JFDA THE JOURNAL OF FINTECH & DIGITAL ASSETS

JFDA, Volume 1 (2), 2026

| Research Article

Financial Fraud Detection: Machine Learning vs. Traditional Rule-Based Systems - A Comparative Analysis of Model.

Haixia Du*

Authors
Haixia Du, Independent Researcher, Auckland, New Zealand.
Email: cassiedu0827@gmail.com
* Corresponding author.

Abstract
With the rapid development and widespread use of electronic payments and digital financial transactions, financial fraud has become increasingly common. It has already caused significant threats and losses to customers and financial institutions. How to detect financial fraudulent activities in a timely and accurate manner has become crucial, as it can not only reduce financial losses but also improves customers' trust in the financial system. Traditional rule-based systems rely on manually defined thresholds and patterns, which can identify anomalous behaviors to a certain extent, but it is difficult to identify complex fraudulent behaviors in the face of continuously changing fraud patterns. In contrast, machine learning provides a dynamic, data-driven approach that is more adaptive and accurate, and it can learn hidden patterns from large datasets. This study uses a simulated credit card transaction dataset generated by the Sparkov tool, trained and tested under the same dataset conditions. Standard pre-processing steps were applied to compare a traditional rule-based system with four machine learning models, including logistic regression, decision trees, random forests and support vector machines (SVM), and to use metrics such as accuracy, precision, recall, F1-score and AUC-ROC to evaluate the model performance. The experimental results show that among all the methods, the random forest model in machine learning performs the best in terms of overall evaluation, showing strong fraud detection ability and low false alarm rate. The decision tree model also performs relatively well with high recall. Although the SVM model has high recall, it falls behind the F1-score and precision. Logistic regression has relatively high recall but very low precision, meaning it produces many false positives. On the other hand, the traditional rule-based system is relatively highly accurate, but its precision and recall are poor, which means it has a high false alarm rate and the actual detection effect is limited. Overall, the results show that machine learning models, especially Random Forests, perform well compared to the traditional rule-based system under the same data conditions. The findings indicate that machine learning models outperform traditional rule-based systems under the same conditions and are more suitable for financial fraud detection in practice.

Keywords: Financial fraud detection, Credit card transactions, Machine learning, Rule-based systems, Random forest
JEL Classification: C45, G21, G28

Article Information
Received: 26 January 2026
Revised: 28 June 2026
Accepted: 10 July 2026
Published: September 2026

Citation:
Du, H. (2026). Financial fraud detection: Machine learning vs traditional rule-based systems - A comparative analysis of model. *The Journal of FinTech and Digital Assets*, 1(2), 45–53.

1. Introduction

Financial fraud issues have become increasingly serious and appear more frequently in daily life. With the growth of different types of fraud, such as credit card theft, forged identities, email scams, and fake transactions, financial fraud not only causes large financial losses to customers and financial institutions but

also creates a crisis of trust in the financial system. Traditional financial systems commonly rely on rule-based fraud detection methods, such as screening transaction behaviors using predefined logical rules. However, the limitations of this approach have become more evident, as it is difficult to quickly and accurately identify fraud in the presence of complex and continuously evolving fraudulent behaviors (West and Bhattacharya, 2016).

Machine learning models provide an alternative to traditional rule-based systems. Unlike rule-based approaches, machine learning models can automatically learn fraud patterns from large volumes of historical transaction data and demonstrate stronger adaptability and predictive capability (Ngai et al., 2011). Many studies have shown that machine learning models achieve higher precision and recall in fraud detection. However, after reviewing existing literature, it can be observed that there is still a lack of experimental research that systematically compares traditional rule-based systems with different machine learning models under the same dataset and evaluation conditions.

This study aims to address this gap by using a publicly available simulated credit card transaction dataset. Under the same dataset and evaluation framework, a traditional rule-based system is compared with several machine learning models, including logistic regression, decision tree, random forest, and support vector machine. The models are evaluated using accuracy, precision, recall, F1-score, and AUC, to identify more effective and adaptive approaches for financial fraud detection.

The remainder of this paper is structured as follows. Section 2 reviews the existing literature on financial fraud detection. Section 3 describes the dataset. Section 4 outlines the methodology. Section 5 presents empirical results. Section 6 concludes the study.

2. Literature Review

Financial fraud has long been one of the key topics studied in the field of financial learning. Early fraud detection systems mainly relied on traditional rule-based methods, which require predefined thresholds to identify suspicious transactions. However, with constantly evolving fraud techniques, traditional rule-based detection methods often lack flexibility and practical applicability.

West and Bhattacharya (2016) conducted a comprehensive review of traditional rule-based systems and pointed out that these detection systems exhibit low responsiveness and limited adaptability. They also emphasized the advantages and potential of machine learning approaches in dealing with dynamic fraud patterns. This perspective provides important theoretical support for the present study. Ngai et al. (2011) proposed an operational classification framework to systematically organize the application of data mining techniques in financial fraud detection. They classified methods such as decision trees, neural networks, and support vector machines, which provided a theoretical basis for algorithm selection and background for this research study.

Bhattacharyya et al. (2011) analyzed the performance of multiple supervised learning models using real-world credit card transaction data. Their results showed that non-linear models typically outperform linear models in fraud detection tasks, which guided this study to prioritize non-linear models such as decision trees and random forests. Ryman-Tubb et al. (2018), by investigating the practical application of artificial intelligence in payment card fraud detection, identified several key challenges in current industry practices, including model interpretability, real-time responsiveness, and system integration. Their findings indicate that model accuracy alone is insufficient, and feasibility in real systems must also be considered for practical deployment.

Al-Hashedi and Magalingam (2021) provided an overview of data mining methods applied to financial fraud detection between 2009 and 2019, summarizing a wide range of mainstream fraud detection approaches. Riskiyadi (2024) and Ramzan and Lokanan (2024) demonstrated the broad applicability of machine learning methods by applying them to financial statement fraud and accounting fraud detection, respectively. However, these studies largely lacked experimental designs that directly compare machine learning approaches with traditional methods. This limitation highlights a current research gap and aligns with the objective of the present study.

A frequently discussed issue identified in the literature is data imbalance. As fraudulent transactions usually account for only a small proportion of total transaction volumes, predictive models tend to favor non-fraudulent cases. Carcillo et al. (2021) proposed a hybrid approach combining unsupervised and supervised learning to address data imbalance and improve model robustness on highly imbalanced datasets. This literature also motivated the use of the Synthetic Minority Oversampling Technique (SMOTE) to oversample the dataset to enhance the model's ability to detect fraudulent behavior.

Afriyie et al. (2023) validated the application of supervised machine learning algorithms in real-time credit card fraud detection and provided useful references for selecting appropriate model evaluation metrics. Stojanović et al. (2021) focused on the strategic application of machine learning for fraud detection in the FinTech domain and emphasized the importance of model performance in real-world applications. This further reinforces the motivation to conduct multi-model comparative experiments, particularly to identify a balance between deplorability and performance.

Although existing studies provide strong theoretical foundations and guidance for methodological selection, most research evaluates only a single machine learning model or conducts limited comparisons. Systematic comparative studies between traditional rule-based systems and machine learning methods under the same dataset and evaluation conditions remain limited. Therefore, this study aims to address this gap by comparing multiple approaches to better understand their respective strengths and weaknesses and to identify more effective solutions for financial fraud detection.

3. Data and Methodology

Before presenting the dataset and experimental procedures, the basic principles of the fraud detection methods are briefly introduced. This study compares a traditional rule-based system with four supervised machine learning algorithms, including logistic regression, decision tree, random forest, and support vector machine (SVM). These methods were selected because they represent different approaches to fraud detection and allow a comparison of their performance under the same experimental conditions (Ngai et al., 2011; Bhattacharyya et al., 2011).

Logistic regression is a supervised learning algorithm widely used for binary classification problems. It estimates the probability that a transaction is fraudulent by using the logistic (sigmoid) function. The probability is calculated as:

$$P(y = 1|x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\cdots+\beta n x_n)}} \qquad (1)$$

where $P(y=1|x)$ represents the probability that a transaction is fraudulent, $x_1, \ldots, x_n$ are the predictor variables, and $\beta_0, \beta_1, \ldots, \beta_n$ are the model coefficients. If the predicted probability is greater than a predefined threshold, the transaction is classified as fraudulent. According to Ngai et al. (2011), logistic regression has a simple structure and is easy to interpret. It is therefore commonly used as a baseline model for comparison with other machine learning algorithms.

A decision tree is a nonlinear supervised learning algorithm that classifies data by splitting the dataset based on selected feature values. During model construction, the Gini index is commonly used to evaluate the quality of each split. The Gini index is calculated as:

$$Gini = 1 - \sum_{i=1}^{C} p_i^2 \qquad (2)$$

where $p_i$represents the proportion of observations belonging to class i. The process continues until a predefined stopping criterion is reached. A decision tree is easy to understand because its classification process can be represented as a series of decision rules. However, when applied to complex datasets, decision trees may become overly complex and are more likely to overfit the data (Bhattacharyya et al., 2011).

Random forest is an ensemble learning algorithm based on decision trees. It constructs multiple decision trees using bootstrap sampling and randomly selected feature subsets. The final classification result is determined by majority voting. Compared with a single decision tree, random forest provides more stable predictions, reduces the risk of overfitting, and achieves better classification performance. Therefore, random forest is widely used in financial fraud detection because of its strong classification performance (Bhattacharyya et al., 2011).

Support vector machine (SVM) is a supervised learning algorithm that classifies data by finding the optimal separating hyperplane. For more complex classification problems, kernel functions can be used to map the data into a higher-dimensional feature space and construct a nonlinear decision boundary. SVM has strong classification capability. However, SVM has a high computational cost when processing large datasets, which may limit its application in some situations (Ngai et al., 2011).

Unlike machine learning models, traditional rule-based systems rely on manually defined IF–THEN rules and predefined thresholds to identify suspicious transactions. These rules are usually developed based on expert knowledge and historical fraud cases. When a transaction meets the predefined conditions, it is flagged for further investigation. Although rule-based systems are simple to implement and easy to understand, they cannot automatically learn from historical data and therefore have limited ability to adapt to new fraud patterns. West and Bhattacharyya (2016) pointed out that the performance of rule-based systems gradually declines as fraud methods continue to evolve.

The following section describes the dataset and experimental procedures used in this study. The dataset used in this study is obtained from a fraud detection benchmark provided by Amazon Science. It is a simulated credit card transaction dataset generated using the Sparkov tool, designed to reflect real-world transaction behaviour with a clear and well-defined structure. The dataset captures key behavioural characteristics of credit card transactions while avoiding the disclosure of sensitive user information, making it suitable for training and testing financial fraud detection models.

The dataset is pre-divided into two subsets: fraudTrain.csv, which is used for model training, and fraudTest.csv, which is used for model testing and performance evaluation. Each transaction contains multiple feature variables, such as transaction amount, merchant category, and user gender, as well as a binary target variable, *is_fraud*, indicating whether a transaction is fraudulent (1) or non-fraudulent (0). This predefined data split facilitates model development and ensures consistent performance comparison across different algorithms. Both subsets preserve the original distribution of fraudulent and non-fraudulent transactions.

Several factors motivated the selection of this simulated dataset. First, it closely mimics real-world credit card transaction behaviour. Second, it avoids issues related to customer privacy and data confidentiality. In addition, the use of a publicly available dataset enhances the reproducibility of this research, enabling comparisons with existing studies and supporting future academic work.

This study adopts a quantitative research methodology based on secondary data and incorporates binary classification supervised learning models to systematically evaluate the performance of various fraud detection approaches. A descriptive comparative research design is employed to compare the performance of traditional rule-based detection methods with multiple machine learning algorithms under the same dataset and experimental conditions, with the aim of identifying more reliable and adaptable solutions for real-world financial fraud detection.

The sample used in this study is derived from the simulated transaction dataset provided by Amazon Science, which contains approximately 200,000 transaction records covering both normal and fraudulent behaviour. Each transaction record is independent, making the dataset suitable for supervised learning modelling.

Prior to model training, standard data preprocessing procedures were applied, including handling missing and duplicate values, encoding categorical variables into numerical representations, and standardising continuous variables such as transaction amounts. As fraudulent transactions represent only a small proportion of the dataset, a severe class imbalance problem exists. To address this issue, the Synthetic Minority Oversampling Technique (SMOTE) was applied to the training set to synthetically generate minority class samples, aiming to achieve a 1:1 ratio between fraudulent and non-fraudulent transactions. As a result, the training dataset was expanded to approximately 2.56 million samples, reducing bias toward the majority class and improving the model's ability to identify fraudulent behaviour.

This study constructs and compares a traditional rule-based detection system and several machine learning models. The selected machine learning algorithms include logistic regression, decision tree, random forest, and support vector machine. All models are trained on the same training dataset and tested on the same testing dataset to ensure comparability of evaluation results.

Model performance is evaluated using consistent metrics, including accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC-ROC). Performance assessment considers overall metrics such as accuracy and AUC, while placing particular emphasis on fraud-sensitive metrics such as recall and F1-score.

As this study uses a publicly available simulated dataset, no field surveys or participant recruitment are required. This reduces data collection time and allows the research to focus on model construction and performance evaluation. The dataset is structurally complete, contains rich features, and does not involve real user information or sensitive data, ensuring compliance with legal and ethical requirements.

The main challenge encountered during the experiment relates to the training efficiency of the support vector machine (SVM) model. In particular, SVM with an RBF kernel requires substantial computational resources. Training the model on the full dataset resulted in memory overflow and program crashes in the Google Colab environment. To address this issue, the sample size was reduced, and different cross-validation strategies were tested. The final approach involved using a subset of 30,000 samples combined with 10-fold cross-validation. Consequently, all models were trained on the same 30,000-sample subset to maintain comparability while addressing computational constraints. The results obtained from this subset were stable and reproducible (random_state = 42), indicating that it was both representative of fraud patterns and computationally feasible.

Overall, the dataset and methodology provide a solid foundation for evaluating fraud detection systems and ensure consistent experimental conditions for comparing different fraud detection methods.

4. Empirical Results and Discussion

To evaluate the effectiveness of traditional rule-based systems and machine learning-based fraud detection methods, this study conducts a systematic comparison of five models: one traditional rule-based system and four supervised learning algorithms, namely logistic regression, decision tree, random forest, and support vector machine. All models are evaluated under consistent experimental conditions using the same test dataset and performance metrics, including accuracy, precision, recall, F1-score, and AUC-ROC, as demonstrated in Figures 1 and 2.

The random forest model achieves the best overall performance across all evaluation metrics. It records the highest recall, F1-score, and AUC, indicating strong capability in identifying fraudulent transactions while maintaining a relatively low false alarm rate. This result suggests that the random forest model provides the most balanced and practical fraud detection performance among the evaluated approaches.

The decision tree model also performs well, achieving high recall and good accuracy. However, its F1-score is lower than that of the random forest model, and its precision is slightly weaker, indicating a less balanced performance and a higher likelihood of false positives. Despite this limitation, its relatively simple structure makes it a potentially interpretable solution in practical applications.

The support vector machine (SVM) model performs well in terms of recall and AUC, demonstrating strong fraud detection capability. However, its precision and F1-score are relatively low, suggesting that although it identifies many fraudulent transactions, it also produces a higher false alarm rate, which may limit its effectiveness in real-world deployment. In addition, due to the high computational requirements of SVM for large datasets, the model initially crashed or timed out during training. By limiting the sample size to 30,000 and applying 10-fold cross-validation, stable and reproducible results were obtained. Although the performance is acceptable under these conditions, the cost–performance ratio of SVM remains inferior to that of the random forest model.

The logistic regression model performs the worst among the four machine learning methods. Although its recall is relatively high, its precision is only 2.29%, resulting in an F1-score of 0.044. This indicates that the model classifies nearly all transactions as fraudulent, leading to an extremely high false positive rate and very low practical usability.

The traditional rule-based system achieves high overall accuracy but exhibits very low recall and the lowest F1-score among all evaluated models. This highlights its limitations in fraud detection and reflects its lack of adaptability to complex and diverse fraud patterns, as fraudulent behaviours often deviate from static predefined rules.

Overall, the results demonstrate that machine learning models significantly outperform traditional rule-based systems in financial fraud detection. Among all evaluated methods, the random forest model shows the best balance between precision and recall, providing strong empirical support for its application in practical fraud detection systems. These findings also offer useful insights for future research exploring more advanced approaches, such as deep learning or hybrid detection models.

	Model	Accuracy	Precision	Recall	F1	AUC
1	LogisticRegression	0.876567	0.022890	0.743124	0.044412	0.908218
2	DecisionTreeClassifier	0.975903	0.131375	0.934266	0.230358	0.955165
3	RandomForestClassifier	0.984935	0.196688	0.941259	0.325383	0.994853
4	SupportVectorClassifier	0.946604	0.063269	0.929604	0.118475	0.980979
5	Rule-Based System	0.946604	0.063269	0.929604	0.118475	0.938137

Figure 1. Model comparison (created by the author, 20/06/2025)

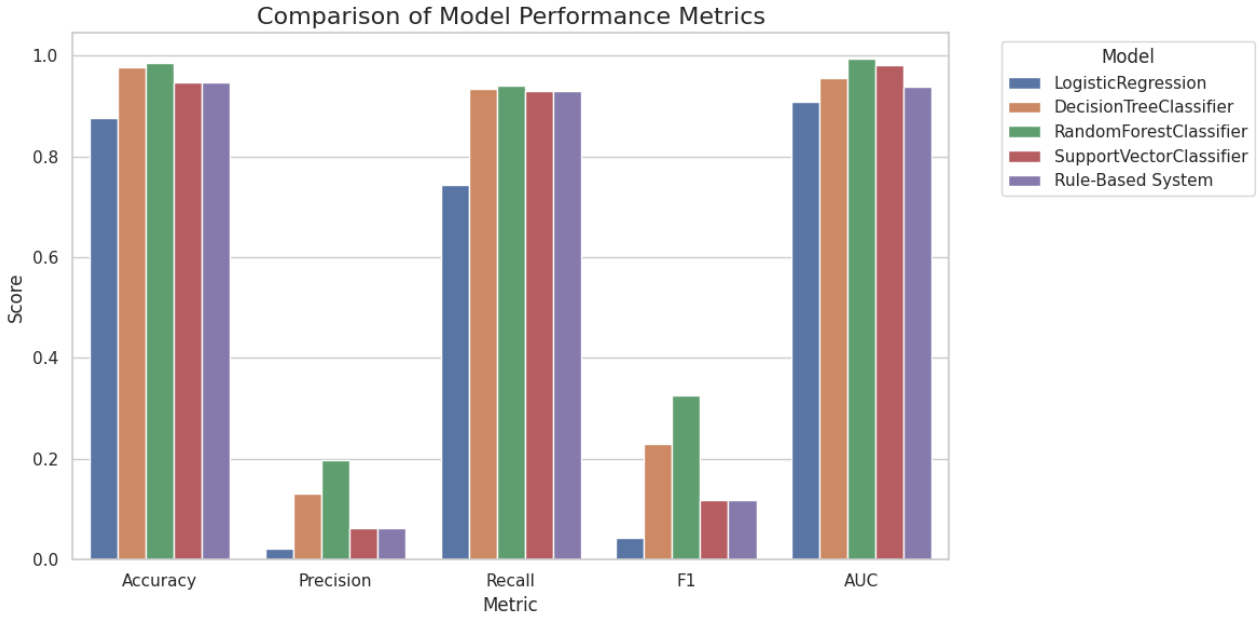

Figure 2. Comparison of model performance metrics (created by the author, 20/06/2025)

5. Conclusion

This study focuses on financial fraud detection and conducts a comprehensive comparative analysis of a traditional rule-based system and four supervised machine learning models, namely logistic regression, decision tree, random forest, and support vector machine, under consistent data conditions. A publicly available and well-structured dataset is employed, and all models are trained and tested under identical experimental settings. Multiple evaluation metrics, including accuracy, precision, recall, F1-score, and AUC, are used to ensure fairness and reliability in performance comparison.

The results clearly demonstrate that machine learning models significantly outperform the traditional rule-based system across key performance metrics, particularly in terms of fraud detection capability and false positive control. Among the evaluated models, the random forest model achieves the best overall performance. It accurately identifies most fraudulent transactions while effectively reducing false positives, making it highly suitable for real-world applications. The decision tree model, with its relatively simple structure, also performs well in terms of recall, which makes it a feasible option in scenarios where interpretability is a primary

requirement. The support vector machine shows stable performance and strong recognition ability; however, it requires substantial computational resources and is therefore less suitable for large-scale deployment in environments with limited capacity. Although acceptable results are achieved through sample down-sampling and cross-validation, deployment cost remains an important constraint.

Compared with machine learning models, the overall detection capability of the traditional rule-based system is relatively weak. Although it achieves high accuracy, it exhibits low recall and F1-score, indicating a high rate of missed fraud cases. This highlights the difficulty of relying on static rules to address the diversity and variability of modern financial fraud patterns. In comparison, traditional rule-based systems are more appropriate as preliminary screening mechanisms rather than primary fraud detection tools.

Overall, integrating machine learning models into financial fraud detection systems can bring significant benefits in terms of improved accuracy, flexibility, and adaptability. To enhance overall detection performance, financial institutions are encouraged to move beyond static rule-based systems and actively explore data-driven and intelligence-based solutions. Future research may extend to areas such as the introduction of deep learning models, real-time streaming data processing, or the development of hybrid approaches that combine rule-based logic with machine learning techniques. When selecting detection models, it is also essential to consider the balance between performance, resource constraints, and real-world deployment scenarios, in order to avoid one-size-fits-all strategies. Through continuous optimisation of algorithms and system architectures, financial fraud detection systems can become more resilient and better equipped to address increasingly complex and dynamic fraud challenges.

Funding

This research received no external funding.

Acknowledgements

The author would like to express sincere gratitude to Dr. David Gabauer and Dr. Cuong Nguyen for their guidance, continuous support, and valuable feedback throughout this research. The author also thanks the anonymous reviewers for their constructive comments and insightful suggestions, which have greatly improved the quality of this manuscript.

Declaration of the Use of Generative AI

Generative AI tools were used solely to improve the language, grammar, and readability of this manuscript. The research design, data analysis, interpretation of the results, and all academic decisions and final decisions related to this manuscript were undertaken solely by the author. The author takes full responsibility for the accuracy, originality, and integrity of this manuscript.

Conflict of Interest

The author declares no conflict of interest.

References

Afriyie, J. K., Tawiah, K., Pels, W. A., Addai-Henne, S., Dwamena, H. A., Owiredu, E. O., Ayeh, S. A., and Eshun, J. (2023). A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. Decision Analytics Journal, 6:100163.

Al-Hashedi, K. G. and Magalingam, P. (2021). Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. Computer Science Review, 40:100402.

Bhattacharyya, S., Jha, S., Tharakunnel, K., and Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. Decision Support Systems, 50(3):602–613.

Carcillo, F., Le Borgne, Y.-A., Caelen, O., Kessaci, Y., Oblé, F., and Bontempi, G. (2021). Combining unsupervised and supervised learning in credit card fraud detection. Information Sciences, 557:317–331.

Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., and 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.

Ramzan, S. and Lokanan, M. (2024). The application of machine learning to study fraud in the accounting literature. Journal of Accounting Literature.

Riskiyadi, M. (2024). Detecting future financial statement fraud using a machine learning model in indonesia: a comparative study. Asian Review of Accounting, 32(3):394–422.

Ryman-Tubb, N. F., Krause, P., and Garn, W. (2018). How artificial intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark. Engineering Applications of Artificial Intelligence, 76:130–157.

Stojanović, B., Božić, J., Hofer-Schmitz, K., Nahrgang, K., Weber, A., Badii, A., Sundaram, M., Jordan, E., and Runevic, J. (2021). Follow the trail: Machine learning for fraud detection in fintech applications. Sensors, 21(5):1594.

West, J. and Bhattacharya, M. (2016). Intelligent financial fraud detection: a comprehensive review. Computers & Security, 57:47–66.