Article

Optimization of Financial Fraud Risk Identification System Based on Machine Learning

Xuanrui Zhang [1,*]

[1] College of Engineering, University of California, Berkeley, Berkeley, CA 94720, USA
[*] Correspondence: Xuanrui Zhang, College of Engineering, University of California, Berkeley, Berkeley, CA 94720, USA

Abstract: With the ongoing digital transformation of the financial sector, the landscape of financial fraud has become increasingly complex and sophisticated, presenting significant challenges to traditional fraud detection systems, which often suffer from low accuracy and delayed response times. Financial fraud risk assessment systems powered by machine learning technology can automatically detect anomalous patterns by analyzing large volumes of historical transaction data, significantly enhancing the precision and timeliness of detection. Currently, machine learning algorithms such as deep learning, decision trees, and support vector machines serve as the core tools for improving detection efficiency and predictive capability. Despite these advancements, challenges such as inconsistent or poor-quality data, model overfitting, and the selection of relevant features continue to constrain system performance and generalizability. Addressing these issues requires comprehensive optimization strategies, including advanced data preprocessing, robust feature engineering, algorithmic fine-tuning, and model validation techniques, to strengthen the system's ability to identify and anticipate fraudulent behaviors. This article proposes a series of targeted improvement measures to tackle the current limitations in financial fraud detection systems and explores the future potential of machine learning technologies to enable proactive, intelligent, and adaptive approaches to financial fraud prevention.

Keywords: financial fraud; risk identification; machine learning; deep learning; data quality; model optimization

Received: 03 October 2025
Revised: 11 October 2025
Accepted: 29 October 2025
Published: 02 November 2025

1. Introduction

In the modern financial system, robust financial fraud risk detection systems are essential for maintaining institutional security and market stability. Traditional risk assessment methods primarily rely on rule-based engines and manual judgment, which, while straightforward, often suffer from limitations in efficiency, scalability, and adaptability when confronting increasingly complex and rapidly evolving fraud schemes. These conventional approaches struggle to identify subtle or novel fraudulent behaviors, leading to delayed responses and increased operational risks for financial institutions.

With the rapid growth of digital financial transactions and the widespread adoption of online and mobile banking, the volume, velocity, and diversity of transaction data have surged dramatically. In this context, machine learning technology has emerged as a critical tool for enhancing fraud detection capabilities. By leveraging advanced data processing, pattern recognition, and predictive analytics, machine learning models can automatically uncover hidden correlations and anomalous patterns in large-scale datasets that traditional methods often overlook. Algorithms such as decision trees, support vector machines, and deep learning architectures enable the system to learn from historical data, adapt to evolving fraud tactics, and improve detection accuracy over time.

Moreover, machine learning-based systems offer the potential for real-time risk assessment, allowing financial institutions to proactively prevent fraud before significant losses occur. Beyond detection, these systems can assist in risk scoring, transaction prioritization, and resource allocation, thereby optimizing operational efficiency and decision-making. However, the effectiveness of machine learning approaches depends heavily on data quality, feature selection, and model design, underscoring the importance of comprehensive optimization strategies.

Overall, integrating machine learning into financial fraud risk detection not only addresses the shortcomings of traditional methods but also establishes a more intelligent, adaptive, and scalable framework capable of responding to the dynamic landscape of modern financial threats.

2. Overview of Financial Fraud Risk Identification System

2.1. Definition of Financial Fraud Risk Identification System

The financial fraud risk identification system leverages advanced data processing, pattern recognition, and artificial intelligence learning technologies to implement real-time monitoring and unconventional transaction screening of financial activities, aiming to detect potential fraudulent behavior. The system collects and analyzes large volumes of transaction data and automatically identifies possible risk points by referencing historical behavior patterns and account attributes. It employs techniques such as classification algorithms, cluster analysis, and outlier detection to filter abnormal behavior patterns from multidimensional data.

By comprehensively evaluating factors such as transaction content, account dynamics, and user behavior, the system is able to identify irregular trading activities effectively. It can detect both known fraud patterns and adapt flexibly to emerging fraudulent methods, demonstrating strong real-time response and adjustment capabilities. As a core component of modern financial anti-fraud frameworks, this system provides robust risk control tools for banks, payment platforms, and other financial institutions, enhancing their capacity to maintain operational security and integrity.

2.2. Importance of Financial Fraud Risk Identification System

With the increasing complexity, concealment, and sophistication of financial fraud, traditional detection methods often struggle to keep pace, leaving financial institutions vulnerable to greater risks and potential losses. The financial fraud risk detection system addresses these challenges by leveraging advanced data analysis and processing techniques to uncover hidden fraudulent activities, thereby ensuring the security and reliability of financial transactions.

This system is capable of processing vast amounts of transactional information quickly and efficiently, enabling rapid identification of diverse fraudulent behaviors, including identity theft, unauthorized fund transfers, fictitious transactions, and other illicit activities. By significantly improving the accuracy and response speed of fraud detection, the system helps financial institutions protect customer assets, maintain regulatory compliance, and safeguard their reputations. Moreover, the deep integration of technology and data analytics in fraud detection not only strengthens operational risk management but also enhances public trust in financial institutions, mitigating negative social and economic impacts.

3. Current Status of Financial Fraud Risk Identification System Based on Machine Learning

3.1. Data Quality Issues Leading to Inaccurate Models

In financial fraud risk identification systems, the quality of data directly determines the accuracy and reliability of predictive models. At present, data quality issues are primarily reflected in incomplete records, missing key values, noisy data, and inherent biases. Improper storage of financial transaction data, transmission errors, or inconsistent data entry may result in the loss of critical details, which can hinder the model's comprehensive understanding of fraudulent behavior during the training phase.

Biases in historical data, such as an overrepresentation of certain transaction types or neglect of others, negatively affect the generalizability of models, weakening their ability to detect new or evolving forms of fraud. This issue is particularly critical in real-time trading scenarios, where data changes rapidly and exhibits high complexity. Without stable and high-quality data, both the predictive accuracy and response speed of the model can be severely compromised, limiting its effectiveness in practical applications.

3.2. Complexity of Feature Selection and Processing

The selection and processing of features play a crucial role in financial fraud detection, but the process is highly complex and challenging. Financial data encompasses multiple levels, including transaction amount, time, counterparties, account characteristics, and user behavioral patterns. The interrelationships among these features are often subtle and difficult to quantify directly. Moreover, fraudulent behavior is inherently secretive and variable, requiring careful consideration of the core elements on which different fraud strategies rely.

Currently, most systems depend on manual feature selection or statistical feature processing methods, which are inefficient for handling large-scale, complex datasets. Inadequate feature selection and processing often limit the model's ability to accurately identify potential fraud risks, particularly when confronted with sophisticated or evolving fraudulent activities. Enhancing feature engineering and developing automated or adaptive selection methods are therefore critical for improving detection performance.

3.3. Poor Interpretability of the Model

Machine learning-based financial fraud detection systems, particularly those involving deep neural networks or complex ensemble models, frequently face significant challenges in interpretability. While these models may achieve high predictive accuracy, they often function as "black boxes" for both ordinary users and financial professionals. When a system flags a transaction as potentially fraudulent, it is often difficult to specify which features or data attributes contributed to the decision.

This lack of transparency poses challenges for financial institutions in meeting regulatory requirements, providing convincing explanations, and maintaining the trust of business personnel. In operational contexts such as compliance monitoring and auditing, interpretability and explainability are essential, yet current systems fall short of these demands. Addressing model interpretability is therefore critical to ensure regulatory compliance, improve operational transparency, and foster confidence in automated fraud detection.

3.4. Insufficient Generalization Ability of the Model

The generalization ability of a predictive model is fundamental to the effectiveness of financial fraud early warning systems. Although many popular models perform well on historical training data, their effectiveness often declines when confronted with new or previously unseen fraudulent methods. Conventional machine learning techniques are prone to overfitting, particularly when financial fraud evolves under different market conditions or temporal contexts.

Models may exhibit excellent performance on existing datasets but fail to adapt to dynamic fraud strategies in real-world applications. With continuous changes in financial markets and the rapid evolution of fraudulent tactics, many models struggle to maintain timely adaptability, resulting in decreased accuracy of early warning systems. Ensuring robust generalization and stability under varying conditions is therefore a critical challenge for current financial fraud detection systems, demanding ongoing optimization of model design, training methods, and data management practices.

4. Optimization Strategy for Financial Fraud Risk Identification System Based on Machine Learning

4.1. Data Cleaning and Filling

Data cleaning and filling are fundamental steps in optimizing financial fraud risk identification systems, aimed at enhancing the completeness, consistency, and reliability of datasets through advanced data preprocessing techniques. Data cleaning typically involves removing redundant or duplicate entries, correcting outliers, and standardizing data formats to ensure that numerical inputs to the model are accurate and consistent.

When handling missing data, adaptive strategies such as weighted estimation or distribution-based imputation can be employed to fill gaps accurately, reducing errors introduced by conventional filling methods. For instance, in credit card transaction fraud detection, missing transaction amount information often results from incomplete data entry or delays in synchronization across systems. To address this issue, intelligent imputation methods can integrate a user's historical transaction records along with the consumption patterns of similar users to estimate missing values.

A comprehensive filling strategy can consider multiple dimensions, such as the user's transaction frequency, time intervals between transactions, and the typical fluctuation range of transaction amounts, thereby producing data that closely approximates the original distribution. Additionally, the imputation process can be further optimized by incorporating the user's specific transaction context, ensuring that the filled data reflects realistic trading behaviors. Such enhanced data cleaning and filling not only improves the model's predictive accuracy but also strengthens its robustness against anomalies and irregularities in real-world datasets.

The filling of missing data can be completed using the following formula:

$$X_{imputed} = Xi + \lambda \times (X_{\text{mean}} - Xi) \quad (1)$$

Among them, $X_{imputed}$ represents the filled data, Xi is the original data where the missing values are located, λ is the interpolated weight coefficient, and X_{mean} is the mean of the feature. This formula adjusts the original data by using the mean and weight coefficients to fill in missing values, thereby minimizing the impact on model training while maintaining the overall distribution characteristics of the data. The cleaning and filling of data have successfully removed the impact of noise and missing data, ensuring the quality of the dataset used for model training and thereby improving the reliability of the operation of financial fraud recognition systems.

4.2. Automatic Feature Selection Method

Selecting appropriate features is essential for improving the performance and predictive accuracy of financial fraud risk identification systems. Automatic feature selection methods enable the system to identify features with high representativeness and predictive value while removing redundant or irrelevant attributes. This not only simplifies model structure but also mitigates the risk of overfitting. Common strategies for feature selection can be categorized into three approaches: screening, combination, and embedded methods.

The screening strategy primarily relies on statistical tests, such as the chi-square test, to evaluate individual features independently. The combination strategy employs algorithms like Recursive Feature Elimination (RFE) to iteratively construct optimal feature subsets. Embedded methods integrate feature selection directly into the model training process, using techniques such as random forests, gradient boosting decision trees (GBDT), and L1 regularization (Lasso).

For example, a financial institution faced challenges in optimizing its fraud detection system due to the high dimensionality of transaction data, which included transaction volume, operation time, location information, device identifiers, and other attributes. The institution applied L1 regularization to eliminate redundant or non-informative features. Following this process, key indicators such as abnormal fluctuations in transaction amounts, frequent late-night trading behavior, and repeated logins from multiple devices were retained, while less relevant information, such as device screen resolution and millisecond-level transaction timestamps, was removed.

By focusing on core features, the system preserved the most informative patterns for fraud recognition, reducing noise during the training phase and significantly enhancing both the accuracy and operational efficiency of the predictive model. The optimization process using L1 regularization can be expressed through the following formula:

$$\hat{\beta} = arg\min_{\beta} (\sum_{i=1}^{n}(y_i - X_i\beta)^2 + \lambda \sum_{j=1}^{p}\left|\beta_j\right| \quad (2)$$

Among them $\widehat{\beta}$ is the optimal feature coefficient, λ is the regularization strength, y_i is the sample label, and X_i is the feature data. Through L1 regularization, the model can effectively filter out the most relevant features for fraud identification, improve the predictive ability of the model, and avoid overfitting.

4.3. Adopting Interpretable Models

Effective detection of financial fraud risk relies on the application of interpretable models, which help overcome the "black box" problem and enhance transparency in predictive decision-making. The key advantage of interpretable models is their ability to clearly demonstrate the reasoning process, allowing users to understand how predictive conclusions are derived and providing a more reliable basis for risk management and operational decisions.

Optimizing model interpretability requires attention to both model selection and result explanation. In terms of model selection, algorithms such as logistic regression and decision trees are highly interpretable and can intuitively reveal the weights and decision rules of different features. For more complex models, such as deep learning networks, specialized interpretation techniques are necessary to analyze the contributions of individual features and understand model behavior.

For example, a financial institution applied the SHAP (Shapley Additive Explanations) metric to interpret the outputs of complex deep learning models in fraud detection. This method translates the influence of various transaction characteristics into easily understandable numerical representations, revealing key indicators such as abnormal fluctuations in transaction amounts, increased nighttime trading activity, and frequent device changes. By leveraging this quantitative information, the institution was able to refine risk prevention strategies, provide actionable operational guidance on identifying fraudulent behaviors, and support business units in making informed decisions.

The core calculation formula for SHAP values is as follows:

$$\alpha_{ij} = \frac{exp(e_{ij})}{\sum_{k=1}^{n} exp(e_{ik})} \quad (3)$$

Among them, α_{ij} represents the importance weight of feature j to the i-th decision, e_{ij} is the feature correlation score obtained through neural network learning, and n is the total number of features. This system helps to clearly explain the judgment logic of deep learning algorithms, significantly enhancing the explanatory power of high difficulty models, and can determine the contribution of main features to the predicted output, thereby indicating the path for model improvement.

The improvement of interpretable models should follow a standardized procedure, aiming to maintain high model interpretability without compromising predictive performance. The process begins with the preliminary organization of input data to ensure quality and consistency, providing a solid foundation for effective model training.

Next, appropriate interpretability analysis tools should be selected, such as rule-based models or functionally enhanced interpretation techniques, to thoroughly investigate the decision-making logic of the model and clarify the relationships between various features and output results. At this stage, evaluating the importance of features and presenting them visually is critical, as this clearly highlights the primary features on which the model relies and their interconnections.

Simultaneously, the model architecture should be adjusted according to actual business requirements to ensure compatibility between explanatory power and specific application contexts. By continuously verifying and optimizing the interpretation results, the system can provide stable and reliable decision support across diverse operational scenarios.

Figure 1 illustrates the complete optimization workflow, from data preprocessing to model interpretability verification, offering a clear and structured operational pathway for the identification of financial fraud risks.

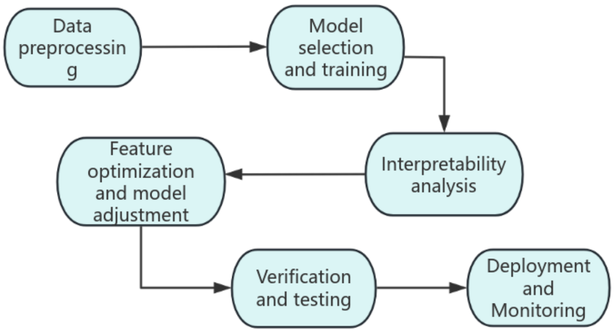
Data preprocessin g
Model selection and training
Interpretability analysis
Feature optimization and model adjustment
Verification and testing
Deployment and Monitoring

Figure 1. Optimization process of financial fraud risk identification system based on interpretable model.

4.4. Cross Validation and Ensemble Learning

Cross validation and ensemble learning are essential techniques for enhancing the performance and robustness of financial fraud risk identification systems. Cross validation divides the dataset into multiple subsets and sequentially uses each subset as training and validation data, allowing comprehensive evaluation of model performance across different data distributions. Ensemble learning, on the other hand, combines the predictive outputs of multiple models to improve overall accuracy and stability. Common ensemble methods include Bagging, exemplified by random forests, and Boosting, represented by gradient boosting decision trees (GBDT). These techniques provide reliable strategies for data processing, model training, and performance optimization.

For instance, an e-commerce platform implemented a hybrid learning strategy that integrates soft voting and stacking in its fraud detection system. Under the soft voting mechanism, the system aggregates prediction probabilities from multiple base models, including logistic regression, support vector machines, and neural networks, and generates a comprehensive prediction through weighted combination. Stacking further enhances performance by using the prediction results of primary models as new input features for a secondary model, such as linear regression, allowing deeper learning and refinement. This approach improves both prediction accuracy and system stability.

Table 1 presents a summary of the optimization steps and specific operations of cross validation and ensemble learning in financial fraud risk identification systems, providing a structured framework for practical implementation.

Table 1. Cross validation and ensemble learning optimization steps.

step	Method or technique	specific operation
Data partitioning	cross validation	Divide the dataset into multiple subsets and evaluate the stability and accuracy of the model through alternating training and validation.

model training	Decision Tree/Random Forest	Train multiple models using decision trees or random forests, perform multiple cross validations, and optimize model performance.
Integration Strategy	Bagging	Train multiple sub models through resampling and ultimately vote or average the results (random forest).
ensemble learning	Boosting	Based on the error adjustment weights of the previous model, gradually optimize the model (gradient boosting tree).
model fusion	weighted fusion	Weight and fuse the results of multiple models to output the final prediction result.
Performance Verification	Cross validation evaluation	Adjust the super parameters in each fold cross validation to ensure the stability of the model, and conduct performance validation.

These ensemble learning strategies enhance the overall model's accuracy and stability by effectively combining the predictions of multiple models. By integrating diverse model outputs, they provide a more robust approach to capturing complex patterns and variations in financial fraud behavior, thereby improving the system's ability to detect and respond to sophisticated fraudulent activities.

5. Conclusion

In the field of financial fraud identification, the application of intelligent algorithms has enabled more efficient, automated, and accurate processing of financial transaction data. By systematically optimizing data quality, implementing advanced feature selection methods, enhancing model interpretability, and adopting strategies such as cross validation and ensemble learning, the predictive accuracy, robustness, and reliability of fraud detection systems have been significantly improved.

These improvements allow financial institutions to better address the increasing diversity, sophistication, and complexity of fraudulent behaviors, effectively constructing a strong risk defense framework. Moreover, the integration of interpretable models provides transparency in decision-making, facilitating compliance with regulatory requirements and strengthening operational confidence.

Looking forward, with the continuous advancement of machine learning techniques, interpretability tools, and the expansion of financial datasets, fraud identification systems are expected to become increasingly intelligent, adaptive, and automated. These developments will further enhance the system's capability to detect emerging fraud patterns in real time, providing a more robust and reliable safeguard for the security and stability of financial markets.

References

1. W. Sun, S. Lei, L. Wang, Z. Liu, and Y. Zhang, "Adaptive federated learning and digital twin for industrial internet of things," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 8, pp. 5605-5614, 2020. doi: 10.1109/tii.2020.3034674
2. A. Iranmehr, H. Masnadi-Shirazi, and N. Vasconcelos, "Cost-sensitive support vector machines," *Neurocomputing*, vol. 343, pp. 50-64, 2019. doi: 10.1016/j.neucom.2018.11.099
3. N. Saeed, M. Ashour, and M. Mashaly, "Comprehensive review of federated learning challenges: a data preparation viewpoint," *Journal of Big Data*, vol. 12, no. 1, p. 153, 2025. doi: 10.1186/s40537-025-01195-6
4. M. Ivanova, P. Petkova, and N. Petkov, "Machine learning and fuzzy logic in electronics: Applying intelligence in practice," *Electronics*, vol. 10, no. 22, p. 2878, 2021. doi: 10.3390/electronics10222878
5. K. Gupta, P. Kumar, S. Upadhyaya, M. Poriye, and S. Aggarwal, "Fuzzy logic and machine learning integration: Enhancing healthcare decision-making," *International Journal of Computer Information Systems and Industrial Management Applications*, vol. 16, no. 3, pp. 20-20, 2024.
6. A. Burduk, D. Lapczynska, J. Kochanska, K. Musial, and J. Husar, "Fuzzy logic in risk assessment of production machines failure in forming and assembly processes," *Journal of Machine Engineering*, vol. 24, no. 2, pp. 34-43, 2024.

7. C. Y. Lee, T. A. Le, and C. L. Hung, "A feature selection approach based on memory space computation genetic algorithm applied in bearing fault diagnosis model," *IEEE Access*, vol. 11, pp. 51282-51295, 2023.
8. P. Boobalan, S. P. Ramu, Q. V. Pham, K. Dev, S. Pandya, P. K. R. Maddikunta, and T. Huynh-The, "Fusion of federated learning and industrial Internet of Things: A survey," *Computer Networks*, vol. 212, p. 109048, 2022.
9. U. Sinha, J. D. P. Rao, S. K. Swarnkar, and P. K. Tamrakar, "Advancing Early Cancer Detection with Machine Learning: A Comprehensive Review of Methods and Applications," *Multimedia Data Processing and Computing*, pp. 165-174, 2023.
10. M. Shahamat, "Applications of Fuzzy Logic in Modern Technology," *Journal of intelligent decision and computational modelling*, vol. 1, no. 1, pp. 27-34, 2025.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of GBP and/or the editor(s). GBP and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.