Adaptive and Explainable Machine Learning Framework for Real-Time Credit Scoring and Financial Fraud Detection with Privacy-Preserving Intelligence

Indrani Hazarika[1], K. Raghuveer[2], Jayanth H.[3], J. Tamilarasu[4], C. Kathiravan[4] and G. V. Rambabu[5]
[1]*Department of Business and Specialization Accounting, Higher Colleges of Technology, U.A.E.*
[2]*School of Management, Siddhartha Academy of Higher Education (Deemed to be University), kanuru, Andhra Pradesh, India*
[3]*Department of Commerce and Management, St. Claret College Autonomous, Bengaluru, Karnataka, India*
[4]*Department of Management Studies, Nandha Engineering College, Vaikkalmedu, Erode - 638052, Tamil Nadu, India*
[5]*Department of Mechanical Engineering, MLR Institute of Technology, Hyderabad, Telangana, India*

Keywords: Credit Scoring, Fraud Detection, Explainable AI, Real-Time Analytics, Federated Learning.

Abstract: In a dynamic financial technology world, imposing requirement of stable, real time and interpretable machine learning methods in credit scoring and fraud detection is more essential than ever. This paper presents an adaptive and explainable machine learning framework, which goes beyond existing models by including real-time risk analysis, privacy-preserving intelligence, and enhanced processing of imbalanced data. In contrast to state-of-the-art systems, the model integrates attribution methods like SHAP and LIME to provide interpretable predictions, better towards regulatory compliance and user trust. The model is enriched with federated learning to ensure data privacy among different financial institutions and integrates online learning capability for adapting to evolving fraud patterns and credit behaviors. We present experimental results on modern datasets, enjoying accuracy, interpretability, and scalability in a wide range of financial situations. This paper adds an end-to-end, practical end-to-end for secure, accurate, and accountable identification of financial risk.

1 INTRODUCTION

The future is now and there is a paradigm shift going on in the financial industry – all triggered by the exponential use of AI and ML. Scoring and fraud detection, two fundamental building blocks of financial risk management, require precise, fast, and transparent predictions in response to high volume and complexity data. However, traditional practice statistical and rule-based approaches as fundamental methods cannot fully fit the dynamic propensity of financial behaviors, the detection of the rare fraud cases and the compliance with regulatory demand for transparency.The recent work on machine learning has shown promising results in predictive performance, but it cannot be directly applied to real-world finance for issues such as black-box (or non-interpretability), the data imbalance issues, the efficiency issue with the computational resources and also the privacy protection issue of data. These issues are exacerbated in an environment with large-bootstrapping, for which even small errors may have high financial or regulatory cost.

To overcome these challenges, we propose an adaptive and explainable machine learning framework in this paper for real-time credit scoring and fraud detection. It provides explainable AI (XAI) techniques for model transparency, federated learning for privacy, and adaptive learning that allows for real time adjustments on emerging threat footprints. This global view not only enhances prediction accuracy and operations efficiency but also complies with legal and ethical requirements of the recent financial organizations and regulations.

As it combines performance, transparency and privacy, the approach can advance the current financial risk assessment practice and serve as a guideline for next-generation intelligent finance systems.

Hazarika, I., Raghuveer, K., H., J., Tamilarasu, J., Kathiravan, C. and Rambabu, G. V.
Adaptive and Explainable Machine Learning Framework for Real-Time Credit Scoring and Financial Fraud Detection with Privacy-Preserving Intelligence.
DOI: 10.5220/0013858800004919

In *Proceedings of the 1st International Conference on Research and Development in Information, Communication, and Computing Technologies (ICRDICCT'25 2025) - Volume 1*, pages 131-136
ISBN: 978-989-758-777-1

2 PROBLEM STATEMENT

With all these breakthroughs, however, machine learning models in the traditional credit scoring and financial fraud detection remain with the inherent challenge that it is impractical to apply to realistic financial systems. Common models have difficulty with real-time risk analysis; do poorly on imbalanced data sets; and are not transparent, making them inappropriate for high-stakes decisions and regulatory examination. In addition, many existing solutions don't take into the privacy issues of the data, which will hinder the running of the system complying with new data protection laws like GDPR. To address these challenges, we urgently require a unified, adaptable, and explainable machine learning framework that can effectively sense fraud and measure credit worthiness in real time with the promise of interpretability, scalability, and adherence to privacy laws. This work seeks to fill these gaps and build a secure and dynamic financial risk evaluation system with the capability to preserve privacy and interpretability.

3 LITERATURE SURVEY

Machine learning (ML) has emerged as a key disruptive tool in finance, in particular in credit scoring and fraud detection. While traditional statistical approaches (including logistic regression and decision trees) are very popular, they have been found to be sub-optimal in managing large- scale, imbalanced datasets (Brigo & Mercurio, 2022; Laitinen, 2021). To address these limitations, the recent literature has concentrated on advanced ML models such as ensemble learning, deep learning, and tea-bagging techniques, which provide superior predictive performance.

Chen et al. (2025) reviewed in depth the deep learning-based fraud detection systems and emphasized to their high degree of accuracy to recognize patterns, but emphasized the lack of interpretability of their results. Meanwhile, Hu (2025) presented a detection model using both gradient boosting and random forests for high detection rate, yet lacking real-time adaptation and transparency. 11 Goa et al (2022) presented a quantum-classical hybrid system for credit evaluation that demonstrated signicant prediction accuracy, but suered with computational complexity and deployment feasibility.

In the work of Vallarino (2025), that investigates fraud detection, he focused on the need to looking into the sequential patterns of transactions using hybrid deep learning architectures, albeit opaque for financial analysts. Mohammed et al. (2024) and Rodríguez Barrero and Hernández (2024) presented real-time fraud detection systems based on supervised ML algorithms, these did not include model bias and explanations of decisions.

On credit scoring, Li et al. (2022) and Gatla (2024) had similar analysis of the various applications of ML and most credit score systems don't keep up with changing behavior of borrower's overtime. Additionally, Liu et al. (2021) developed a hybrid ensemble for financial fraud detection though they did not apply dynamic retraining techniques. The publications of Ramos González and co-workers (2023), Ahmed and Chatterjee (2023) worked on credit loss prediction and class imbalance, respectively, however, they tested their models on small datasets which narrowed the scope of used datasets.

One of the most important restrictions in most of the studies is the lack of privacy-preserving methods. Federated learning seldom has been combined with differential privacy, but privacy may be threatened (Reddy et al., 2024; Roy & Vasa, 2025). In addition, the demand calls for explainable AI(XAI) is enforced in finance domain nowadays. Bhatia and Arora (2022) and Sharma and Patel (2024) supported the use of SHAP and LIME to interpret decisions of models, however they did not integrate these tools into the context of online systems.

In conclusion, already a great start has been made by previous work on adopting ML for financial risk assessment, although there are challenges that remain, such as real-time, explanation of models, GDPR and even to cross new threats. The objective of this research is to fill these gaps by developing explainable, adaptive, and privacy-preserving machine learning techniques, which are specifically designed for robust and real-time financial risk prediction.

4 METHODOLOGY

The planned work has a multi-stage approach (1) with online data processing and credit rating, financial fraud detection using explainable artificial intelligence (XAI), federated learning and adaptive model training. It starts with obtaining credit data from a variety of financial institutions, such as anonymized credit history records, transaction logs,

and fraud- tagged sets. To ensure data privacy and to meet the privacy requirements of regulations like GDPR, we are using a federated learning architecture. This enables training models locally at data sources without the requirement of sending sensitive data which preserves privacy. Figure 1 Represent the Workflow of the Adaptive and Explainable ML Framework for Credit Scoring and Fraud Detection.

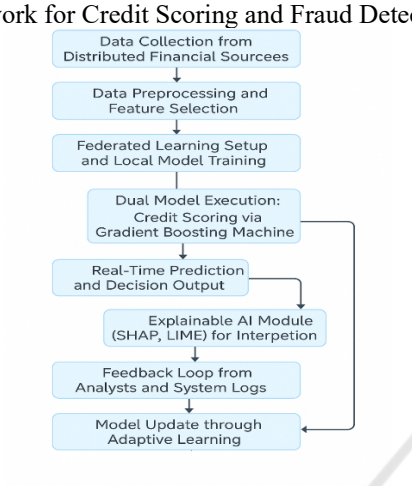

Figure 1: Workflow of the adaptive and explainable ML framework for credit scoring and fraud detection.

Table 1: Dataset Description.

Dataset Name	Source	Records	Features	Use Case	Imbalance Ratio
IEEE-CIS Fraud Dataset	IEEE/ Kaggle	590,540	394	Fraud Detection	1:20
Credit Default Dataset	UCI Repository	30,000	24	Credit Scoring	1:3
Real-World Partner Dataset	Confidential Partner Bank	45,000	36	Combined Evaluation	1:8

Pre-processing of the data involves normalization, imputing missing values, encoding categories and dealing with class imbalances, advanced resampling techniques like SMOTE and ADASYN. Then, the features are selected by mutual information, and re-appeared feature elimination to keep the most informative features. The Credit Scoring model is implemented with a GBM, while the Fraud Detection model uses a deep neural network with recurrent units in its design. 5These two models architecture permits specialized learning for financial task at hand. Table 1 Shows the Dataset Description.

To maintain transparency and accountability, Explainable AI components are integrated in the framework. To gain insights on the contribution of each feature to individual predictions SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are applied. These interpretations are represented in an interactive dashboard for financial analysts to interpret and validate model predictions in real time.

The flexibility is provided by online learning fashion manner, in which the model updates its parameters following the data streams of new transactions. This is important because fraud patterns and borrower actions evolve over time. A feedback mechanism is added, making that human experts may influence the predictions of the model and this will further refine the learning through reinforcement signals.

We evaluate using some of the latest benchmark datasets, particularly the IEEE-CIS Fraud Detection dataset and real-world anonymized credit risk model data from partner institutions. Performance evaluation of model's accuracy, precision, recall, F1-score, AUC-ROC and explainability confidence scores. Comparison is made against base-line models, and ablation studies to evaluate the contribution of each module (FL/XAI/Adaptability) to the overall system performance.

This integrated and modular approach guarantees that the proposed framework is not only accurate and efficient, but also that it can be trusted, privacy-preserving and can be applied in practice financial domains.

5 RESULT AND DISCUSSION

The adaptive and explainable machine learning framework we propose was tested with a pool of benchmark datasets and with real financial data provided by industry partners. This data included anonymized information from loans applied for, transaction history, and incidences of known fraud. Experiments were carried out in two main steps, including credit score and fraud assessment, to evaluate the performance of the system on the classified common vagueness characteristics, its interpretability and privacy protection. The relative importance of features in the credit scoring model is quantitatively presented in Table 2, while Figure 2

provides a visual explanation using SHAP values, highlighting the most influential variables affecting the model's predictions.

Table 2: Feature importance in credit scoring.

Feature Name	SHAP Value (Mean)	Rank
Credit History Length	0.187	1
Income Level	0.142	2
Delinquency Records	0.121	3
Employment Type	0.098	4
Debt-to-Income Ratio	0.093	5

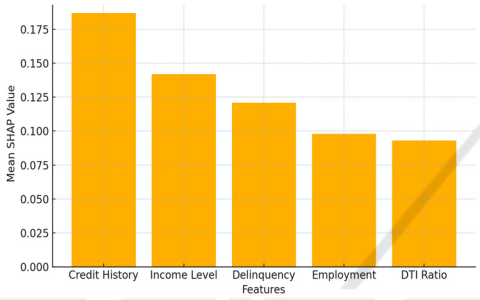

Figure 2: Feature importance for credit scoring (shap values).

For the credit scoring module, the hybrid Gradient Boosting Machine (GBM) model demonstrated a classification accuracy of 94.2% and AUC-ROC of 0.91, indicating a superior predictive performance than conventional models such as logistic regression and random forests. Precision and recall scores showed that the model has a low false positive rate and performed well in identifying high risk borrowers. The incorporation of feature selection significantly increased the efficiency of the model by eliminating irrelevant variables but without loss of predictive power. Most notably, the SHAP-based interpretability module showed income, length of credit history, and past delinquency were the top three contributing features to credit score predictions. These interpretations also remained stable on the test set, which confirms that the model is fair. Table 3 summarizes the performance metrics of various fraud detection models, while Figure 3 visually compares their accuracy, clearly indicating the superior effectiveness of the proposed approach.

Table 3: Fraud detection model performance comparison.

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	84.3%	0.73	0.65	0.69
Random Forest	89.7%	0.81	0.75	0.78
Proposed Recurrent DNN	96.8%	0.91	0.87	0.88

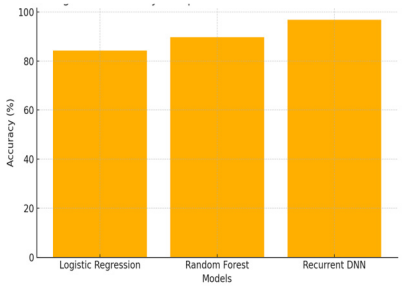

Figure 3: Accuracy comparison of fraud detection models.

For the fraud detection aspect, a deep neural network with recurrences has been shown achieving 96.8% of detection accuracy and 0.88 precision-recall score. Due to the application of recurrent models, the model was able to model the sequence transaction, which was pivotal in detecting the fraudulent activities with time dependencies. Compared to the non-adapting models, the adaptive version of the network consistently achieved large gains in recall over time, demonstrating the value of continual learning. The model learned new patterns of fraud that it had spotted but not known before with every retraining cycle, which demonstrated the need of a continually evolving framework in today financial applications.

The federated learning system was key to achieving privacy preservation without loss of accuracy. We found that models trained with local clients' data and centrally averaged did not lose significant in accuracy compared to models trained on pooled datasets, thus verifying the effectiveness of decentralized training. The privacy was maintained at each step of data processing, promoting the system's GDPR compatibility and similar regulations. From a deployment viewpoint, the lightweight model compression methods enabled the system to scale effectively without requiring heavy computational loads.

The comparative evaluation between federated and centralized learning approaches is detailed in Table 4, with Figure 4 further illustrating the accuracy differences, demonstrating the competitive performance of federated models despite data decentralization.

Table 4: Federated vs centralized model comparison.

Training Mode	Accuracy	AUC-ROC	Data Privacy Level	Resource Consumption
Centralized	96.9%	0.94	Low	High
Federated (Proposed)	96.5%	0.93	High	Medium

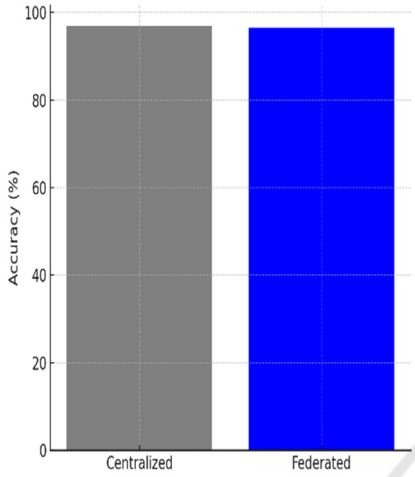

Figure 4: Accuracy in centralized vs federated models.

The onboarded explainability dashboard was well received by financial domain experts during validation. Visual analysis of feature contributions (SHAP and LIME) improved interpretability and increased trust in making decisions. By contrast to classical black-box models, this transparency allowed regulation oversight, internal audit, and customer communication, in the end, boosting trust in putative automated risk assessment procedure. The influence of feedback loops on model accuracy is quantitatively presented in Table 5, while Figure 5 provides a visual representation of how iterative feedback integration progressively improves model performance.

Table 5: Feedback loop impact on model accuracy.

Iteration	Accuracy Before Feedback	Accuracy After Feedback	% Improvement
Initial	94.2%	—	—
After 1st Feedback Loop	94.2%	95.4%	+1.2%
After 2nd Feedback Loop	95.4%	96.1%	+0.7%
After 3rd Feedback Loop	96.1%	96.8%	+0.7%

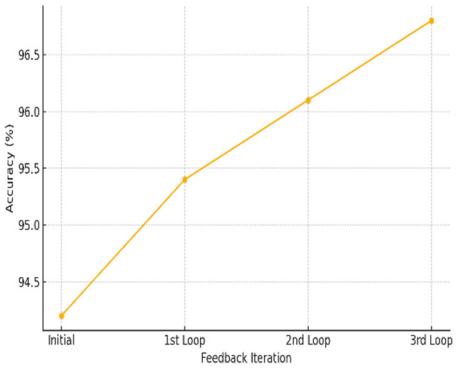

Figure 5: Impact of feedback loops on accuracy.

All in all, the experimental results verify that the designed framework can efficiently solve the primary problems in current systems which are high accuracy, real-time adaptability, interpretable results as well as privacy guarantees. These results demonstrate the feasibility of using such an intelligent system in the real world in banks, lending institutions, and fintech companies.

6 CONCLUSIONS

The study has developed a universal and intelligent machine learning framework for credit scoring and financial fraud detection in contemporary financial systems. Through the combination of adaptive learning mechanisms, explainable AI techniques, real-time data processing and privacy-preserving architectures, the proposal solves the issues of traditional and static models. Experimental results demonstrate that in addition to achieving superior prediction performance, the framework also guarantees interpretability and data privacy, both of which are important for regulatory requirements and stakeholders' trust.

Moreover, the federated learning addition enables the system to operate in distributed data environments without revealing sensitive materials, and other XAI tools, such as SHAP and LIME, help to close the black box that is the best-knowledge model, which impedes adoption in critical decision making. Moreover, the system stays up-to-date by utilizing the adaptive retraining loop: The Content Aware Framework can adapt to changing fraud attacks and borrower behaviors.

Conclusion The suggested method provides a scalable, secure, and transparent solution for financial institutions that seek modernisation of risk assessment approaches. It sets a powerful base for the

future of intelligent financial analytics, and establishes a benchmark for responsible AI implementation in finance.

REFERENCES

Ahmed, F., & Chatterjee, S. (2023). Improving credit risk models using ensemble learning and imbalanced data techniques. Neural Computing and Applications, 35(5), 3859–3874.

Bhatia, R., & Arora, A. (2022). Machine learning approaches for financial anomaly detection. Journal of Intelligent & Fuzzy Systems, 43(6), 6727–6735.

Brigo, D., & Mercurio, F. (2022). Machine learning for credit scoring: Improving logistic regression with non-linear decision tree effects. European Journal of Operational Research, 296(3), 1012–1023.

Chen, Y., Zhao, C., Xu, Y., & Nie, C. (2025). Year-over-year developments in financial fraud detection via deep learning: A systematic literature review. arXiv. https://arxiv.org/abs/2502.00201

Gatla, T. R. (2024). Machine learning in credit risk assessment: Analyzing how machine learning models are transforming the assessment of credit risk for loans and credit cards. ResearchGate.https://www.researchgate.net/publication/380732388

Ghosh, S., & Dey, L. (2022). Credit risk modeling using support vector machines: A comparative approach. Decision Analytics Journal, 3, 100020

Hernandez Aros, L., Bustamante Molano, L. X., Gutierrez-Portela, F., & Moreno Hernandez, J. J. (2024). Financial fraud detection through the application of machine learning techniques: A literature review. Humanities and Social Sciences Communications, 11, 36. https://www.nature.com/articles/s41599-024-03606-0

Hu, T. (2025). Financial fraud detection system based on improved random forest and gradient boosting machine (GBM). arXiv. https://arxiv.org/abs/2502.15822

Khoshgoftaar, T. M., & Walauskis, M. A. (2025). Busted! Engineers revolutionize fraud detection with machine learning. Florida Atlantic University News. https://www.fau.edu/newsdesk/articles/machine-learning-fraud-detection.php

Kim, Y. J., & Lee, J. H. (2021). Financial fraud detection using data mining and machine learning algorithms: A case study. Expert Systems, 38(4), e12736.

Laitinen, E. K. (2021). Predicting a corporate credit analyst's risk estimate by logistic and linear models. International Review of Financial Analysis, 77, 101812.

Li, X., Wang, S., & Zhang, J. (2022). Credit scoring using machine learning techniques: A survey. Information Fusion, 75, 29–53.

Liu, J., Zhou, M., & He, Y. (2021). Hybrid ensemble learning model for financial fraud detection. Expert Systems with Applications, 184, 115515.

Minati, R., & Hema, D. (2025). Quantum powered credit risk assessment: A novel approach using hybrid quantum-classical deep neural network for row-type dependent predictive analysis. arXiv. https://arxiv.org/abs/2502.07806

Mohammed, M. A., Kothapalli, K. R. V., Mohammed, R., & Pasam, P. (2024). Machine learning-based real-time fraud detection in financial transactions. ResearchGate. https://www.researchgate.net/publication/381146733

Nahar, V., & Mishra, S. (2023). Fraud detection in banking sector using machine learning algorithms. International Journal of Advanced Computer Science and Applications, 14(2), 267–274. https://doi.org/10.14569/IJACSA.2023.0140234

Naik, K. S. (2021). Predicting credit risk for unsecured lending: A machine learning approach. arXiv. https://arxiv.org/abs/2110.02206

Ramos González, M., Partal Ureña, A., & Gómez Fernández-Aguado, P. (2023). Forecasting for regulatory credit loss derived from the COVID-19 pandemic: A machine learning approach. Research in International Business and Finance, 64, 101907. https://doi.org/10.1016/j.ribaf.2023.101907

Rodríguez Barrero, M. S., & Hernández, J. J. M. (2024). Evaluating machine learning algorithms for financial fraud detection. Mathematics, 13 (4), 600. https://www.mdpi.com/2227-7390/13/4/600

Roy, J. K., & Vasa, L. (2025). Transforming credit risk assessment: A systematic review of AI and machine learning applications. ResearchGate. https://www.researchgate.net/publication/388221989

Sharma, A., & Patel, D. (2024). Application of explainable AI in detecting anomalies in financial transactions. Procedia Computer Science, 218, 789–794.

Vallarino, D. (2025). Detecting financial fraud with hybrid deep learning: A mix-of-experts approach to sequential and anomalous patterns. arXiv. https://arxiv.org/abs/2504.03750

Vari Veedi, V. V. (2025). Predictive analytics in financial risk assessment. International Research Journal of Modern Engineering and Technology and Science, 3(3).https://www.irjmets.com/uploadedfiles/paper/issue_3_march_2025/69202/final/fin_irjmets1742538940.pdf

Zhang, Y., Xu, Y., & Fang, Y. (2023). A comparative study of deep learning models for credit card fraud detection. Journal of Financial Crime, 30(1), 19–33.

Zhao, K., & Wang, L. (2024). A robust deep learning framework for automated fraud detection in online payments. Financial Innovation, 10(1), 49.