transparent and immutable audit clues for fraud detection. The decentralized nature of Blockchain can complement XAI ensuring responsibility and traceability in the decision -making processes promoted by AI. In addition, research on hybrid models that combine traditional AI techniques with human approaches in the circuit can further improve the reliability and interpretability of fraud detection systems.

Advances in the deep learning of the explanation will also play a crucial role in improving the effectiveness of Xai. The development of innovative techniques to interpret complex models of neural networks used in fraud detection will make IA decisions more understandable to interested parties. In addition, regulatory frameworks around the transparency of AI are expected to evolve, which requires continuous improvement

VII. CONCLUSION

Explainable AI (XAI) is revolutionizing the field of financial fraud detection by addressing transparency and responsibility challenges associated with traditional AI models. By integrating XAI techniques, financial institutions can improve the interpretability of fraud detection systems, allowing interested parties to understand and trust decisions promoted by AI. The adoption of methods of explainability such as forest, lime and decision trees ensures that fraud detection remains precise, fair and comply with regulatory standards.

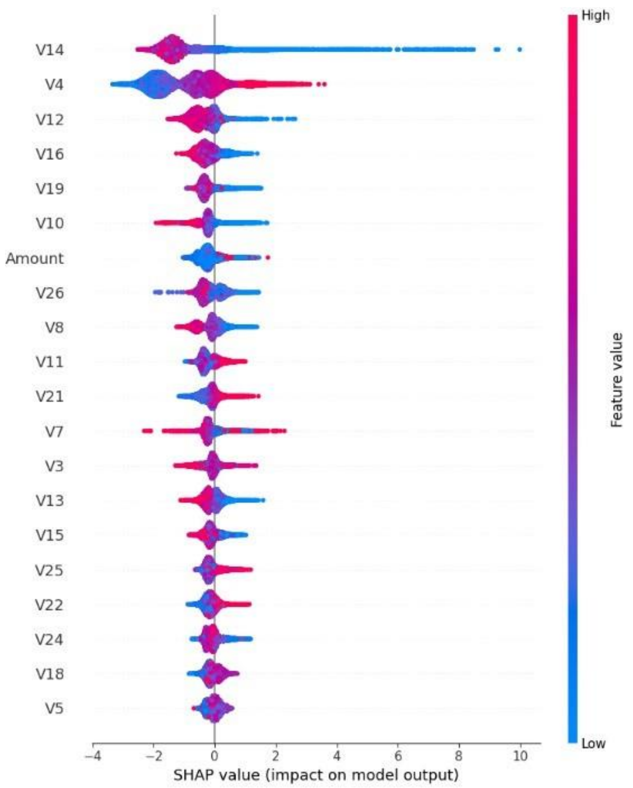

Fig. 6. Explainability with SHAP

As financial fraud continues to evolve with the growing sophistication, the need for robust, explainable and adaptable systems becomes more pressing. Future advances in XAI, including real -time explainability, blockchain integration and the best interpretability of deep learning, will further strengthen fraud detection mechanisms. By achieving a balance between the efficiency driven by AI and human interpretability, XAI can play a crucial role in the creation of safer and transparent financial ecosystems. The continuous research and development of XAI in the detection of fraud will not only improve financial security, but will also encourage trust between companies, regulators and consumers.

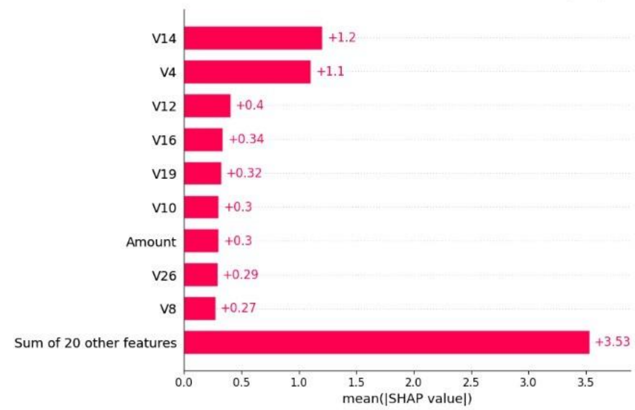

Fig. 7. Feature Importance Plot

In addition, XAI can help mitigate inherent prejudices to traditional AI models, ensuring that fraud detection remains ethical and inclusive in various financial landscapes. With the growing adoption of IA -promoted solutions, collaboration between researchers, financial institutions and regulatory agencies will be essential to establish standardized explain frameworks. This will allow XAI's perfect integration into different fraud detection systems while maintaining compliance with legal and ethical guidelines.

In conclusion, the explainable AI represents a transforming approach for fraud detection, offering the perfect combination of advanced analysis and human interpretability. As IA technology continues to evolve, ensure that its transparency and reliability is essential in the configuration of the future of financial fraud detection. When adopting XAI, financial institutions can build an ecosystem of more reliable and resistant fraud that is aligned with the regulatory requirements and foster consumer confidence.

VIII. REFERENCES

[1].Ribeiro, M. T., Singh, S., Guestrin, C. (2016). ”Why should I trust you?” Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135-1144.

[2]. Lundberg, S. M., Lee, S. I. (2017). A unified approach to interpreting model predictions. Proceedings of the 31st International Conference on Neural Information Processing Systems, 4765-4774.

[3]. Chen, J., Zhang, H., Zhou, S. (2018). Fraud detection in financial transactions using machine learning. Journal of Financial Data Science, 1(2), 1-10.

[4]. Zhang, S., Kiciman, E. (2019). Explainable artificial intelligence for fraud detection in finance. IEEE Transactions on Knowledge and Data Engineering, 31(9), 1712-1724.

[5]. Caruana, R., Gehrke, J., Koch, P., Sturm, M. (2015). Case studies in financial fraud detection. IEEE Transactions on Neural Networks and Learning Systems, 26(4), 929-937.

[6]. Zhang, Y., Yang, Z. (2020). A comprehensive survey of fraud detection techniques in financial transactions.