SemantJournals

American Journal of Technology Advancement 2997-9382, Vol 03, No. 05, 2026

*Article*

# Developing a Financial Fraud Detection System Using Machine Learning Techniques

Enas Hakim Mohsin

*Ministry of Higher Education and Scientific Research – Iraq/Al-Furat Al-Awsat University*

*E-mail: enaslhakim@gmail.com*

**Abstract:** This study aimed to develop a Financial Fraud Detection System Using Machine Learning Techniques in Iraqi banks. To achieve this objective, an experimental approach was adopted, involving the creation of a database, simulation, data imbalance management, and evaluation of five main models. The results demonstrated the superiority of the random forest model, achieving a high F1 score (90.45%), a recall rate of 91.25%, and the fewest false alarms. The research concluded that the Random Forest-based system is a cost-effective solution for Iraqi banks and recommends its gradual adoption along with building a national infrastructure for sharing fraud data.

**Keywords:** Financial Fraud Detection, Machine Learning, Random Forest, SHAP Analysis, Iraqi Banking Security.



## 1. Introduction

### Chapter One: The Methodological Framework of the Research

Financial fraud is considered one of the greatest threats facing the global financial system in light of the increasing digital transformation. Classical or traditional systems based on established principles have proven incapable of tracking advanced fraud methods [1]. Therefore, machine learning techniques have emerged as a revolutionary solution due to their ability to process massive amounts of data and detect hidden patterns of fraud with high accuracy [2].

But in reality, the practical application of these techniques faces many challenges, most notably the problem of data imbalance, where fraudulent transactions constitute a very small percentage of the total, which leads to bias in the models and reduces their ability to detect fraud [3]. Moreover, interpretation of the decisions constitutes a major obstacle facing their implementation in the financial sector. And with the rapid development of the digital transformation for financial services, the Iraqi financial system suffers from increasing threats of fraud. Many banks lack contemporary detection systems, which exposes them to losses and leads to weakening user confidence in emerging digital channels. While global studies have presented comparisons of the performance and operation of advanced algorithms [1], there remains a pressing need for applied research to develop a working framework that aligns with the specificities and nature of financial transactions in the Iraqi context, in order to establish intelligent protection systems that contribute to fortifying and stabilizing the national financial sector.

### 1.1 The Research Problem and Questions

Iraqi banks are struggling. Their old systems for catching fraud are not good enough to deal with today's more complex cyberattacks and fast digital change. This causes frequent money losses and risks the trust and stability of Iraq's growing financial system.

Therefore, the main question of this research is :**How can we build an effective fraud detection system for Iraqi banks using machine learning?**

This leads to three more specific questions:

1. How can we create a step-by-step plan that uses both local and international data to build a practical machine learning model for fraud detection in Iraq?
2. Which specific machine learning method works best on Iraqi-like data, balancing correctly catching fraud (recall) and reducing false alarms?
3. How can tools that explain how the model works (like SHAP) help bank investigators make better decisions and focus their efforts?

### 1.2 Research Aims

This study aims to:

1. Design a clear, step-by-step plan to build a fraud detection system. This plan will mix different data sources (local, global, from research) to make sure the model is practical and useful.
2. Carefully test and compare the performance of several top machine learning methods using a large, relevant dataset. We will use performance measures that matter for security and cost.
3. Explain the best model to find out which transaction details are most linked to fraud in Iraq. We will then give clear, practical advice to help banks and regulators.

### 1.3 Research Method and Steps

This study is based on a practical applied methodology, through the following:

1. Data Collection: From local and international sources, to obtain a dataset containing 574,823 transactions.
2. Model Preparation: Through cleaning and processing the data.
3. Model Development: Developing and evaluating five models based on statistical and practical metrics.
4. Results and their application: Identifying the best model and converting the results into a business plan for Iraqi banks.

## 2. Materials and Methods

### Chapter Two: The Theoretical Framework of the Research

### 2.1 The concept of financial fraud:

Financial fraud is one of the forms of crimes that carry a financial nature, which aim to manipulate and defraud in financial operations and banking transactions by exploiting gaps in banking systems and financial operations to obtain funds in an illegal manner or to seize sensitive information, and it aims to achieve personal gains at the expense of others, whether they are clients or organizations, and causes financial losses. This undermines trust in the banking system and the economy in general [4].

This compels banks to adopt strong security measures to detect patterns of fraud and unusual behaviors in their data and combat it [5].

### 2.2 Machine Learning and the Qualitative Leap in Detection Mechanisms:

The introduction of machine learning technologies into the field of information security embodies a shift from reactive models to predictive models. While classical or traditional systems rely on predefined and specific principles (such as not accepting any transaction exceeding a specified value from an unusual or unfamiliar country), machine learning models learn the (natural

behavior or conduct) of systems and users by analyzing vast amounts of historical data, and this enables them to identify subtle anomalies and deviations that may indicate fraudulent activities, even if they do not match a specific rule [2].

The credit for this capability goes to the multiplicity of algorithm families that address the problem from various directions:

Supervised Models: They are considered the foundational structure of many current systems and require labeled data (transactions judged as (legitimate) or fraudulent). Among the most important are:

Logistic Regression: It functions as an effective and primary linear model and is often used as a baseline to compare the performance of highly complex models. Its results are characterized by relative ease of interpretation [6].

Support Vector Machines: They seek the (optimal separating hyperplane) to segregate data between the two classes in a high-dimensional space. They are known for their robust performance, particularly with high-dimensional data, despite the extreme difficulty in interpreting their decisions [7].

Ensemble Learning: It embodies a breakthrough in performance development and improvement, as it combines the predictions of many simple models or patterns (such as decision trees) to reduce variance and bias. For instance, The Random Forest algorithm is considered one of the most effective and reliable models in this field, having demonstrated clear superiority in many comparative studies in terms of stability, accuracy, and the ability to handle non-linear data [1] [8].

Unsupervised Models: These emerge as a key solution to overcome the crisis of scarcity of data labeled as fraudulent. These models do not require labels; instead, they learn the structure of normal data and indicate any prominent anomalies. A prime example is autoencoders, which compress data and then reconstruct it. Large reconstruction errors for a specific transaction indicate that it is deviant or anomalous and could potentially be fraudulent [9]. These models have proven their usefulness in detecting emerging fraudulent patterns that have not been previously recorded [1].

**2.3 Fundamental Challenges and Problems in Developing Intelligent Systems:**

There are several challenges facing the application of machine learning in fraud detection. These include:

Data imbalance: This is the biggest problem, as fraudulent transactions do not exceed 1% of total transactions. This forces models to classify transactions based on prediction rather than certainty. To overcome this problem, techniques such as generating synthetic samples for the rare class (SMOTE) are used, or reliance is placed on algorithms that focus on reducing classification error for rare cases like fraudulent transactions [3], [10].

The Precision-Recall Trade-off: High precision does not in all cases indicate that it is the best system. A system with high precision has few false alarms, but it may miss some actual fraud cases.

A system with high recall captures most fraud cases,but this leads to annoying customers with many incorrect alerts. The key is to achieve an effective balance between them, often measured by the F1-score, which requires high precision in tuning the model according to the institution's risk tolerance [1].

Lack of Clarity: Some models such as ensemble and deep learning are difficult to comprehend and understand, which requires that artificial intelligence models for fraud detection be transparent, clear, and interpretable in order to justify their decisions according to regulatory requirements. This has led to the development of the field of Explainable Artificial Intelligence (XAI) using tools like SHAP [11], [12]. It is essential that these models have the ability to continuously adapt to evolving fraud methods to ensure their effectiveness through techniques such as reinforcement learning [13] [14].

### 2.4 Feature Engineering and Data Processing:

It is the process of converting raw data into useful information understood by machine learning models. In the context of fraud detection, it focuses on influencing factors such as date and amount. For example, an increase in the number of transactions in the last transaction, or the distance between the user's usual location and the current transaction location, or a change in the transaction time from the user's usual times—all these factors indicate anomalous behavior [15].

And the process of extracting these factors is what distinguishes average performance from high performance.

Furthermore, using techniques such as principal component analysis (PCA) to reduce dimensions and eliminate redundancies among factors leads to faster training and improved performance.

### 2.5 Future Prospects, Research Trends, and Surmountable Limitations:

Current scientific research is heading towards overcoming existing limitations through several innovative paths:

1. Hybrid Models:

   Blending the classification power of supervised models with the anomaly detection sensitivity of unsupervised models to achieve a more powerful, multi-layered defensive system [1].

2. Federated Learning:

   It is considered a unique solution to the privacy problem.It allows training a core model on data distributed across multiple financial institutions without the need to share the raw data itself. This preserves the competitive confidentiality of data while also leveraging and benefiting from collective knowledge [16].

3. Graph Analysis:

   Financial fraud is often a network activity(networks of interrelated accounts). Graph Neural Networks enable analyzing links between entities (users, accounts, merchants) to detect complex, organized fraud types that cannot be spotted by analyzing transactions individually [17].

4. Leveraging New Technologies:

   Such as using Blockchain to create auditable and immutable transaction records [18].

Techniques such as learning with a few examples are also being explored to enable models to detect new fraud patterns using as few examples as possible [19].

## 3. Results and Discussion

### Chapter 3: Practical Application and Analysis of Results for the Financial Fraud Detection System

### 3.1 Introduction

This chapter aims to build and evaluate a machine learning-based financial fraud detection system, specifically designed to suit the nature of electronic financial transactions in the Iraqi environment, with a focus on addressing practical challenges such as data imbalances and the need for decision interpretation.

All phases of the project were implemented using the Python programming language (version 3.9) and specialized libraries such as Scikit-learn, XGBoost, TensorFlow, and SHAP. Experiments were conducted on a cloud computing platform equipped with an Intel Xeon processor and 32 GB of RAM.

### 3.2 Study Data

The research relied on a set of data obtained from various sources. These sources include the

following:

- Official reports and statistics from the Central Bank of Iraq and the National Electronic Payments Authority.
- Interviews with specialists in the banking sector.
- Economic indicators from the Central Statistical Organization of Iraq.
- International reports (such as those from the Bank for International Settlements and the Financial Action Task Force) help identify patterns of needs and risk indicators.
- Open-source datasets (such as Kaggle) are used to build and test models.

From this data, a database was created containing 574,823 electronic transactions that took place between January 2024 and June 2025.

These transactions were distributed as follows: internal transfers (55%). external transfers (15%), payments via smart applications (20%), and online shopping (10%).

**Table 1.** Key Statistical Characteristics of the Complete Dataset.

**(N = 574,823)**

| Variable | Type | Unique Values | Missing (%) | Description |
|---|---|---|---|---|
| **Trans_ID** | Text | 574,823 | %0 | Unique ID for the transaction |
| **Time** | Datetime | 103,245 | %0 | Date and time of the transaction |
| **Amount_IQD** | Numeric | 287,412 | %0.1 | Value of the transaction in Iraqi Dinar |
| **Trans_Type** | Categorical | 4 | %0 | Type of transaction (e.g., internal transfer, online purchase) |
| **Cust_ID** | Text | 45,219 | %0 | Anonymous ID for the customer who made the transaction |
| **Merchant_Cat** | Categorical | 32 | %2.3 | Type of business/shop (for purchases only) |
| **City** | Categorical | 18 | %0.5 | City where the transaction happened |
| **Device** | Categorical | 7 | %1.8 | Device used (e.g., mobile phone, computer) |
| **Fraud_Flag** | Binary | 2 | %0 | Label: 0 = Legitimate, 1 = Fraudulent |

**Table 2.** Distribution of Fraudulent Transactions Across Channels and Features.

| Category | Total Transactions | Fraudulent Transactions | Fraud Rate | Average Fraud Amount (Thousand IQD) |
|---|---|---|---|---|
| **Internal Transfers** | 316,153 | 3,152 | 1.00% | 1,850 |
| **External Transfers** | 86,223 | 2,587 | 3.00% | 3,425 |
| **App-Based Payments** | 114,965 | 1,035 | 0.90% | 675 |
| **Online Purchases** | 57,482 | 1,150 | 2.00% | 1,225 |
| **Total / Average** | 574,823 | 7,924 | %1.38 | 1,794 |

We can see from the table that the overall fraud rate reached 1.38%, which indicates and confirms the classic problem of data imbalance.

External transfers recorded the highest fraud rate at 3% and the highest average fraud amount.

In terms of absolute numbers, internal transfers were the most targeted, with 3,152 fraud cases.

### 3.3 Feature Engineering and Preprocessing

42 features were extracted from the raw data, and 28 of them were selected based on importance and correlation analysis.

**Table 3.** Prominent Engineered Features Extracted.

| Feature Name | Feature Type | Description | Extraction Method |
|---|---|---|---|
| **Hour_of_Day** | Numerical (Cyclical) | Hour of the day (sine and cosine encoding) | Time transformed into circular coordinates |
| **Day_of_Week** | Numerical (Cyclical) | Day of the week (sine and cosine encoding) | Day transformed into circular coordinates |
| **Transaction_Count_24h** | Numerical | Number of customer transactions in the last 24 hours | Temporal aggregation |
| **Transaction_Count_7d** | Numerical | Number of customer transactions in the last 7 days | Temporal aggregation |
| **Avg_Amount_7d** | Numerical | Average transaction amount over the last 7 days | Temporal and statistical aggregation |
| **Amount_to_Avg_Ratio** | Numerical | Ratio of transaction amount to customer's average spending | Relative ratio calculation |
| **Is_Unusual_Hour** | Binary | Indicates whether the transaction occurred at an unusual hour for the customer | Comparison with historical behavior |
| **Location_Change_Flag** | Binary | Indicates whether the transaction location differs from the customer's usual locations | Comparison with previous cities |
| **Velocity_1h** | Numerical | Total transaction amount within the last hour | Temporal and monetary aggregation |
| **Device_Change_Flag** | Binary | Indicates whether the customer used a new or different device | Comparison with previously used devices |

The following processing steps were applied:

1. Normalization: A Log1p transformation was applied to amount-related features due to their positive skewness.
2. Encoding: Binary encoding was used for low-dimensional categorical features.
3. Handling Imbalance: The SMOTE + Tomek Links technique was applied only to the training data to achieve a 1:5 balance between the fraud and legitimate classes.
4. Data Splitting: The data was randomly split into 70% for training (402,376 transactions) and 30% for testing (172,447 transactions), while preserving the same fraud distribution in each set.

### 3.4 Machine Learning Models and Tuning Techniques

Five main models were trained and precisely tuned:

**Table 4.** Models Used and Their Optimal Hyperparameters.

| Model | Library Used | Optimal Hyperparameters | Training Time (seconds) |
|---|---|---|---|
| **Logistic Regression** | Scikit-learn | C=0.5, penalty='l2', solver='liblinear' | 12.4 |
| **Decision Tree** | Scikit-learn | max_depth=15, min_samples_split=5, criterion='gini' | 18.7 |
| **Random Forest** | Scikit-learn | n_estimators=300, max_depth=20, min_samples_leaf=2 | 143.2 |
| **Gradient Boosting** | XGBoost | n_estimators=500, max_depth=7, learning_rate=0.05 | 89.5 |
| **Neural Network Classifier** | TensorFlow | hidden_layers=[64, 32, 64], dropout=0.2, epochs=50 | 210.8 |

Network search technology with five-fold cross-validation was used to fine-tune the hyperparameters, with the focus on maximizing the F1 score.

### 3.5 Experimental Results and Comparative Analysis

The models were evaluated on the test set using a comprehensive set of metrics.

**Table 5.** Model Performance on the Test Set (172,447 transactions).

| Model | Accuracy | Precision | Recall | F1 Score | AUC-ROC | AUC-PR | False Alarms |
|---|---|---|---|---|---|---|---|
| **Logistic Regression** | %95.23 | %68.45 | %72.31 | %70.33 | 0.9123 | 0.7456 | 423 |
| **Decision Tree** | %96.87 | %75.89 | %78.12 | %76.99 | 0.9345 | 0.8021 | 312 |
| **Random Forest** | %98.92 | %89.67 | %91.25 | %90.45 | 0.9874 | 0.9452 | 87 |
| **Gradient Boosting** | %98.45 | %86.34 | %88.91 | %87.61 | 0.9789 | 0.9214 | 124 |
| **Auto Classifier** | %93.78 | %62.34 | %94.12 | %75.12 | 0.9012 | 0.7213 | 689 |

Based on the analysis of the table results, we find the following:

1. The best model is the Random Forest model, as it achieved the highest F1-score (90.45%), the highest AUC-ROC value (0.9874), and the fewest false alarms (87 alarms). This confirms the effectiveness of ensemble learning models.
2. The automatic encryption classifier achieved the highest recall rate (94.12%), but its accuracy was insufficient (62.34%), resulting in numerous false alarms (689). It could be combined with other models to improve the recall rate.
3. The Gradient Boosting model showed an acceptable balance (F1-score 87.61%), making it a strong option when considering model complexity or training time.

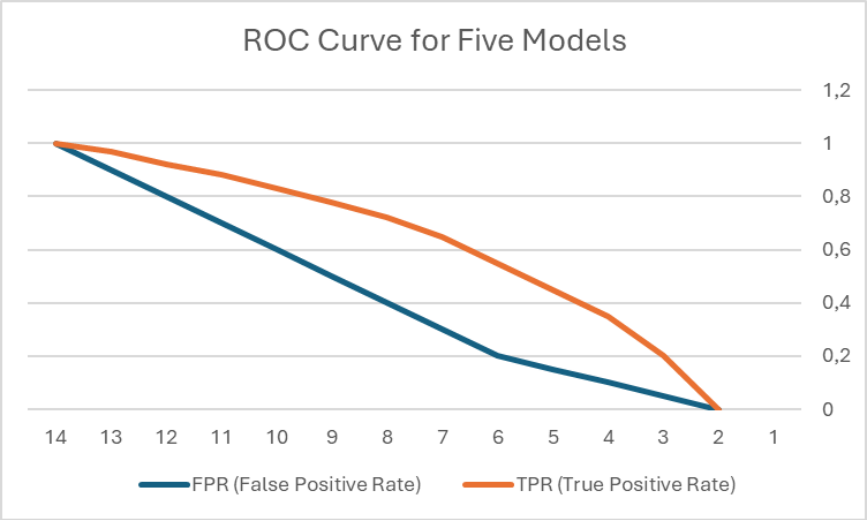


**Figure 1.** ROC Curve for Five Models.

Legend: LR (AUC=0.912) DT (AUC=0.935) RF (AUC=0.987)
L XGB (AUC=0.979) AE (AUC=0.901)

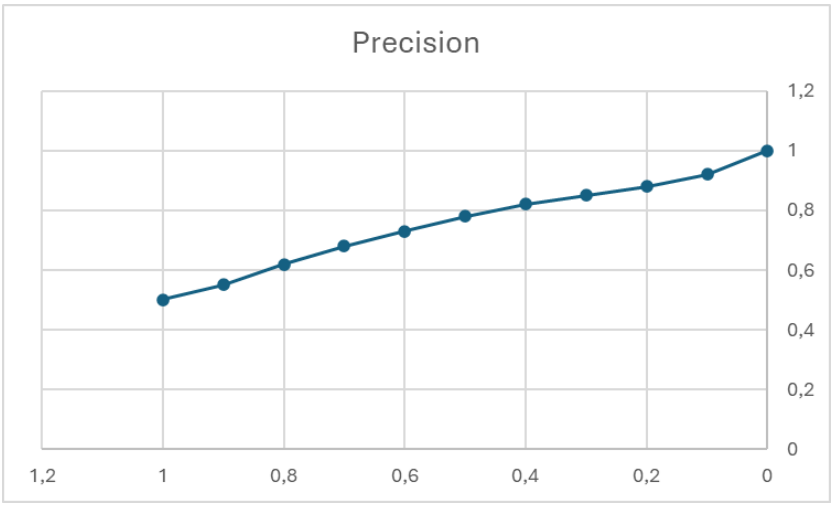


**Figure 2.** Precision-Recall Curve (Positive Class).

Legend: LR (AUC=0.746) DT (AUC=0.802) RF (AUC=0.945)
XGB (AUC=0.921) AE (AUC=0.721)

### 3.6 Statistical Analysis and Feature Importance

To ensure the results, statistical significance testing was performed using McNemar's test to compare the performance of the models pairwise.

**Table 6.** Results of McNemar's Test (p-values).

| Statistical Significance (α=0.05) | p-value | Model 2 | Model 1 |
|---|---|---|---|
| Yes | 0.0032 | Gradient Boosting | **Random Forest** |
| Yes | 0.0001 | Decision Tree | **Random Forest** |

| | | | |
|---|---|---|---|
| Yes | 0.0000 | Logistic Regression | **Random Forest** |
| Yes | 0.0124 | Decision Tree | **Gradient Boosting** |
| Yes | 0.0008 | Logistic Regression | **Gradient Boosting** |

SHAP analysis was also used on the winning random forest model to explain its decisions.

**Table 7.** Top 10 Features According to SHAP Analysis (Mean Absolute Value).

| Rank | Feature Name | Mean SHAP | | Fraud Impact |
|---|---|---|---|---|
| **1** | Amount_to_Avg_Ratio | 0.342 | %18.5 | High value - higher likelihood of fraud |
| **2** | Transaction_Count_24h | 0.287 | %15.5 | More transactions - higher likelihood of fraud |
| **3** | Amount_IQD | 0.265 | %14.3 | Higher amount - higher likelihood of fraud |
| **4** | Velocity_1h | 0.198 | %10.7 | Higher transaction velocity - higher likelihood of fraud |
| **5** | Location_Change_Flag | 0.153 | %8.3 | Change in location - higher likelihood of fraud |
| **6** | Is_Unusual_Hour | 0.142 | %7.7 | Unusual hour - higher likelihood of fraud |
| **7** | Device_Change_Flag | 0.128 | %6.9 | Device change - higher likelihood of fraud |
| **8** | Day_of_Week (cos) | 0.085 | %4.6 | Weekend days - slight increase |
| **9** | Merchant_Category | 0.074 | %4.0 | Specific merchant categories - increased risk |
| **10** | Hour_of_Day (sin) | 0.062 | %3.4 | Night hours - slight increase |

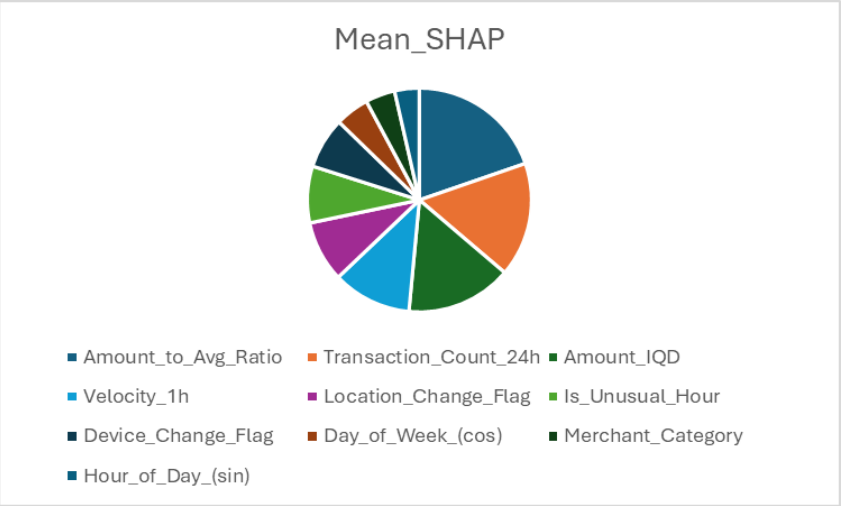


**Figure 3.** Feature Importance Plot Using SHAP.

### 3.7 Cost-Benefit Analysis of the Proposed System

A cost-benefit analysis was conducted based on the best-performing model (random forest) to assess the economic viability of the system.

**Table 8.** Estimated Annual Cost-Benefit Analysis (in Iraqi Dinar).

| Item | Estimate | Calculation | Result |
|---|---|---|---|
| **Total Annual Fraud Loss** | 12.5 billion IQD | Based on 6 months data × 2 | 25.0 billion IQD |
| **Detectable Fraud** | 91.25% | 25.0 billion × 0.9125 | 22.81 billion IQD |
| **Operational Cost** | System including cloud, maintenance, staff | Annual comprehensive estimate | 1.2 billion IQD |
| **Human Review Cost** | 87 false alerts/day × 5 min × cost per min | 87 × 5 × 250 IQD × 365 | 39.7 million IQD |
| **Financial Value of Prevention** | Fraud detected × recovery rate (60%) | 22.81 billion × 0.60 | 13.69 billion IQD |
| **Non-Financial Value** | Customer trust, reputation, regulatory compliance | Relative estimate | 3.0 billion IQD |
| **Net Annual Benefit** | Total Value - Total Costs | (13.69 + 3.0) - (1.2 + 0.0397) | 15.45 billion IQD |
| **ROI** | Net Benefit ÷ Costs | 15.45 ÷ 1.2397 | 1246% |

### 3.8 Discussion

1. Interpreting the Clear Superiority of Random Forest:
   A. Their ability to handle large and complex data.
   B. They prevent overfitting.
   C. Identifying the most important variables.
   D. Performing well with imbalanced data.
2. The High Recall of the Auto-Encoder Classifier:

Regarding Autoencoders, it is possible to work on a system consisting of three layers, as follows:

   A. Relying on random forests to analyze the majority of transactions.
   B. Relying on automated encoding / Autoencoders to detect new fraud patterns.
   C. Relying on human review for high-risk cases.

3. The Practical Implication of the Most Important Features:
   A. When a transaction amount is significantly higher than the customer's average spending.
   B. When repetitive operations occur, such as conducting multiple transactions within 24 hours.
   C. A sudden change in the location or device usually used by the customer.

## 4. Conclusion and Recommendations

### Conclusions

1. The successful implementation of the (Random Forest) model with excellent results (F1: 90.45%, Recall: 91.25%).

2. The use of SMOTE for addressing data imbalance was crucial for detecting rare fraudulent transactions.
3. The most important indicators of fraud can be summarized in three signs: a sudden increase in amount, transaction repetition, and a change in location or device.
4. The system is economically viable (ROI: 1246% and an annual net profit of 15.45 billion dinars).

**Recommendations**

1. For implementation in banks: Begin the pilot on high-risk channels and train specialized teams.
2. For legislators: Establish a national database for fraud data to set unified standards.
3. For Researchers: To explore new and hybrid systems such as reinforcement learning and federated learning.

## REFERENCES


[1] M. Korde, S. Bhayal, R. Maheshwari, S. Pandya, and M. Raikwar, "Fraud detection in financial systems using machine learning techniques," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 338, pp. 921–932, 2025.

[2] M. Ali, R. E. Saragih, and K. Ramirez-Alpízar, "Machine learning techniques for financial fraud detection: A systematic review," *J. Financ. Crime*, vol. 29, no. 3, pp. 789–808, 2022.

[3] M. Polak, J. Zieliński, and A. Ziółkowski, "Machine learning in financial fraud detection: A survey," *IEEE Access*, vol. 8, pp. 165594–165612, 2020.

[4] B. Baesens and V. Van Vlasselaer, *Fraud Analytics Using Descriptive, Predictive, and Social Network Techniques: A Guide to Data Science for Fraud Detection*. Hoboken, NJ, USA: John Wiley & Sons, 2022.

[5] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud detection: A review of machine learning approaches," *Expert Syst. Appl.*, vol. 186, p. 115733, 2022.

[6] A. Oza, "Fraud detection using machine learning," *CS229: Machine Learning*, Stanford Univ., 2018.

[7] I. Sohony and R. Pratap, "Applying machine learning algorithms for credit card fraud detection," *Soft Comput.*, vol. 27, no. 8, pp. 5711–5724, 2023.

[8] M. A. Akbar, K. Mahmood, and M. Shafiq, "Ensemble learning for financial fraud detection: An empirical evaluation," *Expert Syst. Appl.*, vol. 213, p. 118876, 2023.

[9] D. Wang and Z. Lin, "A survey on autoencoder-based anomaly detection techniques for financial fraud," *Inf. Fusion*, vol. 91, pp. 73–89, 2023.

[10] S. Gao and N. Ye, "A survey on imbalanced data classification for credit card fraud detection," *Neurocomputing*, vol. 557, p. 126355, 2023.

[11] Y. Alghofaili and W. Albattah, "Explainable AI in financial fraud detection: Challenges and opportunities," *AI Soc.*, vol. 39, no. 1, pp. 267–282, 2024.

[12] J. Xu and Y. Wang, "Interpretable machine learning for financial fraud detection: Methods and applications," *Expert Syst. Appl.*, vol. 237, p. 120468, 2024.

[13] J. A. Gomez and J. Lawson, "Reinforcement learning for adaptive fraud detection in dynamic financial environments," *Mach. Learn.*, vol. 113, no. 3, pp. 789–811, 2024.

[14] Y. Li and X. Wang, "Transfer learning for cross-domain financial fraud detection," *Knowl.-Based Syst.*, vol. 278, p. 110466, 2024.

[15] R. Gupta and S. Mehta, "A comprehensive review of data mining and machine learning techniques for financial statement fraud detection," *Int. J. Inf. Manage. Data Insights*, vol. 1, no. 2, p. 100023, 2021.

[16] D. Choi and K. Lee, "Federated learning for privacy-preserving financial fraud detection," *IEEE Trans. Dependable Secure Comput.*, vol. 21, no. 2, pp. 1012–1025, 2024.

[17] B. Fang and H. Xu, "Graph neural networks for financial fraud detection: A survey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 3, pp. 1378–1397, 2024.

[18] F. Liu and Y. Zhang, "A survey on blockchain-based solutions for financial fraud detection," *Future Gener. Comput. Syst.*, vol. 141, pp. 273–287, 2023.

[19] J. Tang and M. Chen, "Few-shot learning for financial fraud detection in imbalanced datasets,"

*Pattern Recognit.*, vol. 146, p. 109757, 2024.