IOCRF 2025 Conference

The 1st International Online Conference on Risk and Financial Management

17–18 June 2025 | Online

MDPI

Detection fraudulent transactions using artificial intelligence algorithms

Rachid BENKHELOUF, Dr. (Econ.)1, Fatima Zohra BEKADDOUR Dr. (Acct.) 2
Department of economics, Maghnia, University Centre of maghnia 1
Department of Accounting and Finance, University of Ghardaia 2

INTRODUCTION

Fraudulent transactions pose a significant challenge to the financial sector, threatening both the economic stability of institutions and the trust of individual consumers. The rapid advancement of technology, particularly in artificial intelligence (AI) and machine learning, has opened new avenues for detecting fraudulent activities effectively.

Detecting fraud in real-time is critical for maintaining the integrity of financial systems, especially in online banking and e-commerce platforms. AI algorithms, such as the K-Nearest Neighbors (KNN), Linear Discriminant Analysis (LDA), and various regression-based methods, have emerged as robust tools that can improve the precision and recall of fraud detection efforts, helping organizations quickly respond to potential threats. As financial transactions become increasingly digital, the datasets used for training,

fraud detection models also undergo significant changes, often leading to imbalanced datasets characterized by a minor proportion of fraudulent activities compared to legitimate transactions.

This imbalance can adversely affect the model's predictive performance, demanding innovative solutions to create effective detection systems,

METHOD

We collected the dataset called **AUDIT_DATA** from Kaggle Depository for fraud transaction detection. The dataset consists of **10000** records with 08 features.

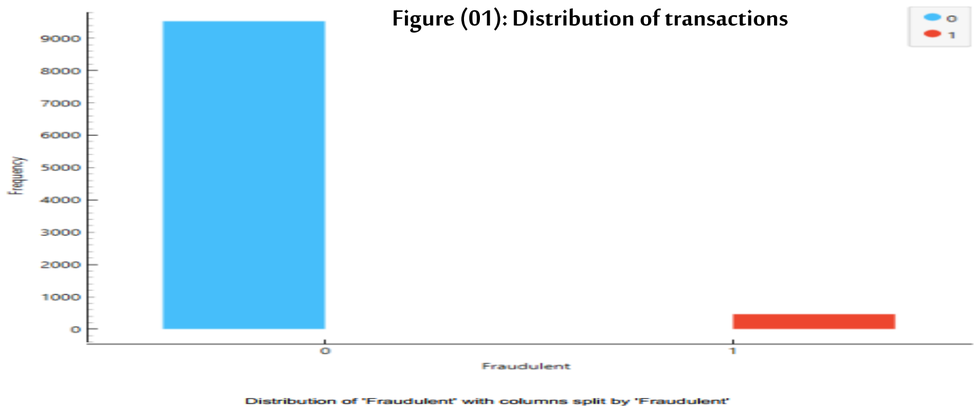
Data Collection: Transaction data was collected from the company's database, including information such as transaction IDs, amounts, transaction types, merchant details, transaction locations, timestamps, account types, and fraud indicators.

The features of the fraudulent transaction dataset are shown in **Table N° 1**

-The dataset exhibits a significant class imbalance, with fraudulent transactions making up just 4.69% of all cases, or 469 transactions, while non-fraudulent transactions accounted for 9,531.

Figure (01): Distribution of transactions

Transactions Feature Name	Description
Transaction ID	Unique identifier for each transaction.
Amount	The monetary value of the transaction
Type	Type of transaction, either "Credit" or "Debit".
Merchant	The merchant involved in the transaction, such as Amazon, Walmart, or Target.
Location	Location of the transaction, categorized as "Local" or "International".
Time	The hour of the day when the transaction occurred
Day_of_Week	The day of the week when the transaction occurred.
Account_Type	Type of account associated with the transaction, either "Personal" or "Business".
Fraudulent	Binary indicator denoting whether the transaction is fraudulent (1) or not (0).



RESULTS & DISCUSSION

The Use of Algorithm in Orange Software

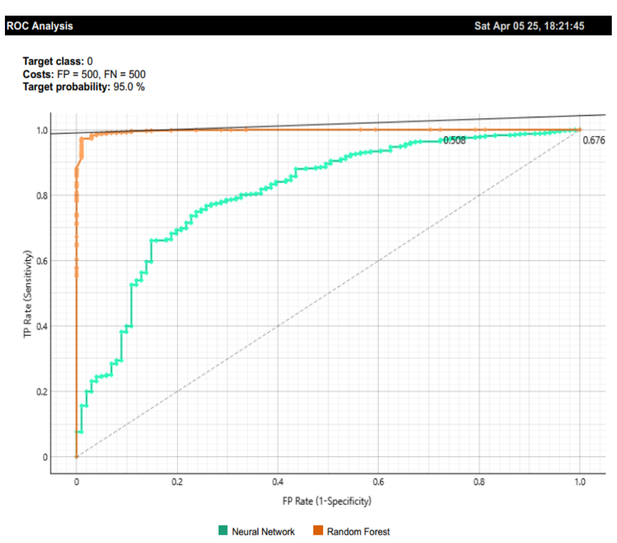
Mechanism for Applying the Artificial Neural Network Algorithm:

After processing the data and imputing missing values, the variables were encoded as a fundamental step. The data was then **split into 80% training data and 20% testing data.** The artificial neural network was designed and trained to classify financial transactions as a set of independent variables. The **Adam** Solver function was used, demonstrating good performance. Additionally, the **ReLU** activation function was adopted, as it outperformed the Tanh and Logistic functions in classifying the types of errors in financial data. The number of neurons in the hidden layers was set to **50** neurons to train the ANN.

the configuration parameters for a Random Forest model:

Number of trees: **10** , Maximal number of considered features: **unlimited** Replicable training: No Maximal tree depth: **unlimited** ,Stop splitting nodes with maximum instances: **5**

Model	AUC	CA	F1	Prec	Recall
Artificial Neural Network	0.808	0.950	0.925	0.902	0.950
Random Forest	0.997	0.964	0.954	0.965	0.964



Confusion Matrix

Confusion matrix for Random Forest

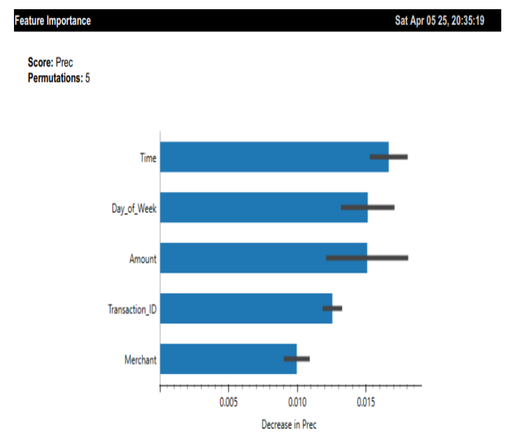
		Predicted		
		0	**1**	**∑**
Actual	**0**	1899	0	**1 899**
	1	73	28	**101**
	∑	**1 972**	**28**	**2 000**

-The accuracy of the artificial neural network (ANN) model reached **90**%, while the area under the ROC curve (AUC) was **0.808.**

-The accuracy of the Random Forest (RF) model reached **96**%, while the area under the ROC curve (AUC) was **0.997.**

-The neural network model correctly predicted 28 True Positives (TP) and 1899 True Negatives (TN), with no False Positives (FP) and 73 False Negatives (FN), indicating high accuracy in identifying error-free financial data but some misclassification of minor errors.

-Time and Amount are particularly important for reducing false positives in fraud detection.



CONCLUSION

This study has demonstrated the effectiveness of machine learning algorithms in fraud detection, with Random Forest outperforming Artificial Neural Network models. The Random Forest achieved 96% accuracy and an exceptional AUC of 0.997, indicating outstanding discriminatory power according to Hosmer & Lemeshow's guidelines. Time and transaction amount emerged as the most critical features for fraud detection. Despite the challenge of class imbalance (only 4.69% fraudulent transactions), the models successfully classified transactions with high precision. The results suggest that Random Forest models provide more reliable probability estimates for fraud detection systems, making them particularly valuable for financial security implementations.

REFERENCES

1/Richardson, A., van Florenstein Mulder, T., & Vehbi, T. (2021). Nowcasting GDP using machinelearning algorithms: A real-time assessment. *International Journal of Forecasting, 37*(2), 941–948

2/ Walton, N., & Nayak, B. S. (2021). Rethinking of Marxist perspectives on big data, artificial intelligence (AI) and capitalist economic development. *Technological Forecasting and Social Change, 166*, 120576

3/ Xiao, S., Chai, H., Wang, Q., Shao, K., Meng, L., Wang, R., Li, B., & Ma, Y. (2021). Estimating economic benefit of sugar beet based on three-dimensional computer vision: A case study in Inner Mongolia, China. *European Journal of Agronomy, 130*, 126378