LIMES PLUS Journal of Social Sciences and Humanities
ISSN: 1820-0869
E-ISSN: 2406-2871
Vol. 23 (2) 2026

Machine Learning Approaches for Fraud Detection and Financial Risk Management

Choi Min-Jae
Department of Data Science, Danubia Research University, Romania

Received: 10.02.2026 Accepted: 12.06.2026 Publishing: 10.08.2026

Abstract
Financial fraud and risk have become increasingly complex as financial transactions move across digital platforms, banking systems, payment networks, and interconnected business environments. Machine learning provides analytical approaches that can identify unusual transaction patterns, classify potentially fraudulent activities, estimate risk, and support continuous monitoring. This paper examines machine learning approaches for fraud detection and financial risk management, focusing on supervised learning, unsupervised anomaly detection, transaction monitoring, credit risk assessment, behavioural analysis, model evaluation, real-time detection, explainability, and implementation challenges. Machine learning can process large volumes of financial data and identify relationships that may be difficult to detect through traditional rule-based systems. Supervised models can learn from historical examples of fraudulent and legitimate transactions, while unsupervised approaches can identify unusual patterns when labelled fraud data are limited. These methods can improve detection speed and support more targeted investigations. However, financial fraud detection involves significant challenges, including imbalanced datasets, false positives, changing fraud strategies, data quality, privacy, model drift, explainability, and regulatory requirements. The paper argues that machine learning should complement rather than completely replace established controls and professional investigation. Effective systems require high-quality data, continuous monitoring, appropriate model evaluation, human oversight, and strong governance. When integrated responsibly with existing financial control systems, machine learning can strengthen fraud prevention, risk assessment, operational efficiency, and organizational resilience.
Keywords
Machine Learning; Fraud Detection; Financial Risk Management; Anomaly Detection; Financial Fraud; Predictive Analytics; Risk Assessment; Cybersecurity; Financial Technology

Introduction

Financial institutions and businesses face persistent risks from fraudulent transactions, identity misuse, cyber-enabled financial crime, credit default, money laundering, and other forms of financial misconduct. The growth of digital payments and online financial services has increased both the speed and scale at which transactions occur.

Traditional fraud detection systems often depend on predefined rules and thresholds. While such rules remain useful, fraud patterns can change rapidly and may involve complex relationships that are difficult to capture through fixed conditions.

Machine learning provides an alternative analytical approach by learning patterns from financial data. Models can classify transactions, identify anomalies, estimate probabilities, and prioritize cases for investigation.

The effectiveness of machine learning depends heavily on data quality and model design. Financial fraud datasets are often highly imbalanced because legitimate transactions greatly outnumber fraudulent ones. Models must therefore be evaluated using appropriate measures rather than accuracy alone.

This paper examines machine learning approaches for fraud detection and financial risk management through nine dimensions: fraud analytics, supervised learning, anomaly detection, transaction monitoring, credit risk, behavioural analysis, evaluation, real-time systems, and governance.

1. Machine Learning and Financial Fraud Detection

Machine learning can support fraud detection by identifying patterns associated with unusual or suspicious financial activity. Relevant information may include transaction amount, time, location, merchant characteristics, account history, device information, and transaction frequency.The objective is not simply to label every unusual transaction as fraud. Instead, models can estimate risk or identify cases that deserve further review.Machine learning can therefore function as a screening and decision-support mechanism within broader financial control systems.

2. Supervised Machine Learning Approaches

Supervised learning uses historical data containing known examples of fraudulent and legitimate transactions. Algorithms can learn relationships between input features and outcome labels and then apply these patterns to new transactions.Common approaches include logistic regression, decision trees, random forests, support vector machines, and gradient-boosting methods. Neural networks can also be used when sufficient data and computational resources are available.The usefulness of supervised models depends on the quality, representativeness, and timeliness of labelled fraud data. New forms of fraud may not resemble historical examples.

3. Unsupervised Learning and Anomaly Detection

Unsupervised approaches can identify unusual observations without requiring every transaction to have a fraud label. Clustering and anomaly-detection methods can highlight transactions that differ substantially from typical behaviour.These approaches can be useful when labelled fraud cases are limited or when organizations need to discover previously unknown patterns.An unusual transaction is not necessarily fraudulent. Therefore, anomaly detection usually works best when combined with investigation, contextual information, and other risk indicators.

4. Transaction Monitoring and Behavioural Analysis

Machine learning can analyze transaction sequences and behavioural patterns over time. Instead of examining individual transactions in isolation, models can consider changes in spending frequency, locations, devices, amounts, or account activity.Behavioural analysis can help identify account takeover, unusual purchasing patterns, and other suspicious activity. Continuous monitoring may allow organizations to detect potential problems sooner. Effective monitoring requires timely data and systems capable of integrating information from relevant channels.

5. Machine Learning for Credit and Financial Risk Assessment

Beyond fraud detection, machine learning can support financial risk management by estimating credit risk, default probability, liquidity pressures, and other indicators. Models can analyze financial histories and behavioural data to identify patterns associated with risk. Predictive risk models can help organizations prioritize reviews and allocate resources. They may also support lending and portfolio-management decisions. Because financial risk decisions can affect individuals and businesses, models should be tested for fairness, stability, and explainability. Historical data may contain biases that influence predictions.

6. Model Evaluation and the Problem of False Positives

Fraud detection systems must balance the identification of genuine fraud with the avoidance of unnecessary alerts. A model that produces too many false positives can create investigation costs and inconvenience legitimate customers. Evaluation should therefore consider measures such as precision, recall, F1 score, specificity, area under the ROC curve, and, where appropriate, precision-recall curves. The most relevant measure depends on the business objective and the relative costs of different errors. Thresholds should also be reviewed as fraud patterns and organizational priorities change.

7. Real-Time Fraud Detection and Financial Technology

Digital financial services increasingly require rapid analysis because transactions can occur within seconds. Machine learning models can be integrated into payment and banking systems to assess risk during or immediately after a transaction.

Real-time detection can support rapid intervention, such as additional authentication or temporary review. However, real-time systems must balance security with customer experience and system performance The reliability of these systems depends on data latency, computational capacity, model availability, and robust technical infrastructure.

8. Explainability, Privacy, and Governance

Financial institutions need to understand how risk models influence decisions, particularly when automated outputs affect customers or financial access. Explainability can help investigators, managers, auditors, and regulators assess whether models are functioning appropriately. Financial data are also highly sensitive. Organizations must implement appropriate access controls, data protection, security measures, and retention policies. Governance should include model validation, documentation, monitoring, bias assessment, change management, and clear accountability for decisions supported by machine learning.

9. Challenges and Future Directions

Machine learning fraud detection faces challenges including data imbalance, adversarial behaviour, changing fraud strategies, model drift, incomplete labels, privacy requirements, and the difficulty of detecting novel attacks. Future systems are likely to combine multiple models, graph-based analysis, behavioural analytics, real-time data, and human investigation. Explainable and adaptive models may become increasingly important as organizations seek both accuracy and accountability. Further research should examine how machine learning can detect emerging fraud while reducing false positives and maintaining fairness, privacy, security, and regulatory compliance.

Conceptual Framework

The following framework presents a simplified relationship between machine learning, financial data and predictive models, and outcomes such as fraud detection, risk assessment, and stronger financial controls. Transaction patterns, model accuracy, false positives, explainability, monitoring, and governance can influence these relationships.

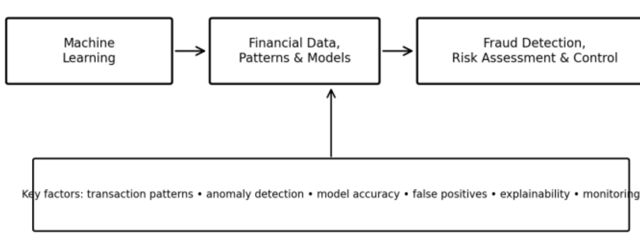

Figure 1. Conceptual framework showing machine learning approaches for fraud detection and financial risk management.

ISSN: 1820-0869
E-ISSN: 2406-2871
Vol. 23 (2) 2026

Conclusion

Machine learning offers significant potential for strengthening fraud detection and financial risk management by analyzing large volumes of transaction and behavioural data. Supervised learning can classify known patterns, while unsupervised methods can identify unusual activity and support the discovery of emerging risks.

The effectiveness of these systems depends on more than algorithmic accuracy. Financial organizations must manage false positives, changing fraud patterns, data quality, privacy, model explainability, and regulatory requirements. Human investigation and established internal controls remain important components of a comprehensive fraud-management system.

Responsible integration of machine learning can improve detection speed, risk assessment, monitoring, and resource allocation. Future approaches should focus on adaptive models, transparent governance, high-quality data, and strong collaboration between technical specialists, financial professionals, investigators, and regulators.

References

Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. Statistical Science, 17(3), 235–255. https://doi.org/10.1214/ss/1042727940

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. Decision Support Systems, 50(3), 559–569. https://doi.org/10.1016/j.dss.2010.08.006

West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. Computers & Security, 57, 47–66. https://doi.org/10.1016/j.cose.2015.09.005

Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. arXiv.

Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.-E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. Expert Systems with Applications, 100, 234–245. https://doi.org/10.1016/j.eswa.2018.01.037

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). 'Why should I trust you?': Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. European Journal of Operational Research, 247(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030