Proceedings of ICFTBA 2026 Symposium: Strategic Management & Business Analytics: Data-Driven Performance Models
DOI: 10.54254/2754-1169/2026.LD38133

Artificial Intelligence Application in Financial Customer Fraud Detection

Yuhan Sun

Wycombe Abbey School Hong Kong, Hong Kong, China
jgao78768@gmail.com

Abstract. With the rapid digitization of financial services, the online transactions have become prevalent, accompanied by a growing number of financial fraud incidents that cause substantial losses to financial institutions and individual users. Traditional rule-based fraud detection methods suffer from low efficiency, poor timeliness and limited adaptability to hidden and emerging fraud behaviors. Benefiting from the powerful self-learning capability, high-dimensional data processing performance and dynamic feature mining advantages of artificial intelligence algorithms, intelligent detection technologies have become the mainstream solution for modern financial fraud risk control. This paper systematically investigates the application of typical machine learning and deep learning algorithms in customer financial fraud detection. It summarizes the practical advantages of different models, analyzes the core application challenges including the data imbalance and poor model interpretability, and proposes targeted future development directions. The research aims to provide the theoretical references for the optimization and industrial application of intelligent financial fraud detection systems.

Keywords: Customer fraud detection, artificial intelligence, finance

1. Introduction

With the rapid development of digital financial services, online transactions have become mainstream in daily financial activities. However, this trend also leads to a continuous rise in customer financial fraud cases, bringing huge economic losses to both financial institutions and individual users. Traditional fraud detection methods mainly rely on fixed manual rules and historical fraud records. These static and experience-dependent methods can hardly cope with increasingly flexible and hidden modern fraud behaviors, and often produce high error rates and delayed detection results [1]. In this context, artificial intelligence technology has gradually become a core tool for financial customer fraud detection. Benefiting from the strong self-learning ability, high-dimensional data processing capability and dynamic feature mining characteristics of artificial intelligence algorithms, AI-based detection tools show unique advantages in processing massive transaction data and identifying abnormal user behaviors [2].

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2754-1169/2026.LD38133

2. Method

This chapter systematically elaborates on mainstream machine learning and deep learning algorithms applied in financial customer fraud detection. These methods are widely adopted to mine hidden risk features in massive transaction data, identify abnormal customer behaviors, and realize accurate classification of fraudulent and legitimate financial activities. Different algorithms have distinct advantages in data adaptability, feature extraction capability and detection efficiency, which are selected and optimized according to the characteristics of financial fraud scenarios.

2.1. Traditional machine learning algorithms

Traditional machine learning algorithms are the basic technical means of financial fraud detection. They rely on manually constructed transaction and user feature indicators for model training. With simple operation, low hardware resource consumption and good interpretability, these algorithms are suitable for processing structured financial data and have long been applied in real-time transaction risk identification.

2.1.1. Decision tree

Decision tree is a basic supervised classification algorithm with intuitive logical reasoning process and strong interpretability. The model constructs a tree-like decision structure through feature splitting, divides data samples into different branches according to feature thresholds, and finally completes the classification of fraud risks [3]. In financial customer fraud detection, decision tree takes customer basic information, historical transaction records and real-time transaction behavior as splitting features, and forms clear risk judgment rules layer by layer.

The biggest advantage of decision tree is that it does not require complex data normalization and preprocessing, and can directly process discrete and continuous financial data. The generated decision rules are consistent with manual risk judgment logic of financial institutions, which is convenient for business staff to understand and verify. However, the single decision tree model has obvious limitations. It is prone to overfitting when processing high-dimensional massive financial data, and is sensitive to local abnormal samples, resulting in poor stability of fraud detection results. Therefore, decision tree is mostly used as a baseline comparison model in current fraud detection research, or as a basic component of ensemble algorithms [4].

2.1.2. Random forest

Random forest is a typical ensemble machine learning algorithm composed of multiple independent decision tree classifiers. The final classification result is determined by the voting mechanism of all decision trees, which effectively overcomes the defects of single model such as low accuracy and overfitting tendency [5]. In financial customer fraud detection research, Wang and Zhang applied random forest model to credit card transaction fraud identification. The research selected multi-dimensional feature variables including customer transaction frequency, transaction amount, transaction location deviation and account login time difference to construct a feature system, and completed model training based on real financial transaction datasets.

Different from single traditional algorithms, the core innovation of random forest in fraud detection lies in its random sampling mechanism. It carries out random sampling of data samples and feature variables during model construction, which enhances the generalization ability of the

DOI: 10.54254/2754-1169/2026.LD38133

model. Meanwhile, the model can output the importance score of each feature, helping financial researchers and practitioners screen core risk indicators and eliminate invalid interference features [6]. Experimental results show that compared with single decision tree and logistic regression models, random forest significantly reduces the false detection rate of financial fraud, and maintains stable detection performance under unbalanced sample distribution of legitimate and fraudulent transactions [5].

2.2. Deep learning algorithms

With the explosive growth of financial big data and the increasingly concealed fraud means, traditional machine learning relying on manual feature construction is difficult to capture complex nonlinear correlation and implicit abnormal patterns in transaction data. Deep learning algorithms realize automatic feature extraction and hierarchical feature learning through multi-layer network structure, which can adapt to high-dimensional, unstructured and sequential financial data, and has become the mainstream technology of high-precision financial fraud detection [7].

2.2.1. MLP

Multi-Layer Perceptron (MLP) is a classic fully connected deep learning model, composed of input layer, hidden layer and output layer. Through forward propagation and back propagation algorithm, MLP continuously optimizes network weights and thresholds, and fits the complex mapping relationship between financial transaction features and fraud risk labels [8]. In existing research on customer fraud detection, scholars take standardized structured financial data such as user account attributes, transaction behavior frequency and transaction amount as MLP input features. The multi-layer hidden layer realizes deep mining of potential correlation between multi-dimensional features, and the output layer completes binary classification of fraudulent and legitimate transactions [8].

MLP makes up for the deficiency of traditional algorithms in nonlinear feature fitting. It can automatically mine the implicit correlation between multi-dimensional transaction features without manual feature screening, and effectively identify potential fraud behaviors that conform to normal single features but have abnormal overall patterns. Compared with ensemble machine learning models, MLP has higher detection accuracy for complex and variable fraud scenarios, and is suitable for large-scale financial transaction batch detection tasks.

2.2.2. ANN

Artificial Neural Network (ANN) simulates the information transmission and processing mechanism of human brain neurons, and is the basic prototype of all deep learning models. Different from the fixed logical structure of traditional algorithms, ANN establishes adaptive nonlinear mapping rules through neuron connection weights, and has strong self-learning and self-adaptive capabilities [9]. In financial fraud detection research, Jurgovsky et al. adopted ANN model to identify customer credit fraud behaviors. The model took user credit record, repayment behavior and asset status as input features, and realized accurate identification of default and fraud risks [8].

The unique advantage of ANN is its strong fault tolerance. It can still maintain stable detection performance when individual financial data is missing or abnormal. Through iterative training of massive historical transaction data, the model can continuously summarize fraud risk laws and adapt to the updated fraud means of criminals. However, the shallow ANN model has limited feature

DOI: 10.54254/2754-1169/2026.LD38133

extraction ability and is easy to fall into local optimal solutions, so it is usually combined with other optimization algorithms to improve detection stability in practical applications.

2.2.3. CNN

Convolutional Neural Network (CNN) is a feed-forward deep learning model with convolution and pooling structures, which is widely used in high-dimensional data feature extraction and pattern recognition. Different from fully connected networks, CNN extracts local key features of data through convolution kernel sliding scanning, and reduces data dimension and parameter quantity through pooling operation, avoiding model overfitting [7]. In financial customer fraud detection, researchers convert sequential transaction data and multi-dimensional user behavior data into matrix feature graphs, and use CNN to extract spatial and local correlation features of data.

CNN performs excellently in identifying hidden and disguised fraud behaviors. It can capture subtle abnormal feature differences that are difficult to distinguish by traditional algorithms and fully connected neural networks, such as abnormal transaction interval rules and irregular multi-terminal login behaviors. At the same time, the sparse parameter characteristic of CNN greatly improves the operation efficiency, which can meet the real-time detection requirements of high-frequency financial transactions of banks and payment institutions, and has high practical application value in industrial scenarios.

3. Discussion

Based on the above analysis of mainstream detection algorithms, this chapter summarizes the core challenges faced by current AI-based financial customer fraud detection technology in practical application, and puts forward targeted future development prospects, so as to provide reference for the optimization and industrial promotion of financial fraud detection models.

3.1. Challenges

3.1.1. Imbalanced distribution of financial data

Data imbalance is the most prominent problem restricting the performance of fraud detection models. In actual financial business scenarios, legitimate customer transactions account for more than 99% of total transaction data, while fraudulent transaction samples are extremely scarce [6]. The long-term unbalanced sample distribution will lead to model training bias. The model tends to learn the features of legitimate transactions, ignores the rare fraud feature laws, and finally leads to high false negative rate. A large number of hidden fraud behaviors cannot be effectively identified, which brings potential property risks to financial institutions and customers. Although some resampling optimization methods have been applied, they are prone to over-simulation of fraud samples and distortion of feature distribution, affecting the generalization ability of the model.

3.1.2. Poor interpretability of deep learning models

Most high-precision deep learning fraud detection models belong to "black box models". The internal feature extraction logic and risk judgment rules are opaque [10]. Financial fraud detection is a high-risk regulatory business, which requires clear and traceable risk identification basis. However, CNN, MLP and other models cannot output explicit risk feature contribution and judgment logic. When the model outputs fraud risk early warning, practitioners cannot accurately

DOI: 10.54254/2754-1169/2026.LD38133

judge the core abnormal factors leading to risk identification, which makes it difficult to verify and correct model errors. Poor interpretability also hinders the compliance application of AI detection models in the strictly regulated financial industry.

3.1.3. Low adaptability to new fraud patterns

Financial fraud means show obvious dynamic and innovative characteristics. Criminals continuously update fraud strategies according to existing detection rules, and generate new disguised fraud behaviors that are different from historical fraud patterns [4]. Current detection models rely on historical fraud sample data for training, and can only effectively identify existing and similar fraud modes. For unknown new fraud behaviors, the model has serious cold start problem, low recognition accuracy and poor dynamic adaptability. It is difficult to realize real-time iteration and active early warning of fraud risk rules, resulting in obvious lag in risk prevention and control.

3.2. Future prospects

3.2.1. Construction of balanced and multi-scene financial dataset

In view of the problem of data imbalance, future research will focus on optimizing data enhancement and sample balancing technology. On the basis of traditional SMOTE sampling algorithm, combined with generative models to generate high-quality simulated fraud samples consistent with real business characteristics [6], so as to balance the proportion of positive and negative samples. Meanwhile, financial institutions will establish multi-scene unified fraud datasets covering credit card transactions, account transfer and payment fraud, realize data sharing under the premise of ensuring user privacy and data security, and provide high-quality data support for model training.

3.2.2. Integration of interpretable technology and hybrid model optimization

Improving model interpretability is the key direction of industrial application of fraud detection technology. Future research will combine explainable artificial intelligence (XAI) technology to visualize the feature extraction process and risk judgment basis of deep learning models [10]. At the same time, hybrid model architecture will be widely used, combining the high interpretability of traditional machine learning algorithms with the high accuracy of deep learning models. The complementary advantages of different models can not only ensure the accuracy of fraud detection, but also realize traceable risk judgment, meeting the compliance requirements of financial supervision [9].

3.2.3. Dynamic adaptive intelligent detection system construction

To solve the problem of poor adaptability to new fraud patterns, the future development direction is to build real-time iterative dynamic detection system. By introducing incremental learning and transfer learning algorithms, the model can automatically learn new fraud features from real-time transaction streaming data, update model parameters and risk rules in real time [7]. In addition, combined with user relationship graph analysis technology, the system can mine potential association fraud behaviors from the perspective of user network correlation, realize active prediction and early warning of unknown fraud risks, and build a full-cycle intelligent risk prevention and control system.

DOI: 10.54254/2754-1169/2026.LD38133

4. Conclusion

Digital financial transformation has greatly improved the convenience and efficiency of public financial transactions, yet it has also contributed to increasingly diversified, concealed and updated financial fraud behaviors, which poses severe challenges to traditional rule-based and manual fraud detection methods. This paper comprehensively investigates and compares the application performance of classical machine learning and mainstream deep learning algorithms in financial customer fraud detection. Traditional machine learning models represented by decision trees and random forests exhibit outstanding interpretability and stable performance in processing structured financial transaction data, which is suitable for basic risk identification and baseline comparison in financial scenarios. In contrast, deep learning techniques including MLP, ANN and CNN are capable of automatic hierarchical feature extraction, effectively capturing implicit nonlinear correlations and subtle abnormal fraud features that cannot be identified by traditional algorithms, thereby significantly improving the detection accuracy of complex and emerging fraud behaviors.

Nevertheless, the practical deployment of existing AI-based fraud detection models still faces three key limitations. First, the severe imbalance of real financial transaction data easily causes training bias and increases model false negative rates. Second, most deep learning models lack interpretability, which hinders regulatory compliance and practical error verification in financial risk control. Third, static model training based on historical samples leads to poor adaptability to newly emerging fraud patterns. In response to the above challenges, this study proposes three promising future research directions: constructing balanced and multi-scene financial datasets via data enhancement technologies, optimizing hybrid models integrated with explainable artificial intelligence to balance accuracy and interpretability, and building dynamic adaptive detection systems based on incremental and transfer learning.

In summary, artificial intelligence has irreplaceable advantages in intelligent financial fraud detection. With the continuous optimization of data quality, model structure and dynamic iteration mechanisms, AI-based risk detection technology will achieve higher accuracy, better interpretability and stronger environmental adaptability. It can not only effectively reduce financial fraud losses of institutions and users, but also provide reliable technical guarantees for the standardized, safe and intelligent development of modern digital financial systems.

References

[1] Li, H., et al. (2023). Research on traditional financial fraud detection methods and their application bottlenecks. *Financial Technology Research, 9*(2), 78-85.

[2] Chen, J., & Zhou, M. (2021). Core advantages of artificial intelligence in massive financial data mining. *Computer Engineering and Applications, 57*(18), 234-240.

[3] Zhang, L., & Liu, M. (2023). Comparative analysis of traditional classification algorithms in customer financial risk identification. *Computer Engineering and Applications, 59*(11), 234-241.

[4] Chen, J., et al. (2025). Challenges and improvements of traditional machine learning for dynamic financial fraud recognition. *Frontiers in Artificial Intelligence, 8*, 1643292.

[5] Wang, Y., & Zhang, H. (2024). Performance optimization of random forest in unbalanced credit card fraud detection. *Journal of Financial Risk Management, 13*(2), 89-102.

[6] Li, S., et al. (2025). Imbalanced data processing strategies for machine learning-based financial fraud detection. *Scientific Reports, 15*(1), 4218.

[7] Zhou, K., & Wang, L. (2024). Deep learning applications and innovations in financial big data fraud detection. *IEEE Access, 12*, 78912-78925.

[8] Jurgovsky, M., et al. (2023). Multilayer perceptron-based abnormal transaction detection for retail banking customers. *Journal of Banking and Financial Technology, 7*(3), 112-125.

DOI: 10.54254/2754-1169/2026.LD38133

[9] Xiao, Y., et al. (2025). Hybrid deep learning framework for hidden financial fraud pattern recognition. *TechRxiv Preprint*.

[10] Sun, H., et al. (2024). Explainable AI for compliant financial fraud detection: A systematic review. *International Journal of Financial Innovation, 10*(4), 56-71.