Perbanas International Conference on Economics, Business, Management, Accounting and IT (PROFICIENT) 2026
"Integrity-Driven Innovation: Reimagining Business, Finance, and Digital Transformation for Sustainable Impact"

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

Predictive Modeling for Fraud Detection in Mobile Payment Systems with Decision Tree Method

Chelsea Aulia Zahra[1,a], Ines Kartika Dewi[2,b], Mercurius Broto Legowo[3,c]

[1,2,3] Data Science Study Program, Faculty of Information Technology, Perbanas Institute

[a)]Corresponding author : chelsea.aulia03@perbanas.id
[b)]ines.kartika05@perbanas.id
[c)] mercurius@perbanas.id

Abstract. *The rapid growth of digital technology has significantly increased the use of mobile payment systems as a fast, practical, and efficient financial transaction method. However, the growing volume and frequency of transactions have also led to a higher risk of fraudulent activities, which may cause financial losses, reduce user trust, and threaten the security of digital payment platforms. Detecting fraudulent transactions is a critical issue because transaction data is typically large, complex, and continuously generated. To address this problem, predictive modeling using machine learning can be applied as an effective solution. This study aims to develop a fraud detection model for mobile payment systems using the Decision Tree method. The research method includes data collection, preprocessing, feature selection, data splitting, model training, and evaluation using performance metrics such as accuracy, precision, recall, and F1-score. The expected results of this study are the successful classification of fraudulent and legitimate transactions with satisfactory performance, as well as the identification of the most influential features in fraud detection. This research is expected to contribute theoretically to the application of machine learning in fraud detection and practically to support financial technology providers in improving transaction security, decision-making processes, and more effective fraud prevention strategies in modern digital payment ecosystems.*

Keywords: Fraud Detection, Mobile Payment Systems, Predictive Modeling, Decision Tree, Machine Learning

INTRODUCTION

The rapid advancement of digital technology has significantly transformed financial transaction activities, particularly through the increasing adoption of mobile payment systems. Mobile payment platforms are widely used because they offer convenience, speed, flexibility, and efficiency in supporting cashless transactions across various daily activities. Recent industry reports indicate that digital payments and digital wallets continue to grow globally and are becoming dominant transaction methods in both online and point-of-sale environments (Allayannis et al., 2018; Eremeeva & Akulova, 2024). This phenomenon shows that mobile payment systems have become an essential part of the modern digital economy and financial ecosystem.

Despite these advantages, the growth of mobile payment systems has also been accompanied by increasing fraud risks. Fraudulent activities in digital payment environments may include unauthorized transactions, account takeover, phishing-based manipulation, identity misuse, and suspicious behavioral patterns. This issue is becoming more serious because payment fraud continues to be recognized as a major concern in digital financial services, while transaction data are typically high-volume, complex, and continuously generated in real time, making manual detection increasingly difficult (Grudniewski et al., 2025; Taavitsainen,2025). If not addressed properly, fraud can lead to financial losses, weaken user trust, and reduce the reliability of digital financial platforms.

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

To address this issue, predictive modeling using machine learning offers a practical and effective solution. Machine learning models can analyze historical transaction data, recognize hidden fraud patterns, and classify suspicious transactions more efficiently than conventional manual approaches. Among various classification techniques, the Decision Tree method is considered appropriate because it is relatively simple, interpretable, and capable of producing clear decision rules for fraud classification (Manorom et al., 2024). In addition, Decision Tree can help identify the most influential attributes in distinguishing fraudulent and legitimate transactions, which is useful for both detection and decision-making purposes.

Several previous studies conducted within the last three years have demonstrated the relevance of machine learning for fraud detection. A study by Xu et al. (2023) highlighted the importance of tree-based approaches in fraud detection because of their interpretability and strong classification capability, especially in imbalanced financial datasets (Xu et al., 2023). Manorom et al. (2024) compared several machine learning algorithms for fraudulent financial transaction detection and showed that Decision Tree remained a competitive and understandable classification method (Manorom et al., 2024). More recently, systematic reviews by Chen et al. (2025) and George et al. (2025) emphasized that fraud detection research continues to rely heavily on supervised learning methods, while also identifying ongoing challenges related to class imbalance, model performance, and interpretability (Chen et al., 2025; George et al., 2025). These findings indicate that Decision Tree remains relevant as a foundational method for predictive fraud modeling.

Based on this background, this study aims to develop a predictive modeling approach for fraud detection in mobile payment systems using the Decision Tree method. The target of this study is transaction data in mobile payment systems, which are classified into fraud and non-fraud categories and analyzed using the Decision Tree method. This research is expected to provide a clearer understanding of how transaction data can be classified into fraudulent and legitimate categories using interpret-able machine learning techniques. Furthermore, the study is expected to contribute theoretically to the application of machine learning in fraud detection and practically to support financial technology providers in improving transaction security, fraud prevention strategies, and data-driven decision-making. The originality of this research lies in applying the Decision Tree method to build a predictive model for detecting fraudulent transactions in mobile payment systems, which exhibit distinct data characteristics and transaction patterns compared to those of conventional banking systems. The novelty of this study lies in the development of a decision tree model that uses mobile payment-specific attributes to classify transactions as fraudulent or legitimate and to generate decision rules for real-time fraud detection.

LITERATURE REVIEW

Mobile Payment Systems

Mobile payment systems are digital financial services that allow users to perform transactions. The growing adoption of digital wallets and online payment platforms reflects a shift in financial behavior (Hafez et al., 2025), (Vlaicu, 2025). However, as the number of users and transactions increases, mobile payment systems also become more vulnerable to security threats and fraudulent activities.

Fraud Detection in Digital Transactions

Fraud detection refers to the process of identifying suspicious or unauthorized activities. In mobile payment systems, fraud may occur through unauthorized payments and abnormal transaction behavior. Transaction data are typically highly imbalanced (Hafez et al., 2025), (Taavitsainen, 2025).

Machine Learning for Fraud Detection

Machine learning has become an approach for fraud detection because it enables systems to learn patterns and classify transactions automatically. They rely on labeled datasets to distinguish fraudulent and legitimate transactions.

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

Previous studies indicate that machine learning improves detection performance (Hafez et al., 2025), (George et al., 2025). Evaluation metrics such as accuracy, precision, recall, and F1-score are essential.

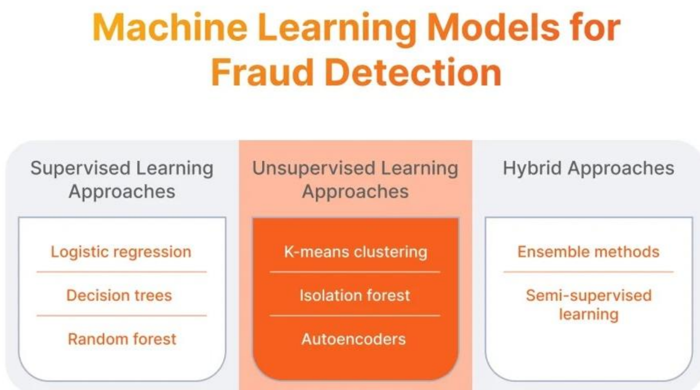

FIGURE 1. Machine Learning Model for Fraud Detection

The Figure 2, illustrates various **machine learning models used for fraud detection**, which are generally categorized into: supervised learning, unsupervised learning, and hybrid. **Supervised learning methods**, such as logistic regression, decision trees, and random forest, rely on labeled datasets where. These methods are widely used (Han et al., 2012), (Géron, 2022). On the other hand, **unsupervised learning methods**, including K-means clustering, Isolation Forest, and autoencoders, do not require labeled data. These techniques are used to detect anomalies patterns in transaction behavior (Géron, 2022), (Hafez et al., 2025). Furthermore, **hybrid approaches** combine both supervised and unsupervised techniques. Methods such as ensemble learning and semi-supervised learning (Hafez et al., 2025), (Ngai et al., 2011).

Decision Tree Method

A decision tree is a supervised learning algorithm used for classification. In fraud detection, a decision tree helps identify key attributes influencing fraudulent behavior. Compared to complex models, it provides better transparency, which is important in financial decision-making (Manorom et al., 2024).
The figure shows machine learning models for Decision Tree :

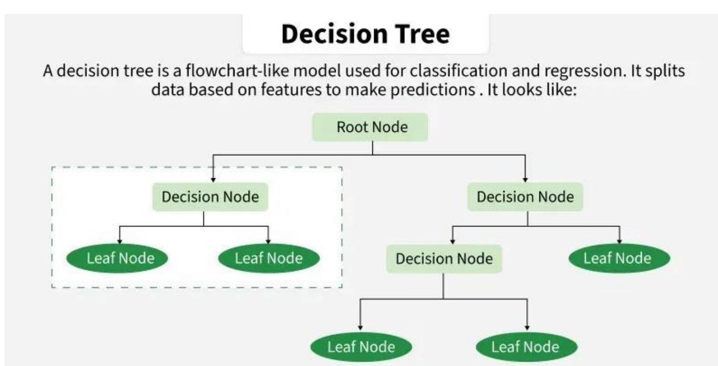

FIGURE 2. Decision Tree

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

A Decision Tree is a supervised machine learning method used for classification and regression tasks. Each path in the tree represents a decision rule used for classification, such as fraud or non-fraud transactions. Due to its simplicity and interpretability, Decision Tree is widely used in fraud detection systems.(Breiman et al., 1984), (Han et al., 2012)

Previous Research

Several studies machine learning in fraud detection. Compared multiple algorithms and found that Decision Tree remains competitive (Manorom et al., 2024). Demonstrated that tree-based models are effective for handling imbalanced fraud datasets Highlighted that machine learning continues to play a key role in fraud detection (Hafez et al., 2025), (George et al., 2025). Decision Tree is still relevant for predictive fraud modeling.

Research Gap

Although many studies have focused on fraud detection, several gaps remain. Most research emphasizes credit card fraud rather than mobile payment systems. Additionally, some high-performance models lack interpretability, making them less suitable for financial decision-making. The issue of imbalanced datasets also remains a major challenge, (Hafez et al., 2025), (George et al., 2025). Therefore, this study focuses on applying Decision Tree as an interpretable model for fraud detection in mobile payment systems.

Conceptual Framework

This study assumes that transaction data contain patterns. Through preprocessing, feature selection, and classification using Decision Tree, the model is expected to generate decision rules. The output is a predictive model evaluated using accuracy, precision, recall, and F1-score.

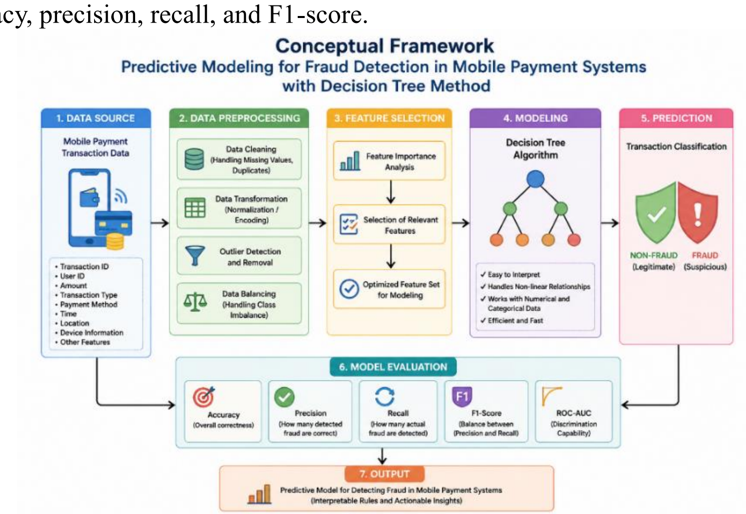

FIGURE 3. Conceptual Framework.

The conceptual framework illustrates the process of predictive modeling for fraud detection in mobile payment systems using the Decision Tree method. The process begins with collecting mobile payment transaction data (Dahlberg et al., 2015), (Han et al., 2012). The data then undergo preprocessing, which involves data cleaning, and handling class imbalance (Han et al., 2012), (Géron, 2022). Next, feature selection is performed to identify the most relevant variables (Géron, 2022), (Kuhn & Johnson, 2013). The selected features are then used to build a classification model using the Decision Tree algorithm (Breiman et al., 1984). To assess the model's performance, evaluation metrics such as accuracy, precision, recall, F1-score, and ROC-AUC are applied (Géron, 2022), (Hafez et al., 2025). Finally, the output of this process is a predictive model capable of detecting fraudulent transactions and providing actionable insights to improve security and decision-making in mobile payment systems (Hafez et al., 2025),(George et al., 2025).

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

METHODS

Research Design

This study employs a quantitative research approach using a machine learning-based predictive modeling technique to detect fraudulent transactions in mobile payment systems. The research focuses on building a classification model using the Decision Tree algorithm to distinguish between fraudulent and legitimate transactions (Géron, 2022), (Hafez et al., 2025).

Data Source

The data used in this study consist of mobile payment transaction data, including transaction ID, user ID, transaction amount, transaction type, payment method, time, location, and device information. The dataset contains labeled data indicating whether a transaction is classified as fraud or non-fraud (Han et al., 2012), (Hahsler, 2024).

Data Preprocessing

Data preprocessing is conducted. The preprocessing steps include data cleaning to handle missing values and remove duplicate records, and data balancing to address class imbalance between fraud and non-fraud transactions. This step is essential to improve model performance and accuracy (Han et al., 2012), (Kuhn & Johnson, 2013).

Feature Selection

Feature selection is performed to identify the most relevant variables that influence fraud detection. Feature selection helps model efficiency (Hahsler, 2024), (Kuhn & Johnson, 2013).

Model Development

The classification model is developed using the Decision Tree algorithm, a supervised learning method that constructs a tree-like structure based on decision rules. The dataset is divided into training and testing data to evaluate model performance. Decision Tree is chosen due to its interpretability and ability to handle both numerical and categorical data (Breiman et al., 1984), (Hahsler, 2024).

Model Evaluation

The performance of the model is evaluated using several classification metrics, including accuracy, precision, recall, F1-score, and ROC-AUC. These metrics are important to ensure that the model performs well, especially in handling imbalanced datasets where fraud cases are typically fewer than non-fraud cases (Géron, 2022), (Hafez et al., 2025).

Research Procedure

The research procedure consists of several stages, including data collection, data preprocessing, feature selection, data splitting into training and testing sets, model training using the Decision Tree algorithm, model evaluation, and interpretation of results. This structured process ensures that the predictive model is developed systematically and can be validated effectively (Han et al., 2012), (Géron, 2022).

Expected Outcome

The expected outcome of this study is a predictive model capable of accurately classifying mobile payment transactions into fraud and non-fraud categories. In addition, the model is expected to identify important features that

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

influence fraud detection and provide insights for improving transaction security in mobile payment systems (Hafez et al., 2025), (Manorom et al., 2024).

RESULTS AND DISCUSSION

Results

The model evaluation results are obtained after applying the Decision Tree algorithm to classify transactions into fraudulent and non-fraudulent categories. The performance of the model is measured using several evaluation metrics, including accuracy, precision, recall, and F1-score.
The evaluation results of the Decision Tree model :

```
Accuracy : 0.9994557775359011
Precision: 0.8941176470588236
Recall   : 0.7755102040816326
F1 Score : 0.8306010928961749
ROC-AUC  : 0.8876759658562355
```

FIGURE 4. Evaluation

The evaluation results of the Decision Tree model show a very high level of performance in classifying transactions. The model achieved an **accuracy of 99.94%**, indicating that almost all transactions were correctly classified. The **precision value of 89.41%** shows that most transactions predicted as fraudulent are indeed fraud. The **recall value of 77.55%** indicates that the model is able to detect a majority of fraudulent transactions. The **F1-score of 83.06%** reflects a good balance between precision and recall. In addition, the **ROC-AUC value of 88.77%** indicates that the model has good capability in distinguishing between fraud and non-fraud classes.
The confusion matrix results are as follows:

	Predicted Non-Fraud	Predicted Fraud
Actual Non-Fraud	56855	9
Actual Fraud	22	76

FIGURE 5. Confusion Matriks

From the confusion matrix, it can be observed that the model correctly classified a large number of non-fraud transactions (56855) and fraud transactions (76).
The Decision Tree model is shown in the following figure :

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

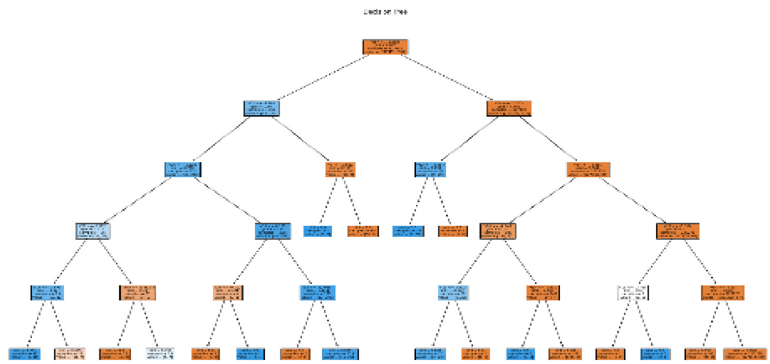

FIGURE 6. Illustrates the structure of the Decision Tree model

The figure illustrates the structure of the Decision Tree model used to classify transactions into fraudulent and non-fraudulent categories. Each internal node represents a **decision rule**, where the data are divided according to certain thresholds. The branches indicate the possible outcomes of these decisions, while the **leaf nodes** represent the final classification results. The Decision Tree structure allows the model to be easily interpreted, as each path from the root to a leaf node forms a clear decision rule for classifying transactions (Breiman et al., 1984), (Han et al., 2012).
The Confusion Matrix generated from the model is shown in Figure 7. :

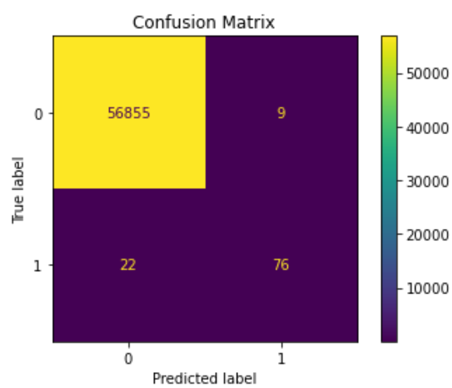

FIGURE 7. Confusion Matrix Illustrates

The confusion matrix illustrates the classification performance of the model by comparing actual and predicted values. The model correctly classified **56,855 non-fraud transactions (True Negatives)** and **76 fraud transactions (True Positives)**. However, there are **9 false positives**, and **22 false negatives**. The confusion matrix provides a detailed evaluation of classification performance beyond overall accuracy (Sokolova & Lapalme, 2009).
The Evaluation results of the model are presented in the following Figure 8 :

Proceedings of the Perbanas International Seminar on Economics, Business, Management, Accounting and IT
[June 2026] [ISSN 3047-0951]

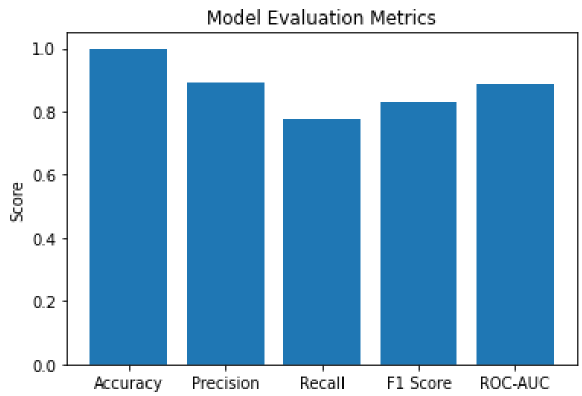

FIGURE 8. Performance Metrics of the Decision Tree model

The bar chart presents the performance metrics of the Decision Tree model, including accuracy, precision, recall, F1-score, and ROC-AUC. The model achieves a very high accuracy, indicating strong overall performance. Precision is also relatively high, meaning that most predicted fraud cases are correct. However, the recall value is comparatively lower. The F1-score shows a balance between precision and recall, while the ROC-AUC value reflects the model's ability to distinguish between fraud and non-fraud classes. These metrics provide a comprehensive evaluation of the model's effectiveness (Géron, 2022). The ROC Curve of the model is shown in the following Figure 9.

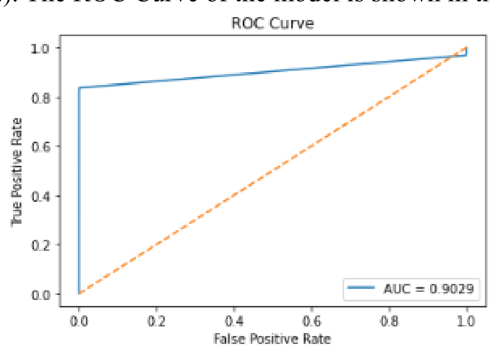

FIGURE 9. ROC Curve Illustrates

The ROC (Receiver Operating Characteristic) curve illustrates the trade-off between the True Positive Rate (TPR) and False Positive Rate (FPR). The AUC (Area Under the Curve) value of approximately 0.90 indicates that the model has a strong ability to distinguish between fraudulent and non-fraudulent transactions. A higher AUC value reflects better classification performance, making this model suitable for fraud detection tasks (Fawcett, 2006).

Discussion

The results show that the Decision Tree model performs very well in terms of overall accuracy. However, in fraud detection tasks, accuracy alone may be misleading because the dataset is highly imbalanced, with non-fraud transactions dominating the data. The high precision value indicates that the model is reliable in identifying fraud cases with minimal false alarms, while the recall value shows that some fraudulent transactions are still missed. This is reflected in the confusion matrix, where 22 fraud cases were classified as non-fraud.

Despite this limitation, the model demonstrates strong overall performance and provides interpretable decision rules through the Decision Tree structure. Important features such as transaction amount, transaction patterns, and account balance contribute significantly to the classification process. To further improve performance, especially recall, future research may apply data balancing techniques such as SMOTE or use more advanced models like Random Forest or Gradient Boosting.

CONCLUSIONS

Based on the results of data processing and model evaluation, the Decision Tree algorithm is able to effectively classify transactions in mobile payment systems into fraudulent and non-fraudulent categories. The model achieved an accuracy of **99.94%**, precision of **89.41%**, recall of **77.55%**, F1-score of **83.06%**, and a ROC-AUC value of approximately 0.90, indicating good performance in identifying fraud patterns and distinguishing between fraud and non-fraud transactions. The confusion matrix results also show that the model correctly classified **56,855 non-fraud transactions** and **76 fraud transactions**, although there were **still 22 undetected fraud cases**.

Furthermore, the preprocessing stage and feature selection process significantly contributed to improving the model's performance by identifying important variables related to fraud detection. The Decision Tree model also provides interpretable decision rules, making it suitable for practical implementation in financial systems. However, limitations such as class imbalance and the possibility of overfitting may still affect the model's ability to detect all fraudulent transactions. Therefore, future research is recommended to apply data balancing techniques or more advanced models to further improve fraud detection performance.

Future research is also recommended to compare Decision Trees with other state-of-the-art classification models or hybrid approaches to further improve fraud detection performance (Chen et al., 2025; George et al., 2025).

References

Allayannis, G. Y., Yemen, G., & Lane, D. (2018). *Fintech, payments innovation, and the acquisition of WorldPay*.

Breiman, L., Friedman, J., Stone, C., & Olshen, R. (1984). *Classification and regression trees (crc, boca raton, fl)*.

Chen, Y., Zhao, C., Xu, Y., Nie, C., & Zhang, Y. (2025). Year-over-year developments in financial fraud detection via deep learning: A systematic literature review. *ArXiv Preprint ArXiv:2502.00201*.

Dahlberg, T., Guo, J., & Ondrus, J. (2015). A critical review of mobile payment research. *Electronic Commerce Research and Applications*, *14*(5), 265–284.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874.

George, M. Z. H., Alam, M. K., & Hasan, M. T. (2025). Machine learning for fraud detection in digital banking: a systematic literature review REVIEW. *ArXiv Preprint ArXiv:2510.05167*.

Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. " O'Reilly Media, Inc."

Grudniewski, T., Wojciechowska-Filipek, S., Brążkiewicz, D., Ciekanowski, Z., & Wysokińska, A. (2025). *Contemporary threats in the financial market*.

Hafez, I. Y., Hafez, A. Y., Saleh, A., Abd El-Mageed, A. A., & Abohany, A. A. (2025). A systematic review of AI-enhanced techniques in credit card fraud detection. *Journal of Big Data*, *12*(1), 6.

Hahsler, M. (2024). An R companion for introduction to data mining. *Journal of Open Source Education*, *7*(82), 223.

Han, J., Kamber, M., & Pei, J. (2012). Data mining: Concepts and. *Techniques, Waltham: Morgan Kaufmann Publishers*, 2012–2013.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26). Springer.

Manorom, P., Detthamrong, U., & Chansanam, W. (2024). Comparative assessment of fraudulent financial transactions using the machine learning algorithms decision tree, logistic regression, naïve bayes, k-nearest neighbor, and random forest. *Engineering, Technology & Applied Science Research*, *14*(4), 15676–15680.

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.

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks.

Information Processing & Management, *45*(4), 427–437.
Taavitsainen, N. (2025). *Trust to Exploitation: The Legal and Societal Implications of Payment Fraud and the Path Toward Stronger Protections*.
Vlaicu, R. (2025). *Digital Payments Adoption by Consumers and Firms: Implications for Financial Inclusion*.
Xu, B., Wang, Y., Liao, X., & Wang, K. (2023). Efficient fraud detection using deep boosting decision trees. *Decision Support Systems*, *175*, 114037.