Hybrid Explainable AI for Financial Risk Prediction: Integrating Ensemble Learning, CNN–LSTM Networks, and Financial Sentiment Analytics

Dr. Yogesh Kumar Jain
Postdoctoral Researcher, Lincoln University College, Malaysia Professor & Deputy Dean, School of Management, IILM University, Greater Noida, India
Pdf.yogesh@lincoln.edu, dryoge@gmail.com

Abstract

The paper tackles the issue of increasing complexity associated with risk assessment in financial transactions due to limitations related to conventional machine learning approaches that cannot handle nonlinearity, context-based sentiment information, and imbalanced classes. Thus, an AI model is proposed for predicting financial risks with machine learning and deep learning algorithms utilizing both structured and unstructured data sources. In particular, structured information was gathered from the "Home Credit Default Risk" dataset consisting of 307,511 samples and 122 financial features, and unstructured information was analyzed by employing FinBERT for sentiment analysis of financial news and social media content.

Data preprocessing included handling missing values, applying label encoding and MinMax scaling, creating feature fusion, and balancing classes using the SMOTEENN method. Ensemble machine learning approaches such as XGBoost, LightGBM, and Random Forest were used through voting and stacking classification techniques, while the deep learning technique was applied by implementing CNN-LSTM architecture. The research was evaluated according to various measures, which include accuracy, precision, recall, F1-score, ROC-AUC, confusion matrix, and explainable artificial intelligence tools. According to experimental results, the best-performing classifier based on the above-mentioned measures was the stacking classification approach with an ROC-AUC of 0.832 and an F1 score of 0.61, which exceeded the CNN–LSTM model in terms of ROC-AUC of 0.821 and F1 score of 0.57. Sentiment-based context features improved predictive performance, yet not significantly.

This paper digs into how artificial intelligence has been used to fight financial fraud, looking at studies from 2015 to 2025 using a systematic review approach based on PRISMA. With digital payments, online banking, DeFi, and algorithmic lending on the rise, fraudsters keep finding new angles. The authors sifted through 245 focused studies, zeroing in on how machine learning, deep learning, graph learning, hybrid AI, explainable AI, and federated learning are being put to work against all kinds of financial scams. They didn't just stick to one type of fraud; the review covers credit card scams, anti-money laundering, insurance fraud, banking fraud, crypto crimes, false financial statements, and shady DeFi schemes. Along the way, they mapped out a clear taxonomy of AI approaches, built a matrix to synthesize all that literature, and pinpointed where research is still falling short. They also stitched together a framework that ties in data realism, explainability, governance, choosing the right model, and handling deployment headaches.

What's interesting is the shift after 2019: researchers started moving away from basic machine learning and dove into graph-based models, deep sequential algorithms, hybrid setups, and, more recently, explainable AI. That said, even with all these cool advances, the field still struggles with issues like keeping models reliable over time, setting standard benchmarks, handling changes in fraud patterns (concept drift), keeping data private, and nailing down what makes an explanation actually helpful.

In the end, the paper makes it clear—moving forward, it's not just about making models accurate. Fraud detection needs to double down on transparency, fair practices, explainability, and being able to scale up

operations. This review should help researchers, industry folks, policymakers, and financial institutions navigate the push for smarter, more dependable AI-powered fraud defenses.

Keywords

Financial Risk Assessment; Hybrid AI; Machine Learning; Deep Learning; CNN-LSTM; Ensemble Learning; XGBoost; LightGBM; Random Forest; SMOTEENN; Credit Risk Prediction; Explainable AI

1. Introduction

It becomes a growing problem for the financial industry to assess the creditworthiness of their clients because of the fast-changing digital finance landscape and economic instability alongside the heterogeneous nature of available data. The traditional methods used to estimate financial risk, such as logistic regression and decision-making rules, fail to deal with nonlinear interactions, data imbalance, and unstructured data such as news and financial sentiments from social media. The goal of the current research is the development of an integrated framework based on artificial intelligence that will be able to merge both structured data such as financial records and unstructured data like sentiments from social media in order to enhance the performance of predictions. Financial systems have been experiencing fast changes associated with their digitalization processes, leading to an upsurge in the number of possible cases of financial fraud in such areas as digital banking, electronic transactions, mobile banking, insurance systems, DeFi solutions, and corporate financial reports. Existing rule-based systems are becoming outdated when dealing with new forms of financial fraud.

The artificial intelligence technology has become one of the most important means of recognizing the existing complex structures behind frauds in financial data characterized by heterogeneity and complexity. Machine Learning, Deep Learning, Graph Analytics, and Explainable AI approaches have played a significant role in the development of research on financial fraud detection. The goal of the paper is to provide a systematic analysis of AI-based research on financial fraud detection over the past ten years.

2. Related Work

In the literature review, we find that prior studies in financial risk modeling have used ML algorithms like SVMs, random forests, and logistic regression. Prior studies on deep learning-based models such as LSTM, CNN, and autoencoders were considered for financial risk modeling. In recent times, it was proven through studies that hybrid methods of applying ML and DL algorithms gave better results than individual models. Studies have also been done to use sentiment analysis and NLP techniques in bankruptcy prediction and financial forecasting. The problem with most of the previous research was that they had problems like insufficient data balance, lack of explainability, and inadequate combination of structured and unstructured data sets.

In this study, we have tried to overcome these limitations by using:

- Ensemble learning techniques
- CNN–LSTM deep learning architecture
- SMOTEENN imbalance correction
- Feature fusion using sentiment analysis
- Explainable AI using SHAP analysis

Most prior reviews focused on particular fraud domains:

- Detection of credit card fraud Explainable AI in finance Analytical Banking Fraud: Fraud Detection using Machine Learning

None of the previous reviews had:

- Comprehensive multi-domain coverage Integrated framework design Evaluation of explainability
- Evaluation focused on governance Comparative analysis of ML, DL, graph learning, and hybrid AI
- This review contributes to the existing body of knowledge by reviewing 245 publications on different domains of fraud. Including the PRISMA framework
- Proposing a task-oriented AI framework
- Analyzing explainability and governance
- Offering practical recommendations for implementation

3. Key Contributions

- Built a mixed AI structure combining ML and DL to assess financial risk.
- Merged structured financial data with unstructured text-based sentiment data.
- Used SMOTEENN to deal with high class imbalance.
- Developed voting and stacking classifiers combining XGBoost, LightGBM, and Random Forest.
- Designed CNN–LSTM to address interaction between features.
- Used SHAP in Explainable AI to assess the influence of features.
- Showcased the benefits of explainable AI and the positive impacts of governance and interpretability on AI performance.
- Conducted a PRISMA-based review on 245 studies from 2015 to 2025.
- Created a task-oriented classification of AI in financial fraud detection.
- Analyzed ML, DL, graph-based learning, and hybrid AI.
- Assessed explainable AI and the impacts of governance.
- Explored constraints of datasets, concept drift and privacy, and operational deployment.
- Provided a holistic constructive model of AI performance, governance, and interpretability.
- Outlined the next steps for the design of fraud detection systems that are scalable and trustworthy.

4. Methods, Experiments & Results

a) Methodology

Dataset

The Home Credit Default Risk dataset included:

- 307,511 loan applications - 122 features - 1 binary categorical target (default or non-default) Other (unstructured) data were included from:

- Financial news posting - Social media post - NewsAPI.org

Preprocessing

The preprocessing steps defined included:

- Removal of features with too many missing values
- Median value imputation for missing numerical features
- Label encoding for categorical features
- MinMax scaling

- Sentiment analysis using FinBERT
- Combining structured data and sentiment features

Imbalance Handling

To handle class imbalance, SMOTEENN was applied, consisting of:
- SMOTE and
- Edited Nearest Neighbor (ENN)

Machine learning models

The ML process involved:
- XGBoost Classifier
- LightGBM Classifier
- Random Forest Classifier
- Voting Classifier
- Stacking Classifier with Logistic Regression as a meta-learner

Deep Learning Model

The DL structure was made up of:
- 1D CNN - MaxPooling - LSTM - Dense Sigmoid

Evaluation Metrics

The metrics used to assess performance included:
- Accuracy - Precision - Recall - F1-score - ROC-AUC - Confusion Matrix - SHAP and interpretability

Study Design

The study design for the systematic review relied on the PRISMA 2020 standards.

Information Sources

- Scopus - Web of Science Scope 2015 - 2025

b) Methodology

Search terms included:

- Financial Fraud - Credit Card Fraud - Banking Fraud - Anti-Money Laundering - Cryptocurrency Fraud
- Machine Learning - Deep Learning - Graph Neural Networks - Explainable AI

Study Criteria

- Studies for AI-based Fraud Detection

- Peer-reviewed articles in the English language - Studies with experiments and/or comparisons
- Studies that describe the datasets and assessment criteria

Criteria for Elimination

- Studies on cybercrime that are not financially related - Editorials and opinion articles
- Replications - Studies that are poorly explained methodologically

Remaining Studies—Records after initial search: 1,500 - Records after review: 245

AI Techniques

Machine Learning - Logistic Regression - Support Vector Machines, Random Forest, XGBoost, Decision Trees

Deep Learning - CNNs - LSTMs - Autoencoders - Transformers

- Recurrent Neural Networks

Graph Learning - Graph Convolutional Networks - Graph Attention Networks - Heterogeneous Graph Learning

Hybrid AI - ML + DL - NL Fusion

- Federated Learning + Explainable AI

Explainable AI - SHAP – LIME - AI with Attention - Counterfactuals

5. Result and Discussion

The study showed that using machine learning models together, also known as ensemble learning approaches, really helped to improve the performance of financial risk prediction. The stacking classifier model did the best with results like

*** ROC-AUC: 0.832 * F1-score: 0.61.**

The CNN-LSTM model also did well with

*** ROC-AUC: 0.821 * F1-score: 0.57.**

Some key things were noticed.

* Using sentiment-based features helped us understand the context better.
* SMOTEENN really helped us detect the minority class.
* How long someone has been employed, their income level, age, the type of organization they work for and their job were all things that helped us predict financial risk.
* Using artificial intelligence helped us be more transparent and understand the results better.

It was also found out that younger people who borrowed money, people with skills, and people who had unstable jobs were more likely to default on their loans.

The review identified several important trends:

Main Findings

1. The fraud domain that is benchmarked the most is still credit card fraud.
2. Deep learning adoption accelerated after 2019.
3. High performance of graph learning on AML and cryptocurrency fraud detection.
4. Hybrid AI systems showed better adaptability within complex fraud environments.
5. Explainable AI has become more important for regulatory compliance.

Dataset Difficulties: The review identified key challenges, including: • Strongly imbalanced classes

• Real-world datasets are limited

• Privacy restrictions

• Concept drifts

- Absence of temporal validation

• Bad standardization of benchmarks The Explainability Challenge

Many studies used SHAP and LIME but were missing. • Stability analysis • Human-centered evaluation

• Governance-oriented assessment

• Validation of deployment in the real world Operational Problems

The review concluded that real-world fraud detection systems require the following:

• Prediction with low latency

• Easily scalable deployment

• Privacy protection
• Institutional governance
• Resilience against changing fraud patterns

6. Discussion

Our framework did a job of combining how well the model predicted things with how well we could understand the results. By using both unstructured data, we got a better understanding of the context. Using models together also helped reduce bias and made the results more robust. The explainable artificial intelligence also helped people trust the results and made sure we followed the rules. Even though the CNN-LSTM model was good, at finding patterns the ensemble machine learning methods still worked better for financial datasets. The study also showed that it is really important to balance how well the model performs with how transparent it is, especially when it comes to big financial decisions. The financial risk prediction performance of the learning approaches was really improved. The ensemble learning approaches and the stacking classifier model were very effective.

The study notes that predictive accuracy is not enough in high-stakes financial environments. Future fraud detection systems will need to balance the following:
• Precision • Explainability
• Equity • Privacy • Governance • Expandability

The review also pointed out that no single AI paradigm dominates all fraud domains. Rather, the suitability of AI models depends on: • Data type
• Fraud pattern
• Deployment platform
• Regulatory compliance
• Computational limitations
The most promising future directions appear to be hybrid AI systems that combine ML, DL, graph learning, and explainability mechanisms.

7. Conclusion

The study successfully addresses the limitations of traditional financial risk assessment systems.

- Hybrid AI models integrating ML and DL demonstrated enhanced predictive capability.
- Ensemble learning methods performed better than deep learning alone.
- Contextual prediction accuracy was improved by unstructured data based on sentiment.
- SMOTEENN is a good solution for problems of class imbalance.
- Interpretability and transparency were enhanced with explainable AI methods.
- Future work may include real-time data integration, transformer-based architectures, and cross-institutional validation.
- AI has revolutionized the detection of financial fraud in many industries.
- Machine learning continues to rule for structured tabular fraud detection.
- Deep learning and graph learning are becoming increasingly important for complex fraud scenarios.
- Explainable AI is a necessity for trusted financial systems.
- Privacy, benchmark realism, concept drift, and deployment governance remain significant challenges.
- Hybrid AI frameworks have the most potential for the future.
- Future directions include evaluation of explainability, privacy-preserving learning, design of multimodal benchmarks, federated learning and adaptive fraud monitoring systems.

References

- Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. Journal of Network and Computer Applications, 68, 90–113. https://doi.org/10.1016/j.jnca.2016.04.007
- Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models. arXiv preprint arXiv:1908.10063.
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities, and challenges toward responsible AI. Information Fusion, 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012
- Batista, G. E. A. P. A., Prati, R. C., & Monard, M. C. (2004). A study of the behavior of several methods for balancing machine learning training data. ACM SIGKDD Explorations Newsletter, 6(1), 20–29.
- Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324
- Brown, I., & Mues, C. (2012). An experimental comparison of classification algorithms for imbalanced credit scoring data sets. Expert Systems with Applications, 39(3), 3446–3453.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.
- Doshi-Velez, F., & Kim, B. (2017). Towards a rigorous science of interpretable machine learning. arXiv preprint arXiv:1702.08608.
- Dou, Y., Liu, Z., Sun, L., Deng, Y., Peng, H., & Yu, P. S. (2020). Enhancing graph neural network-based fraud detectors against camouflage. Proceedings of the 29th ACM International Conference on Information and Knowledge Management, 315–324.
- Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654–669.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 30.
- Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring. European Journal of Operational Research, 247(1), 124–136.
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.
- Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. Decision Support Systems, 50(3), 559–569.
- Yang, X., Liu, C., & Wang, Y. (2020). Financial sentiment analysis based on FinBERT and deep learning techniques. Expert Systems with Applications, 159, 113567.
- Zhang, Y., Aggarwal, C., & Qi, G. J. (2019). Stock price prediction via discovering multi-frequency trading patterns. Proceedings of the ACM International Conference on Information and Knowledge Management.