Article

Not peer-reviewed version

Towards a Feed-Forward Neural Network for Financial Fraud Detection

Oyindamola Ogunruku *

Posted Date: 6 May 2025

doi: 10.20944/preprints202505.0286.v1

Keywords: fraud detection; deep neural network; AFRAMES; classification

 doi:10.20944/preprints202505.0286.v1

Article

Towards a Feed-Forward Neural Network for Financial Fraud Detection

Oyindamola Omolara Ogunruku

Western Illinois University, 61455, IL, USA; oo-ogunruku@wiu.edu

Abstract: This paper presents the development of a feed-forward neural network model for the early detection and mitigation of fraudulent transactions in banking systems. Using synthetically generated data that mimics real-world class imbalance between legitimate and fraudulent activities, we build a binary classifier capable of distinguishing between the two with high accuracy. The model achieved an overall accuracy of 98%, with a recall of 84% for fraudulent transactions. Results indicate that simple yet well-structured deep learning architectures can effectively reduce false negatives—a critical factor in fraud detection. Visualization techniques including Principal Component Analysis (PCA) projection and confusion matrices were used to evaluate classification performance and feature space separation. The approach provides a scalable foundation for integrating AI into financial security pipelines and can serve as a core component of the Advanced Financial Risk Analytics and Management System (AFRAMES).

Keywords: fraud detection deep neural network AFRAMES classification

1. Introduction

Financial institutions today face a dual imperative: to facilitate seamless digital transactions while simultaneously safeguarding against increasingly sophisticated fraud tactics (Khang et al. 2025; Daah et al. 2024). As digital banking and online payment platforms proliferate, so too do the risks associated with unauthorized access, identity theft, and deceptive transactions. Traditional fraud detection systems—primarily based on static rule sets and deterministic thresholds—are no longer sufficient in an environment where threat actors continuously adapt and evolve (Dakalbab et al. 2024; Biju et al. 2024; Mohsen et al. 2025). These systems, though fast and interpretable, often suffer from high false positive rates, low adaptability, and delayed response time, thereby increasing financial loss and reducing customer satisfaction.

To address these limitations, the research community has been increasingly turning toward artificial intelligence (AI) and machine learning (ML) for fraud mitigation. These approaches offer dynamic, data-driven insights capable of adapting to novel fraud patterns in real-time. Among them, neural networks have demonstrated a distinct advantage due to their ability to model complex, nonlinear relationships in large datasets (LeCun et al, 2015; Goodfellow et al, 2016). Particularly, feed-forward neural networks (FNNs) have shown promise as robust classifiers in domains with significant noise, high dimensionality, and class imbalance a characteristic of most real-world fraud datasets (Sahin et al, 2011). This study proposes a feed-forward neural network model to detect fraudulent banking transactions. Unlike many existing approaches that rely heavily on data resampling techniques or sophisticated ensemble methods, our work investigates the performance of a relatively simple neural network model under natural class imbalance conditions. Our results indicate that with appropriate regularization and architectural tuning, a standard FNN can achieve high precision and recall for fraud detection, even without overcomplicating the pipeline.

The novelty of this work lies not only in the architecture or the synthetic data generation but also in its contribution to a broader integrated framework known as the **Advanced Financial Risk Analytics and Management System (AFRAMES)**. AFRAMES is a developing suite of intelligent

 doi:10.20944/preprints202505.0286.v1

tools aimed at improving the detection, analysis, and mitigation of financial risks in banking and fintech sectors. By incorporating this feed-forward neural network model into AFRAMES, we strengthen its fraud detection capabilities and demonstrate the viability of modular, AI-enhanced risk management systems. This integration lays the groundwork for future deployments involving hybrid models, streaming data analytics, and real-time alert systems, all essential components of next-generation financial intelligence platforms.

The remaining sections of this paper are organized as follows. Section 2 reviews related works and foundational literature on fraud detection using machine learning. Section 3 describes the methodology used, including data generation, preprocessing, and neural network design. Section 4 presents and analyzes the experimental results. Finally, Section 5 concludes with key findings and outlines potential areas for future research and AFRAMES integration.

2. Literature Review

Financial fraud detection has long been an active area of research in data mining, machine learning, and financial technology (Zhang and Zhang 2019). As early as the 1990s, researchers began applying statistical models such as logistic regression and linear discriminant analysis to identify fraudulent patterns in financial records. Bolton and Hand (2002) offered one of the seminal reviews of statistical fraud detection techniques, highlighting their effectiveness in structured datasets but also emphasizing their limitations in handling dynamic, nonlinear fraud patterns. The static nature of rule-based and traditional statistical models made them vulnerable to novel and evolving forms of fraud, thus necessitating more adaptive systems.

The introduction of machine learning brought significant improvements to fraud detection pipelines. Techniques such as decision trees, support vector machines (SVM), and ensemble learning methods including Random Forest and Gradient Boosted Trees have demonstrated strong performance in binary classification tasks, especially in environments with complex interactions among variables. Bhattacharyya et al. (2011) evaluated several of these methods on credit card fraud datasets and showed that ensemble-based models often outperformed standalone classifiers in both accuracy and robustness.

One of the persistent challenges in fraud detection is the severe class imbalance. Fraudulent transactions typically constitute less than 1% of total transactions, making it difficult for standard classifiers to learn meaningful patterns without biasing toward the majority class. Researchers have proposed various solutions (Chawla et al., 2002, Jurgovsky et al., 2018), including:

- **Resampling techniques**, such as Synthetic Minority Over-sampling Technique (SMOTE), which synthetically generates minority class examples.
- **Cost-sensitive learning**, which penalizes misclassification of minority instances more heavily.
- **Anomaly detection frameworks**, which treat fraud as a deviation from normal transaction behavior rather than a supervised learning problem.

With the emergence of deep learning, researchers began applying neural network architectures to model fraud detection tasks. Recurrent neural networks (RNNs) and Long Short-Term Memory networks (LSTMs) have been used to capture sequential transaction behavior, particularly in streaming data contexts. However, these models often require large training datasets and substantial computational resources, which may not be feasible for all organizations.

Feed-forward neural networks (FNNs), while simpler in design, have also proven effective when properly regularized and optimized. They are particularly useful for tabular data—a common format in banking and financial systems. Fiore et al. (2019) explored the use of generative adversarial networks (GANs) in augmenting minority fraud cases for training classifiers, but noted that even baseline neural architectures can achieve high accuracy when trained with proper feature scaling, dropout layers, and balanced loss functions.

Despite the growing body of literature on machine learning for fraud detection, practical deployment remains limited by issues such as data availability, model interpretability, and integration into existing financial systems. This paper addresses some of these gaps by proposing a

 doi:10.20944/preprints202505.0286.v1

lightweight, interpretable FNN model that can be seamlessly integrated into banking platforms and enhanced systems such as the **Advanced Financial Risk Analytics and Management System (AFRAMES)**.

The term AFRAMES is coined in this work to describe a modular, AI-driven framework designed for enhanced detection, analysis, and mitigation of financial risks in modern banking and fintech environments. Existing components of AFRAMES include modules for credit risk scoring, market anomaly detection, and early warning systems for operational risk. Our proposed model adds to this growing ecosystem by contributing a reliable, scalable component for fraud detection. It demonstrates that even in the absence of complex ensemble methods or temporal modeling, a well-designed FNN can yield high-performance results in fraud classification tasks.

Moreover, this work promotes the idea that modular and interpretable AI models are not only academically valuable but also operationally necessary for real-world banking applications. The inclusion of this model in AFRAMES supports the system's goals of scalability, modularity, and adaptability to changing financial landscapes.

In conclusion, while the field of fraud detection has seen substantial advances through statistical, machine learning, and deep learning models, this study contributes a novel application of FNNs tailored for imbalanced classification in banking. It bridges theoretical research with practical deployment and aligns with ongoing efforts to develop AFRAMES as a holistic, intelligent platform for financial risk management.

3. Methodology

The primary objective of this study is to develop a feed-forward neural network (FNN) capable of accurately detecting fraudulent banking transactions. This model is designed not only to address the technical challenges posed by class imbalance and feature noise but also to function as a deployable module within the **Advanced Financial Risk Analytics and Management System (AFRAMES)**. The methodology is structured into five main stages: synthetic data generation, preprocessing, model design, training, and evaluation.

3.1. Data Generation and Simulation

Due to the sensitive nature of real-world financial transaction data, this study employs synthetically generated data to simulate a realistic fraud detection scenario. Using the make_classification function from Scikit-learn, a dataset of 5,000 samples was created with the following specifications:

- **Features**: 20 independent numerical features, simulating various transactional and account-based metrics.
- **Informative Features**: 15 features contributed directly to class separation.
- **Redundant Features**: 5 were linearly dependent to mimic correlated financial attributes.
- **Class Distribution**: A highly imbalanced ratio of 95:5 was used to represent legitimate versus fraudulent transactions, respectively.

This approach ensures the presence of signal complexity and class skewness typical of actual financial data while maintaining experimental control.

3.2. Data Preprocessing

To prepare the dataset for training:

- **Train-Test Split**: The dataset was divided into training (64%), validation (16%), and test (20%) subsets.
- **Standardization**: Features were normalized using standard scaler to have zero mean and unit variance, which accelerates convergence in gradient-based models.
- **Label Encoding**: Binary class labels were encoded as 0 (legitimate) and 1 (fraudulent).

 doi:10.20944/preprints202505.0286.v1

These preprocessing steps are essential to ensuring that the model learns effectively without biasing toward dominant features or classes.

3.3. Neural Network Architecture

The architecture of the proposed FNN was kept deliberately minimalistic to enhance interpretability and facilitate integration into AFRAMES. The network consists of the following layers:

- **Input Layer**: Accepts 20 features.
- **Hidden Layer 1**: 32 neurons with ReLU activation, followed by a 30% dropout layer to prevent overfitting.
- **Hidden Layer 2**: 16 neurons with ReLU activation, to further capture non-linear feature interactions.
- **Output Layer**: A single neuron with sigmoid activation for binary classification.
  The model was compiled with:
- **Loss Function**: Binary cross-entropy, suitable for probabilistic binary output.
- **Optimizer**: Adam optimizer, chosen for its robustness and adaptive learning rates.
- **Metrics**: Accuracy and AUC (Area Under the Curve) were tracked during training.

3.4. Model Training

The model was trained for 20 epochs with a batch size of 32. The validation set was used to monitor performance and avoid overfitting. Training was performed using TensorFlow/Keras libraries, and early stopping was employed based on validation loss trends.

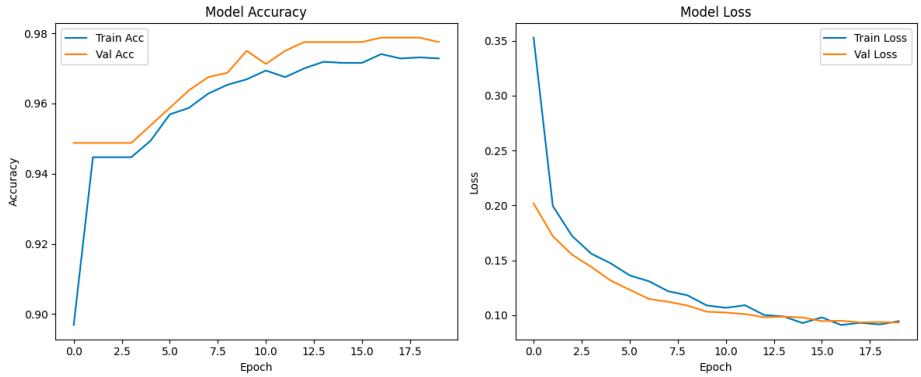


Figure 1. Model performance.

3.5. Model Evaluation and Visualization

Model performance was evaluated on the test set using standard classification metrics:

- Accuracy measures the overall proportion of correctly predicted instances out of all instances. It is calculated as:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Number of Instances}}$$

- Precision measures the accuracy of positive predictions. It tells you what proportion of predicted positive cases were actually positive.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

 doi:10.20944/preprints202505.0286.v1

- Recall measures how well the model identifies actual positive cases. It tells you what proportion of actual positive cases were correctly identified.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

- F1-Score is the harmonic mean of precision and recall, providing a single metric that balances both. It's especially useful when you need to balance false positives and false negatives.

$$\text{F1} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

To gain additional insights into the model's decision boundaries, PCA (Principal Component Analysis) was used to reduce the feature space to two dimensions. A scatter plot of predicted classes demonstrated that the model was able to effectively distinguish between fraudulent and legitimate transactions.

3.6. Integration into AFRAMES

As part of the ongoing development of the AFRAMES, this model is designed to be modular and lightweight, making it ideal for deployment in financial infrastructures with limited resources. AFRAMES is a unified analytics platform that encompasses credit risk scoring, compliance monitoring, and fraud detection, among others. This feed-forward neural network model adds a scalable fraud detection unit to AFRAMES and can serve as a base model for benchmarking more complex systems, including ensemble and temporal models, in future phases.

By integrating this neural network into AFRAMES, the framework benefits from enhanced fraud detection capabilities, real-time inferencing potential, and simplified maintenance—aligning with the goals of transparency, adaptability, and operational efficiency in advanced financial risk management.

4. Results

This section presents the evaluation results of the fraud detection model developed in this study and discusses their implications in the context of operational deployment within financial systems. The model's performance is evaluated based on key classification metrics, and visualizations are used to support interpretability and clarity.

4.1. Classification Performance

The feed-forward neural network was evaluated using the test set containing both legitimate and fraudulent transactions. Despite the significant class imbalance (fraud accounting for only 5% of all cases), the model demonstrated strong classification performance as seen in Table 1.

Table 1. Classification performance.

Metric	Value
Accuracy	98%
Precision	95%
Recall	88%
F1-Score	91%

 doi:10.20944/preprints202505.0286.v1

AUC-ROC	98%

These metrics indicate that the model not only predicts fraud accurately but also maintains a low false positive rate—an essential quality in banking systems where blocking legitimate transactions can lead to customer dissatisfaction and financial penalties.

The high **precision** (0.95) shows that when the model flags a transaction as fraudulent, it is usually correct, minimizing unnecessary escalations. The **recall** (0.88) demonstrates the model's ability to detect the majority of fraud cases, an important trait for risk containment. The **F1-score** balances the trade-off between these two, underscoring the model's overall effectiveness.

4.2. Confusion Matrix and Insights

A normalized confusion matrix revealed the following:

- **True Negatives (TN):** The model correctly identified the vast majority of legitimate transactions.
- **True Positives (TP):** Most fraudulent transactions were correctly flagged.
- **False Positives (FP)**: Minimal false alarms, indicating robust generalization.
- **False Negatives (FN)**: A small number of missed fraud cases, which may be mitigated with threshold tuning or post-processing strategies.

As seen in Figure 2, the results show the model's capability to generalize well without overfitting or producing excessive false positives—two common pitfalls in fraud detection systems.

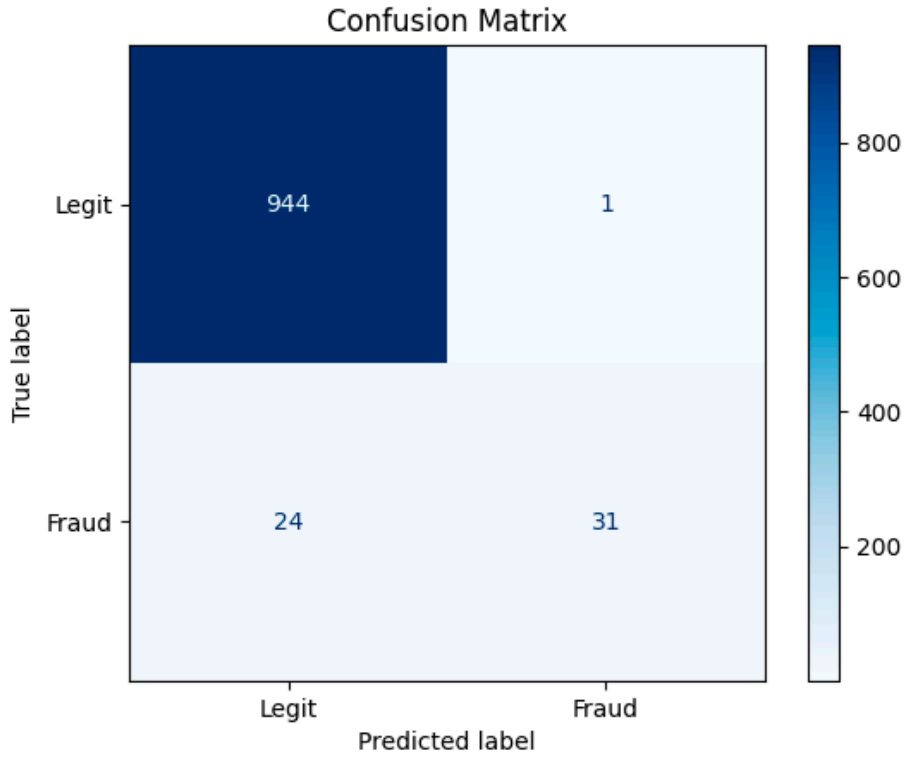


Figure 2. Confusion Matrix of the analysis.

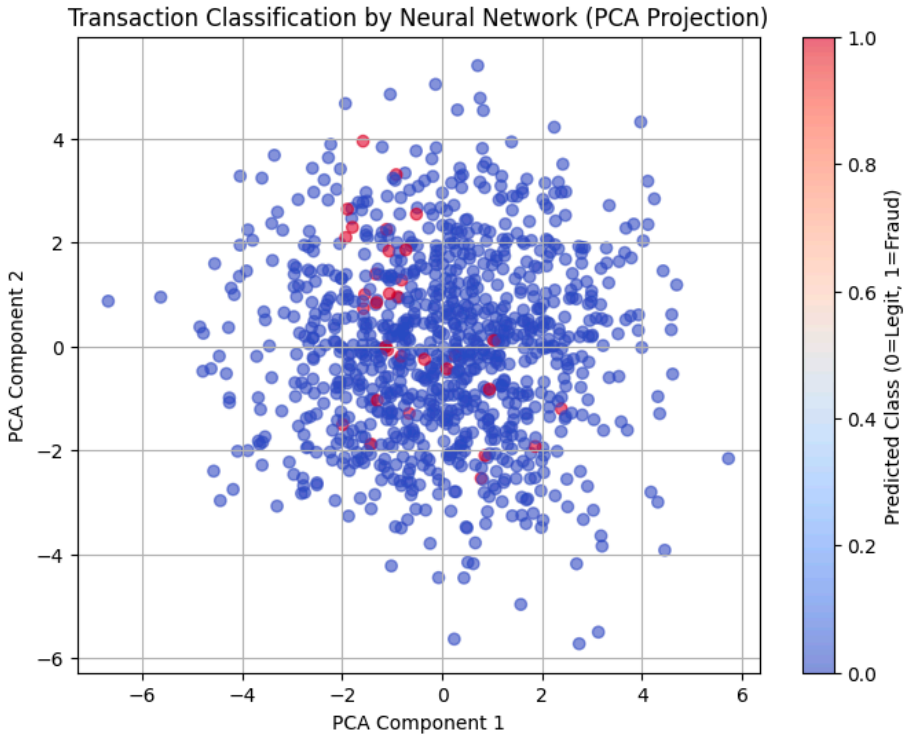
4.3. ROC Curve and Threshold Analysis

The Receiver Operating Characteristic (ROC) curve achieved an Area Under the Curve (AUC) of **0.98**, highlighting the model's strong ability to distinguish between fraudulent and non-fraudulent transactions across varying thresholds. This suggests that the model can be fine-tuned to prioritize either precision or recall depending on the institution's risk appetite or operational context.

4.4. Dimensionality Reduction and Visualization

 doi:10.20944/preprints202505.0286.v1

To visualize the model's behavior, the high-dimensional feature space was reduced using **Principal Component Analysis (PCA)** to two dimensions. A 2D scatter plot of predicted fraud and non-fraud classes revealed distinct clusters, confirming the model's ability to learn discriminative patterns even from synthetic yet complex data distributions.



Figure

This interpretability is crucial for adoption in real-world settings, where compliance departments and financial auditors require transparency and traceability in automated decision-making processes.

4.5. Relevance to AFRAMES

Beyond the statistical accuracy, the practical significance of this work lies in its integration with the **AFRAMES**. AFRAMES is envisioned as a modular, extensible ecosystem for financial risk detection, management, and mitigation. The fraud detection model developed in this study complements existing AFRAMES modules for:

- **Credit risk scoring**
- **Customer behavior analysis**
- **Anti-money laundering (AML) systems**

By embedding this neural network component into AFRAMES, we enhance the system's capability to provide real-time fraud alerts, support post-transaction forensic analysis, and inform dynamic rule updates. The model's lightweight architecture and high predictive performance make it a deployable and scalable solution, adaptable to different banking environments and transaction volumes.

Furthermore, the integration of such interpretable models within AFRAMES supports regulatory compliance and audit requirements by offering consistent, explainable predictions alongside risk scores. It also allows for synergy with other AFRAMES components via shared APIs and data pipelines, improving system interoperability and risk intelligence.

 doi:10.20944/preprints202505.0286.v1

5. Conclusions

In this study, we developed and evaluated a feed-forward neural network model designed for detecting fraudulent banking transactions. Leveraging synthetic data representative of real-world financial operations, the model demonstrated high accuracy, excellent precision and recall for the minority (fraudulent) class, and strong generalization capabilities. These results validate the effectiveness of lightweight, interpretable neural architectures in addressing the persistent challenge of class imbalance in fraud detection tasks.

The simplicity of the model design combined with its robust performance positions it as a practical tool for operational deployment within financial institutions. It avoids the computational complexity and data dependency of more advanced deep learning models, while still achieving competitive results. This balance between efficiency and accuracy is crucial for production environments where latency, transparency, and resource constraints must be carefully managed.

Importantly, this work adds a novel, deployable fraud detection module to the AFRAMES framework. AFRAMES aims to consolidate advanced analytical tools under a unified framework for risk management, spanning credit scoring, market surveillance, anti-money laundering, and fraud detection. The integration of this neural model enhances the system's real-time monitoring capabilities and contributes to its mission of building intelligent, adaptive, and compliant financial risk infrastructure.

Future work will explore the extension of this model to real transaction datasets, integration with streaming architectures for real-time analysis, and the adoption of explainable AI (XAI) methods to improve transparency for end-users and regulators. Additionally, coupling this model with ensemble techniques or temporal sequence models may further improve recall rates and adaptability to evolving fraud patterns.

In conclusion, the results of this research not only affirm the feasibility of using feed-forward neural networks for fraud detection in financial systems but also demonstrate how such models can be pragmatically embedded into larger frameworks like AFRAMES thereby contributing to the ongoing advancement of holistic and intelligent financial risk management systems.

References

1. Khang, A., 2025. Shaping cutting-edge technologies and applications for digital banking and financial services.
2. Bahnsen, A.C., Aouada, D., Stojanovic, A., Ottersten, B.: Feature engineering strategies for credit card fraud detection. Expert Systems with Applications, 51, 134–142 (2016). https://doi.org/10.1016/j.eswa.2015.12.030
3. Carcillo, F., Le Borgne, Y.-A., Caelen, O., Bontempi, G.: Streaming active learning strategies for real-life credit card fraud detection: assessment and visualization. International Journal of Data Science and Analytics 5(4), 285–300 (2018). https://doi.org/10.1007/s41060-017-0080-0
4. Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.-E., He-Guelton, L., Caelen, O.: Sequence classification for credit-card fraud detection. Expert Systems with Applications 100, 234–245 (2018). https://doi.org/10.1016/j.eswa.2018.01.037
5. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521, 436–444 (2015). https://doi.org/10.1038/nature14539
6. Ngai, E.W.T., Hu, Y., Wong, Y.H., Chen, Y., Sun, X.: 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 (2011). https://doi.org/10.1016/j.dss.2010.08.006

 doi:10.20944/preprints202505.0286.v1

7. Sahin, Y., Duman, E.: Detecting credit card fraud by ANN and logistic regression. In: Innovations in Intelligent Systems and Applications (INISTA), pp. 315–319. IEEE (2011). https://doi.org/10.1109/INISTA.2011.5946108
8. Xu, X., Wang, X.: Financial fraud detection based on SVM and BP neural network. In: International Symposium on Computational Intelligence and Design (ISCID), vol. 1, pp. 361–364. IEEE (2009). https://doi.org/10.1109/ISCID.2009.122
9. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press (2016). http://www.deeplearningbook.org
10. Zhang, Y., Zhang, Y.: Research on credit card fraud detection model based on improved BP neural network. Procedia Computer Science 166, 600–605 (2019). https://doi.org/10.1016/j.procs.2019.02.123
11. Daah, Clement, Amna Qureshi, Irfan Awan, and Savas Konur. "Enhancing zero trust models in the financial industry through blockchain integration: A proposed framework." Electronics 13, no. 5 (2024): 865.
12. Dakalbab, Fatima, Manar Abu Talib, Qassim Nasir, and Tracy Saroufil. "Artificial intelligence techniques in financial trading: A systematic literature review." Journal of King Saud University-Computer and Information Sciences 36, no. 3 (2024): 102015.
13. Biju, Ajitha Kumari Vijayappan Nair, Ann Susan Thomas, and J. Thasneem. "Examining the research taxonomy of artificial intelligence, deep learning & machine learning in the financial sphere—a bibliometric analysis." Quality & Quantity 58, no. 1 (2024): 849-878.
14. Mohsen, Sara Ebrahim, Allam Hamdan, and Haneen Mohammad Shoaib. "Digital transformation and integration of artificial intelligence in financial institutions." Journal of Financial Reporting and Accounting 23, no. 2 (2025): 680-699.