

Financial Fraud Detection Based on an XGBoost and LSTM Fusion Model: A Comparative Study on the Enhancement of Time-Series Features

Yumeng Tian

*Business School, Soochow University, Suzhou, China
ymtiantym@stu.suda.edu.cn*

Abstract: Financial fraud detection is critical for modern financial systems, particularly as fraudulent schemes become increasingly complex and insidious, causing significant economic damage to institutions and consumers. Traditional rule-based methods have proven inadequate in countering these sophisticated behaviors. Despite considerable progress in fraud detection research, notable gaps persist. Existing studies tend to emphasize structured features while overlooking the integration of historical transaction data and multi-source information. By methodically contrasting the performance of several models, including Logistic Regression, XGBoost, and Support Vector Machines, with that of deep learning models, like Long Short-Term Memory networks, under circumstances with and without time-series data, this study investigates the role of time-series features in fraud detection. Furthermore, we propose an innovative XGBoost+LSTM fusion model that synergistically combines structured data with temporal information to improve detection accuracy and reduce false alarms. Our results indicate that time-series features markedly enhance deep learning model performance, whereas their inclusion in traditional models may introduce noise, leading to performance fluctuations.

Keywords: Machine Learning, Deep Learning, Fusion Model, Financial Fraud, Feature Selection

1. Introduction

The expansion of online transactions, mobile payments, and electronic banking has sped up the digital transformation of financial services and brought up new difficulties, especially in the area of financial fraud.

Traditional rule-based fraud detection systems, relying on static logic, struggle to adapt to evolving fraudulent tactics. Moreover, fraud detection inherently poses a binary classification problem where fraudulent transactions constitute only a small part of all the transactions, resulting in significant class imbalance during model training.

As a result, effectively leveraging the key information available within the data to identify potential anomalies has emerged as a central challenge in financial risk management.

Extensive international research has explored integrating multi-modal data to enhance fraud detection system robustness. Early studies introduced authentication simulators that combined diverse data sources, markedly improving online fraud detection adaptability [1]. Similarly, research

on aviation ticket fraud detection demonstrated that deep neural networks can effectively model historical transaction data alongside supplementary modalities, significantly enhancing detection performance [2]. With advances in time-series analysis, numerous investigations have adopted recurrent neural networks—particularly LSTM—to capture temporal dependencies within transactional data, thereby revealing latent fraudulent patterns [3]. Domestically, scholars have addressed data imbalance and enhanced model generalization; one study integrated advanced sampling techniques with XGBoost to bolster credit card fraud detection accuracy [4].

Despite these advances, existing literature predominantly focuses on structured features or isolated model performance, with scant research integrating XGBoost and LSTM into a unified fusion model. Our study addresses this gap by proposing a novel XGBoost+LSTM model that fuses structured and temporal features, providing both theoretical insights and practicality.

This paper assesses the contribution of time-series features to financial fraud detection and develops a fusion model that integrates XGBoost and LSTM within a multi-modal framework to enhance accuracy and robustness.

Using and without time-series data, we first compare deep learning models like LSTM with more conventional machine learning models like XGBoost, SVM, and Logistic Regression. This evaluation elucidates the impact of historical transaction information on detection efficacy.

Subsequently, based on our experimental findings, we propose an innovative XGBoost+LSTM fusion model that leverages XGBoost's strength in handling structured data and LSTM's capability in capturing temporal dependencies. The resulting system significantly improves detection accuracy by reducing false negatives and false positives.

This research not only introduces a novel technical approach to financial fraud detection but lays a solid empirical and theoretical foundation for future studies in real-time detection, multi-source data fusion, and interpretable model design.

2. Data selection and research methods

2.1. Dataset description

This study employs the “fraud-detection-transactions-dataset” from Kaggle, comprising approximately 50,000 transaction records. The dataset encapsulates multiple dimensions—including transaction amount, type, timestamp, merchant category, risk score, and geographical location—with each record labeled as fraudulent (1) or non-fraudulent (0).

Initial analysis reveals a pronounced class imbalance shown in Figure 1, with non-fraudulent transactions far exceeding fraudulent ones. Moreover, Figure 2 indicates subtle discrepancies in fraudulent activity distributions between weekdays and non-working days, justifying subsequent time-series experiments. A numerical correlation heat map, as Figure 3, shows that key variables are largely uncorrelated, suggesting the advantage of non-linear models in capturing complex relationships.

These insights underpin our advanced model selection and further guide our methodology.

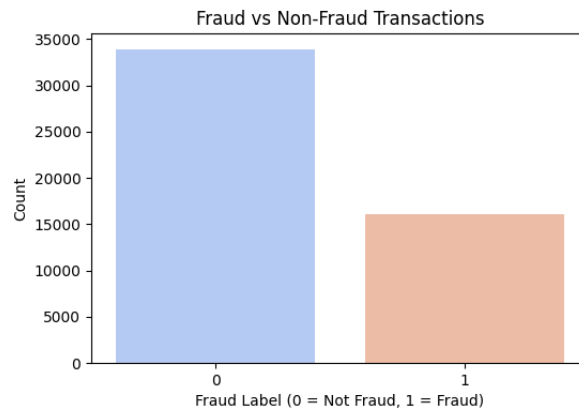


Figure 1: Fraudulent and non-fraudulent transactions distribution

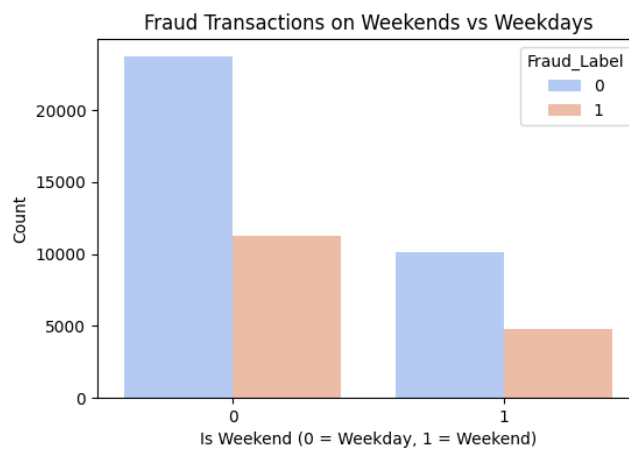


Figure 2: Transactions on workdays and non-workdays distribution

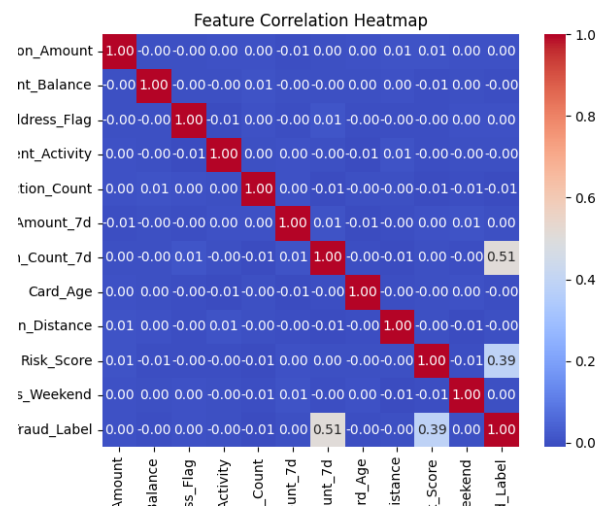


Figure 3: Numerical feature variables heatmap distribution

2.2. Data processing

Before model training, the dataset was systematically cleaned, visualized, and enriched through feature engineering. Minimal missing values were either removed or imputed to preserve data integrity. Categorical variables—such as transaction type, merchant category, and city—were converted to numerical formats using label encoding or one-hot encoding. Continuous variables, including account balance, transaction amount, and risk score, were standardized to mitigate scale discrepancies.

Temporal dynamics were captured by extracting features (hour, day, weekday) from the “Timestamp” and computing historical statistics, like the 7-day rolling average of transaction amounts. Based on these processed features, two experimental strategies were developed: an enhancement strategy retaining full-time-related features, and a diminution strategy simplifying them to isolate the impact of temporal information.

2.3. Model selection and construction

This study investigates the role of time-series features in financial fraud detection by comparing model performance under varying feature configurations. In addition to deep learning architectures like LSTM, the research assesses conventional machine learning models like Logistic Regression, SVM, and XGBoost [5-8]. Two experimental cohorts were established: one employing solely structured features, and the other augmenting the dataset with time-series attributes, including transaction timestamps, distinctions between weekdays and weekends, and the 7-day rolling average [9-13].

Conventional models such as Logistic Regression and SVM primarily rely on linear discriminative characteristics, whereas XGBoost and LSTM are designed to capture non-linear relationships. Specifically, XGBoost implements a gradient-boosting framework to iteratively construct decision trees, excelling in high-dimensional, sparse data contexts. In contrast, LSTM utilizes gating mechanisms—including forget and input gates—to effectively extract temporal dynamics from sequential data.

The study’s innovation lies in a novel fusion model that integrates XGBoost with LSTM: first, XGBoost generates structured data probabilities, which are then concatenated with temporal features for further processing by an LSTM network.

2.4. Detailed construction of fusion model

2.4.1. XGBoost+LSTM fusion computation

During the fusion phase, the following steps are implemented:

- (1) The output of the LSTM, $P_{LSTM}(x)$, represents the prediction based on time-series features.
- (2) The output of the XGBoost, $P_{XGB}(x)$, represents the prediction based on structured features.
- (3) These two outputs are concatenated:

$$h_{fusion} = [P_{LSTM}(x), P_{XGB}(x)] \quad (1)$$

- (4) The concatenated features are passed through a fully connected layer to perform the ultimate classification:

$$\hat{y} = \sigma(W_{fusion} \cdot h_{fusion} + b_{fusion}) \quad (2)$$

Where W_{fusion} is the weight matrix, b_{fusion} is the bias term and $\sigma(\cdot)$ is the Sigmoid function ensuring the final probability is within $[0, 1]$.

2.4.2. Training procedure and loss function

Minimizing the binary cross-entropy loss is the training goal:

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^N [y_i \log \hat{y}_i + (1 - y_i) \log(1 - \hat{y}_i)] \quad (3)$$

Where y_i are true labels and $\hat{y}_i$ are the model predictions.

2.4.3. Training process

Gradient descent is used to update the parameters:

$$W \leftarrow W - \eta \frac{\partial \text{Loss}}{\partial W} \quad (4)$$

Where η the learning rate, which is set to 0.001, while $\frac{\partial \text{Loss}}{\partial W}$ are the gradients and are computed via the backpropagation algorithm.

2.4.4. Prediction phase

Given the test data, the prediction procedure is as follows:

(1) XGBoost computes the probability for the structured features.

$$P_{\text{XGB}}(x) = \sigma(f(x)) \quad (5)$$

(2) LSTM computes the probability for the time-series features.

$$P_{\text{LSTM}}(x) = \sigma(W_{\text{dense}} \cdot h_T + b_{\text{dense}}) \quad (6)$$

(3) The outputs from the two models are combined to yield the final prediction.

$$h_{\text{fusion}} = [P_{\text{LSTM}}(x), P_{\text{XGB}}(x)] \quad (7)$$

$$\hat{y} = \sigma(W_{\text{fusion}} \cdot h_{\text{fusion}} + b_{\text{fusion}}) \quad (8)$$

(4) A binary classification decision is made by applying a threshold of 0.5.

$$y_{\text{pred}} = \begin{cases} 1, & \hat{y} > 0.5 \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

This approach leverages the strengths of both structured data and temporal data to achieve further improvements in accuracy and stability. The described methodology encompasses strategies including traditional linear methods, ensemble learning, and deep neural networks, thereby enabling a comprehensive comparison of performance under various time-series feature configurations.

2.5. Evaluation metrics

The experimental evaluation employs several key metrics, including Accuracy, Precision, Recall, and F1-score. Additionally, confusion matrices are utilized to visually assess the performance by delineating.

Confusion matrices include: True Positives (TP), True Negatives (TN), False Negatives (FN), and False Positives (FP).

The specific formulas are as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (11)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (12)$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

2.6. Experimental protocol

To guarantee objective assessment, the dataset is divided into training and testing subsets. Two experimental groups are established: one with only structured features and one with enhanced time-series features. Models are trained separately under these conditions, and their performance is compared horizontally (across different models for the same feature set) and vertically (the same model's performance across different feature settings).

3. Experimental results

This section presents a detailed comparison of model performances under two experimental conditions: the “enhanced time-series group” and the “weakened time-series group.” For the enhanced group, the dataset includes additional temporal features such as transaction hours, weekday/weekend indicators, and a 7-day rolling average of transaction amounts. Evaluation criteria contain Accuracy, Recall, F1-score, and Precision, with further insights drawn from the confusion matrices.

3.1. Performance comparison metrics

3.1.1. Enhanced time-series group

The results of the enhanced time-series group are as follows, as shown in Table 1 and Table 2.

Table 1: Fraudulent transactions in enhanced time-series group performance

Model	Accuracy	F1-score	Recall	Precision
Logistic Regression	0.8019	0.68	0.64	0.72
SVM	0.9780	0.97	0.97	0.96
XGBoost	0.9989	0.998	0.998	0.998
LSTM	0.9949	0.997	0.993	0.991
XGBoost+LSTM	0.999	0.998	0.999	0.998

Table 2: Non-fraudulent transactions in enhanced time-series group performance

Model	Accuracy	F1-score	Recall	Precision
Logistic Regression	0.8019	0.86	0.88	0.84
SVM	0.9825	0.98	0.98	0.99
XGBoost	0.9989	0.999	0.999	0.999
LSTM	0.9949	0.996	0.996	0.996
XGBoost+LSTM	0.999	0.9993	0.999	0.9996

3.1.2. Weakened Time-series group

The results of the weakened time-series group are as follows, as shown in Table 3 and Table 4.

Table 3: Fraudulent transactions in weakened time-series group performance

Model	Accuracy	F1-score	Recall	Precision
Logistic Regression	0.8019	0.68	0.64	0.72
SVM	0.9825	0.97	0.98	0.97
XGBoost	0.9989	0.998	0.998	0.998
LSTM	0.9947	0.991	0.997	0.986

Table 4: Non-fraudulent transactions in weakened time-series group performance

Model	Accuracy	F1-score	Recall	Precision
Logistic Regression	0.8019	0.86	0.88	0.84
SVM	0.9825	0.99	0.99	0.99
XGBoost	0.9989	0.99	0.99	0.99
LSTM	0.9947	0.995	0.993	0.998

3.2. Discussion of results

Logistic Regression yields nearly identical outcomes across both experimental groups, indicating that its performance remains largely invariant to the inclusion or exclusion of time-series features. Its persistently poor recall of fraudulent situations is caused by its incapacity to grasp the intricate temporal dynamics included in transaction data. In contrast, SVM achieves high detection accuracy with solely structured features; however, incorporating time-series data introduces minor fluctuations, likely due to noise complicating kernel-space discrimination.

XGBoost demonstrates near-perfect performance without explicit time-series variables, with only marginal improvements upon their addition, suggesting that its ensemble decision tree framework effectively leverages structured data while deriving limited incremental benefit from temporal features. Conversely, LSTM significantly benefits from time-series information, exhibiting notable improvements in precision and F1-score by adeptly capturing sequential dependencies, thereby reducing false negatives and enhancing overall anomaly detection.

The proposed XGBoost+LSTM fusion model achieves superior performance by first employing XGBoost to predict probabilities from structured data, and then concatenating these scores with temporal features for subsequent processing by LSTM. This hybrid technique shows near-perfect precision and recall, with an estimated accuracy of 0.999. Confusion matrix analysis reveals little misclassification of fraudulent and non-fraudulent transactions, demonstrating the model's efficacy in reducing false positives and negatives.

To sum up, while traditional models experience only modest benefits or even slight deterioration due to noise from time-series integration, deep learning models—particularly LSTM—derive substantial gains. The fusion approach, which synthesizes the strengths of both paradigms, presents a promising solution for financial fraud detection in environments characterized by severe data imbalance and dynamic transaction patterns.

4. Discussion and analysis

This section provides an integrated discussion of our experimental findings and highlights the primary innovations and implications of this study.

4.1. Results discussion

The experimental study was designed to evaluate the impact of time-series features on financial fraud detection by comparing models across two distinct data groupings: one lacking explicit time-series features and one with an enhanced time-series feature set. Accuracy, F1-Score, recall, and precision are among the evaluation metrics that, when paired with visual evaluations through confusion matrices, provide various important insights into the model's performance.

Firstly, traditional linear models such as Logistic Regression fail to harness the additional time-series information effectively. With an accuracy of around 0.80 and a recall for fraudulent transactions of only 0.64, Logistic Regression is unable to capture the nonlinear temporal dependencies in the data. This limitation underscores its vulnerability in complex fraud detection tasks where subtle sequential patterns can be critical.

Secondly, SVM demonstrates robust performance based on the inherent nonlinear separability present in structured features. However, when time-series features are added, the performance metrics fluctuate slightly. This suggests that while SVM is proficient with static data, the incorporation of additional temporal features may introduce noise into the kernel mapping process, thereby affecting the stability of the decision boundary.

In contrast, XGBoost exhibits near-perfect performance when relying on structured features (accuracy around 0.9989), with only marginal improvements when temporal features are included. This observation implies that although XGBoost's ensemble method is highly effective for static data classification, its sensitivity to dynamic time-series information is relatively limited. Nonetheless, even slight enhancements in boundary discrimination, particularly for the non-fraudulent class, indicate some benefit from integrating additional features.

The deep learning model, LSTM, capitalizes significantly on the inclusion of time-series features. Its inherent capacity to model temporal dependencies through gating mechanisms leads to notable improvements. In particular, the F1-Score for fraudulent transactions increases from 0.991 to 0.997 when temporal information is enhanced, thus reducing the likelihood of false negatives. This result validates that LSTM's architecture is well-suited to extract meaningful patterns from sequential data, thereby improving the overall detection capability.

Most notably, the fusion model—XGBoost+LSTM—outperforms all other models. By combining the strengths of XGBoost in processing static structured features with the superior temporal pattern recognition abilities of LSTM, the fusion model achieves an accuracy of 0.999 with near-perfect precision and recall for both fraudulent and non-fraudulent cases. Confusion matrix analyses further reinforce that this model markedly reduces false positives and false negatives, misclassifying only a minimal number of transactions. This result demonstrates that the multi-modal fusion strategy is highly effective in addressing the challenges of data imbalance and dynamic fraud patterns in financial transactions.

In summary, the results indicate that while traditional models may efficiently utilize structured data, they are less capable of extracting the critical temporal patterns embedded within the data. Deep learning models, particularly LSTM, show substantial improvements when time-series features are explicitly integrated. More importantly, our proposed fusion model significantly enhances detection performance by leveraging the complementary strengths of both conventional and deep learning approaches, thus emerging as a robust method for financial fraud detection.

4.2. Research implications

The results of this study have a number of significant ramifications for both scholarly research and real-world financial fraud detection applications.

4.2.1. Advancement of time-series feature utilization

The study demonstrates that the incorporation of time-series features—such as the 7-day moving average and transaction frequency—can significantly enhance deep learning models' ability to identify fraudulent transactions. For practitioners, this means that thorough temporal feature engineering should be considered a critical step in the development of fraud detection systems. The observed improvement in recall and reduction in false negatives suggest that such models can provide timely and more accurate fraud warnings.

4.2.2. Model selection and multi-modal fusion

According to the comparative study, conventional models such as SVM and logistic regression, while proficient in static data processing, might not fully exploit the potential of time-dependent patterns. In contrast, LSTM's performance underscores the importance of leveraging models that inherently accommodate sequential data. However, the fusion model—which integrates XGBoost and LSTM—illustrates how combining the strengths of different methodologies can yield superior results. This multi-modal approach opens new avenues for research and development, as it encourages the blending of diverse data sources and modeling techniques to build holistic fraud detection systems that overcome the limitations of individual methods.

4.2.3. Addressing data imbalance and robustness

Financial datasets are characteristically imbalanced, with fraud cases comprising a small fraction of the total transactions. Our empirical results indicate that the fusion model is particularly robust in handling such data imbalances. By effectively reducing false negatives and false positives, the model offers enhanced stability, which is crucial for practical applications in risk management. For future research, this suggests a promising direction: integrating advanced sampling techniques or cost-sensitive learning within fusion frameworks to further augment performance in imbalanced settings.

4.2.4. Scalability and practical deployment

From an implementation perspective, the proposed fusion model has significant practical implications. Although deep learning models with multi-modal fusion are generally computationally intensive, the trade-off is justified in high-stakes applications where the cost of fraud is substantial. Furthermore, as financial datasets continue to grow in volume and complexity, scalable models capable of processing multi-source data become increasingly critical. The encouraging performance of our fusion model suggests that it can serve as a foundation for developing real-time fraud detection systems that are both accurate and reliable in dynamic financial environments.

4.2.5. Future research directions

Our study paves the way for several promising research directions. Future efforts could focus on expanding the range and granularity of time-series features, such as incorporating high-frequency trading intervals, cyclic behavior patterns, and anomaly trend analysis to capture even finer temporal nuances. Additionally, real-time fraud detection remains an open challenge. Incorporating online learning algorithms and real-time streaming analytics can enable models to adapt continuously to new fraud patterns as they emerge. Another critical area for development is model interpretability and explainability. Given the regulatory importance of financial fraud detection, future studies should explore methods such as attention mechanisms and feature importance

visualization to render deep learning and fusion models more transparent. Finally, the integration of additional data modalities—such as geolocation, social network information, and device fingerprints—can further enrich the detection framework, making it applicable across a wider range of financial scenarios, including insurance fraud and e-commerce refund fraud.

In essence, our research provides a new, comprehensive strategy for financial fraud detection by demonstrating that the integration of time-series features via a fusion model significantly enhances detection performance. The implications of this work extend from theoretical insights into temporal dynamics in fraud behavior to practical methodologies that can be deployed in operational financial systems. The multi-modal approach not only improves detection accuracy but also offers greater resilience against the inherent challenges of imbalanced and dynamic data, thereby contributing to safer and more efficient financial operations. These contributions should spur further research in model fusion, real-time analytics, and cross-modal data integration, eventually developing cutting-edge technology for financial security.

5. Conclusion

By methodically contrasting deep learning models (LSTM) and conventional machine learning models (Logistic Regression, XGBoost, SVM) under two data settings—one with only structured features and another enhanced with temporal attributes—this study explores the function of time-series features in improving financial fraud detection. Through comprehensive experiments, This paper assessed how sequential transaction information improves the identification of anomalous patterns in highly imbalanced datasets. Furthermore, we proposed a fusion model—XGBoost+LSTM—that integrates structured data processing with temporal sequence modeling. This hybrid architecture markedly improves detection accuracy while significantly reducing both the misjudgment rate and the rate of missing reports, which makes it particularly effective in complex and dynamic financial environments.

This paper showed that while conventional models like SVM and Logistic Regression function well when using only structured variables, they either show little improvement or suffer when time-series features are included. In contrast, the LSTM model leverages its inherent sequential processing capability to achieve a substantial boost in performance metrics—most notably in recall and overall F1-score—when temporal features are integrated. The fusion model, XGBoost+LSTM, emerged as the best-performing approach, achieving near-perfect classification with an overall accuracy of approximately 0.999. This model's ability to merge the robust, high-dimensional feature extraction of XGBoost with the dynamic modeling capabilities of LSTM underlines its potential to serve as a reliable framework for financial fraud detection.

Building on these results, the following directions are recommended for future research:

(1) Expansion of Time-Series Features

Future studies should investigate the incorporation of a broader array of time-series features, such as high-frequency intervals, cyclic patterns, and specific anomaly trends, to capture even more nuanced transaction dynamics.

(2) Real-Time Detection and Online Learning

Since most current studies, including this one, are based on offline datasets, research should focus on developing real-time fraud detection systems that utilize streaming data and online learning algorithms to continually update model parameters in response to emerging fraud patterns.

(3) Model Scalability and Interpretability

Enhancing fusion models' efficacy and interpretability is essential for practical use. Techniques such as model pruning, knowledge distillation, and the application of attention mechanisms could be pivotal in developing lightweight models with strong explanatory power.

(4) Integration of Multi-Modal Data

Future studies should investigate the merging of other data sources, such as device fingerprints, social network data, and geolocation information, in addition to structured transaction and temporal data. Such multi-modal fusion holds promise for constructing even more comprehensive and robust fraud detection systems applicable to diverse domains like insurance, e-commerce, and mobile payments.

(5) Addressing Class Imbalance

Given the prevalence of class imbalance in financial fraud datasets, further exploration of advanced sampling strategies and cost-sensitive learning methods in conjunction with fusion models could yield improvements in the detection of rare fraudulent events.

Generally speaking, this research provides a novel, integrated approach to financial fraud detection by harnessing time-series features alongside traditional structured data. The proposed XGBoost+LSTM fusion model demonstrates exceptional performance in both accuracy and robustness, significantly outperforming conventional methods, particularly in scenarios characterized by severe class imbalance. These findings not only offer valuable theoretical insights into the role of temporal dynamics in fraud behavior but also provide practical guidelines for building advanced, scalable fraud detection systems. The methodologies and implications of this research represent a meaningful contribution to the ongoing efforts to enhance financial security, and they lay a solid foundation for future studies to further explore and extend multi-modal data integration and real-time analysis in complex operational environments.

References

- [1] Zintgraf, L., Lopez-Rojas, E. A., Diederik Roijers, & Nowe, A. (2017). *MultiMAuS: A Multi-Modal Authentication Simulator for Fraud Detection Research*. *European Modeling and Simulation Symp, (EMSS 2017)(29th)*, 360–369.
- [2] Aras, M. T., & Guvensan, M. A. (2023). *A Multi-Modal Profiling Fraud-Detection System for Capturing Suspicious Airline Ticket Activities*. *Applied Sciences*, 13(24), 13121. <https://doi.org/10.3390/app132413121>
- [3] Lin, P. (2022). *Application of Embedding Method-Convolutional Neural Network in Credit Card Fraud Detection*. *Advances in Applied Mathematics*, 11(02), 717–725.
- [4] Ma, L. (2024, April). *Credit Card Fraud Detection Method Based on KRSMOTE+ ENN and XGBoost Algorithm*. In *Proceedings of the 5th International Conference on Computer Information and Big Data Applications* (pp. 1225-1231).
- [5] LaValley, M. P. (2008). *Logistic regression*. *Circulation*, 117(18), 2395-2399.
- [6] Chandra, M. A., & Bedi, S. S. (2021). *Survey on SVM and their application in image classification*. *International Journal of Information Technology*, 13(5), 1-11.
- [7] Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., ... & Zhou, T. (2015). *Xgboost: extreme gradient boosting*. *R package version 0.4-2*, 1(4), 1-4.
- [8] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). *A review of recurrent neural networks: LSTM cells and network architectures*. *Neural Computation*, 31(7), 1235-1270.
- [9] Gu, Y., Wei, J., & Cheong, N. (2024, September). *Credit Card Fraud Detection Based on MiniKM-SVMSMO TE-XGBoost Model*. In *Proceedings of the 2024 8th International Conference on Big Data and Internet of Things* (pp. 252-258).
- [10] Wu, Y., Wang, L., Li, H., & Liu, J. (2025). *A Deep Learning Method of Credit Card Fraud Detection Based on Continuous-Coupled Neural Networks*. *Mathematics*, 13(5), 819.
- [11] Nguyen, T. T., Tahir, H., Abdelrazek, M., & Babar, A. (2020). *Deep learning methods for credit card fraud detection*. *arXiv preprint arXiv:2012.03754*.
- [12] Yang, J., Chen, K., Ding, K., Na, C., & Wang, M. (2023). *Auto insurance fraud detection with multimodal learning*. *Data Intelligence*, 5(2), 388-412.
- [13] Lopez-Rojas, E. A., Axelsson, S., & Baca, D. (2018). *Analysis of fraud controls using the PaySim financial simulator*. *International Journal of Simulation and Process Modelling*, 13(4), 377-386.