International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 21 No. 3 (1) 2025

ijerst.editor@gmail.com

editor@ijerst.com

ISSN 2319-5991 www.ijerst.com
Vol. 21, No. 3(1), 2025

Research Paper

Enhancing Fraud Detection In E-Commerce Transactions In Multi Perspective User Behaviour Analysis And Process Mining

J. Prashanthi[1], A. Pallavi[2], N. Ramya[2], B. Shekar Reddy[2], K. Hemanth[2]
[1]Assistant Professor, [2]UG Student, [1,2]Department of Computer Science and Engineering (Data Science)
[1,2]Sree Dattha Group of Institutions, Sheriguda, Ibrahimpatnam, 501510, Telangana.

ABSTRACT

The increasing volume of digital transactions and the growing sophistication of fraudsters have made real-time financial fraud detection a critical concern for financial institutions. Traditional systems, which relied heavily on rule-based approaches and manual oversight, often failed to adapt to emerging fraud patterns. These limitations resulted in high false-positive rates and delayed responses. Earlier detection methods, such as statistical models and threshold-based systems, proved inadequate in identifying complex and evolving fraudulent behaviors. With the advent of machine learning, however, fraud detection has advanced significantly. AI-powered systems can now learn from historical transaction data and uncover subtle fraud indicators with improved accuracy. The primary motivation behind developing AI-based solutions is the urgent need for automated, real-time fraud detection systems capable of rapidly identifying threats and mitigating risks. These systems help reduce human error and financial losses while addressing the limitations of traditional methods—particularly their poor adaptability, high false-positive rates, and scalability challenges. The proposed AI-driven approach utilizes machine learning models such as Support Vector Machines (SVM) and Decision Trees to analyze real-time transaction data. This enables quicker and more accurate fraud detection. By processing data instantly, the system not only enhances detection speed and reduces financial loss but also provides a scalable and efficient solution for combating fraud in today's fast-paced digital landscape.

Received: 14-7-2025 Accepted: 21-8-2025 Published: 28-8-2025

1. INTRODUCTION

In today's rapidly evolving digital economy, the surge in online financial transactions has introduced significant convenience but also heightened the risk of fraudulent activities, especially in banking and e-commerce sectors. This project addresses the critical need for intelligent, real-time fraud detection by developing a web-based system using the Django framework, integrated with machine learning algorithms such as Support Vector Machine (SVM) and Random Forest.

The system enables users to upload transaction datasets, preprocess data, train models, and evaluate their performance using metrics like accuracy, precision, recall, and F1-score. It also incorporates interactive data visualizations and process mining techniques to enhance transparency and traceability of fraud patterns.Motivated by the increasing sophistication of fraud tactics and limitations of traditional rule-based systems, this project aims to create a scalable and adaptive solution that can be deployed across various domains including banking, e-commerce, insurance, and government sectors, ultimately reducing financial losses and building greater trust in digital financial systems.

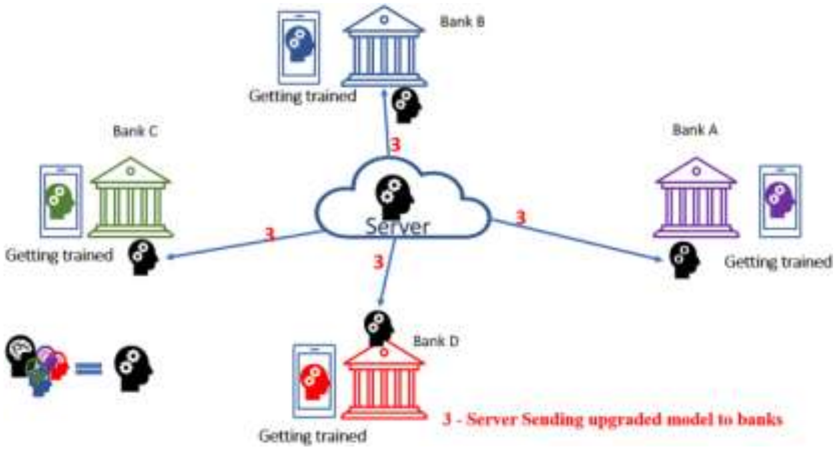

Figure 1: Application

2. LITERATURE SURVEY

Financial fraud is the act of gaining financial benefits by using illegal and fraudulent methods [1,2]. Financial fraud can be committed in different areas, such as insurance, banking, taxation, and corporate sectors [3]. Recently, financial transaction fraud [4], money laundering, and other types of financial fraud [5] have become an increasing challenge among companies and industries [4]. Despite several efforts to reduce financial fraudulent activities, its persistence affects the economy and society adversely, as large amounts of money are lost to fraud every day [5]. Several fraud detection approaches were introduced many years ago [1]. Most traditional methods are manual, and this is not only time consuming, costly, and imprecise but also impractical [6]. More studies are conducted to reduce losses resulting from fraudulent activities, but they are not efficient [7]. With the advancement of the artificial intelligence (AI) approach, machine learning and data mining have been utilized to detect fraudulent activities in the financial sector [8,9]. Both unsupervised and supervised methods were employed to predict fraud activities [4,10]. Classification methods have been the most popular method for detecting financial fraudulent transactions. In this scenario, the first stage of model training uses a dataset with class labels and feature vectors. The trained model is then used to classify test samples in the next step [1,2,5].

Thus, the study attempts to identify machine-learning-based techniques employed for financial transaction fraud and to analyse gaps to discover research trends in this area. Recently, some reviews have been conducted to detect fraudulent financial activities. For instance, Delamaire et al. [11] conducted a review on different categories of fraudulent activities on credit cards, which include bankruptcy and counterfeit frauds, and suggested proper approaches to address them. Similarly, Zhang and Zhou [12] investigated ML methods for fraud transactions, which include the stock market and other fraud detection processes in financial sectors. Raj and Portia. [13] explored several ML approaches used for credit card fraud detection. Phua et al. [14] conducted a comprehensive survey to explore data mining and machine learning techniques to detect frauds in various aspects, including credit card fraud, insurance fraud, and telecoms subscription fraud.

Recently, there has been a significant increase in fraud activities in health sectors [15]. Abdallah et al. [16] introduced a review to investigate different approaches for uncovering fraudulent activities in the health care domain based on statistical approaches. Popat and Chaudhary [17] presented an extensive review work on credit card fraud detection. The authors provide a detailed analysis of various ML classification methods with their methodology and challenges. Ryman-Tubb et al. [6] reviewed several state-of-the-art methods for detecting payment card fraudulent activities using transactional volumes. The study showed that only eight approaches have a practical implication to be used in the industry. A study by Albashrawi and Lowell [3] analyzed several studies for one decade covering fraud detection in financial sectors using data mining techniques. However, this was not exhaustive

IJERST International Journal of Engineering Research and Science & Technology

and comprehensive enough as they ignored the method of evaluations and the pros and cons of data mining techniques, among others.

3. PROPOSED SYSTEM

The proposed fraud detection system utilizes Support Vector Machines (SVM) to identify fraudulent financial transactions by analyzing historical data and learning patterns indicative of suspicious activity. The process begins with a comprehensive dataset, such as the Credit Card Fraud Detection dataset from Kaggle, which includes transaction details like amount, time, and merchant information. Preprocessing is a critical step involving the removal of null values, label encoding of categorical variables, and standardization of features to ensure uniformity. The SVM algorithm is then applied to classify transactions by finding the optimal hyperplane that separates fraudulent and legitimate cases, effectively handling complex, non-linear relationships. Following model training, performance is evaluated using metrics such as accuracy, precision, recall, and F1-score, and compared with other algorithms like Decision Trees, Logistic Regression, and Neural Networks. This comparative analysis aids in selecting the most suitable model for deployment. The dataset is typically split in an 80-20 ratio for training and testing, ensuring the model's ability to generalize to new, unseen transactions.

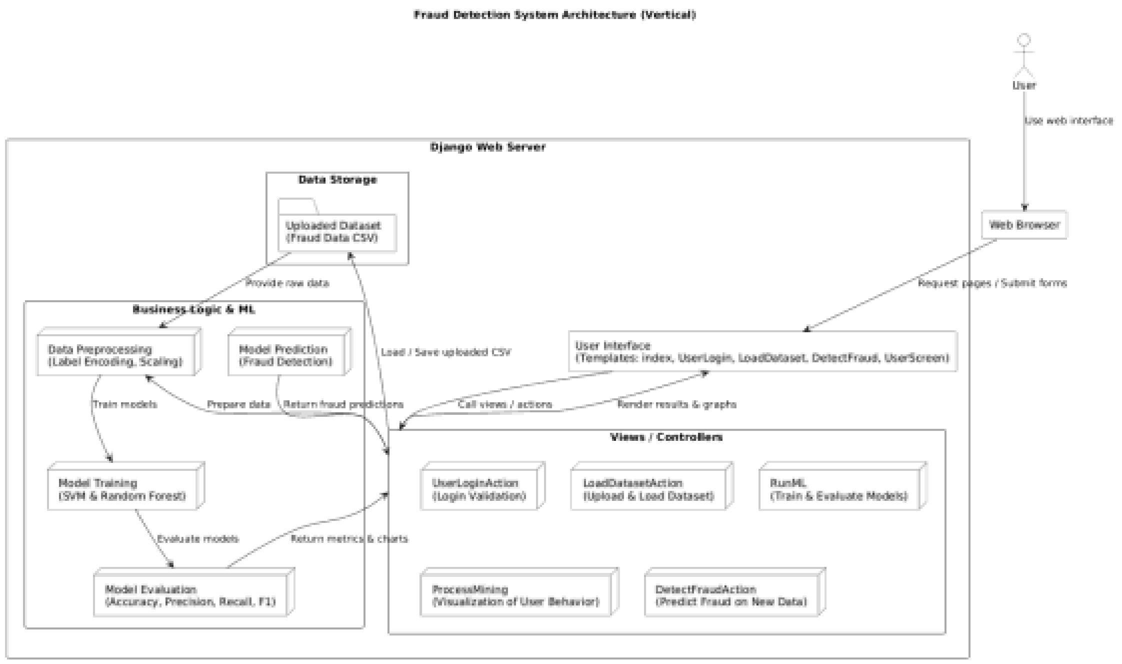

Figure 2: Proposed Block Diagram

3.1 Random Forest Classifier

The fraud detection system leverages a Random Forest Classifier trained on a preprocessed dataset of historical e-commerce transactions, where each transaction is represented by features such as transaction amount, device type, location, time, product category, and account history. The input matrix X_train undergoes label encoding, normalization, and missing value handling to ensure consistency, while the target vector y_train contains binary labels indicating whether a transaction is genuine (0) or fraudulent (1). During training, the Random Forest algorithm constructs an ensemble of decision trees using bootstrap sampling and feature subspace splitting, and makes final predictions through majority voting, enabling the model to capture complex patterns and reduce overfitting. The trained model is then tested on unseen data **X_test**, which mirrors the structure and preprocessing of **X_train**, to classify new transactions. Predictions are evaluated against the true labels **y_test** using key metrics such as accuracy, precision, recall, F1-score, and a confusion matrix, offering a comprehensive understanding of the model’s ability to detect fraudulent behavior effectively.

Advantages of the Random Forest Classifier

Random Forest offers a powerful and reliable solution for fraud detection due to its high predictive accuracy and robustness, achieved through ensemble voting and decorrelated decision trees. It excels in handling imbalanced and noisy datasets—common in fraud scenarios—by effectively distinguishing rare fraudulent transactions from genuine ones. Unlike linear models, it captures complex, non-linear feature interactions and is resistant to overfitting thanks to its use of random feature selection and aggregated predictions. Its scalability and support for parallel processing make it ideal for high-volume e-commerce environments. Additionally, Random Forest provides built-in feature importance rankings, simplifying feature selection and aiding in interpreting fraud signals. It requires minimal hyperparameter tuning, performs well out-of-the-box, and is compatible with both binary and multiclass classification problems. Its support for probabilistic outputs enables threshold-based fraud risk scoring, and with proper optimization, it can be adapted for real-time or batch-based detection in production systems.

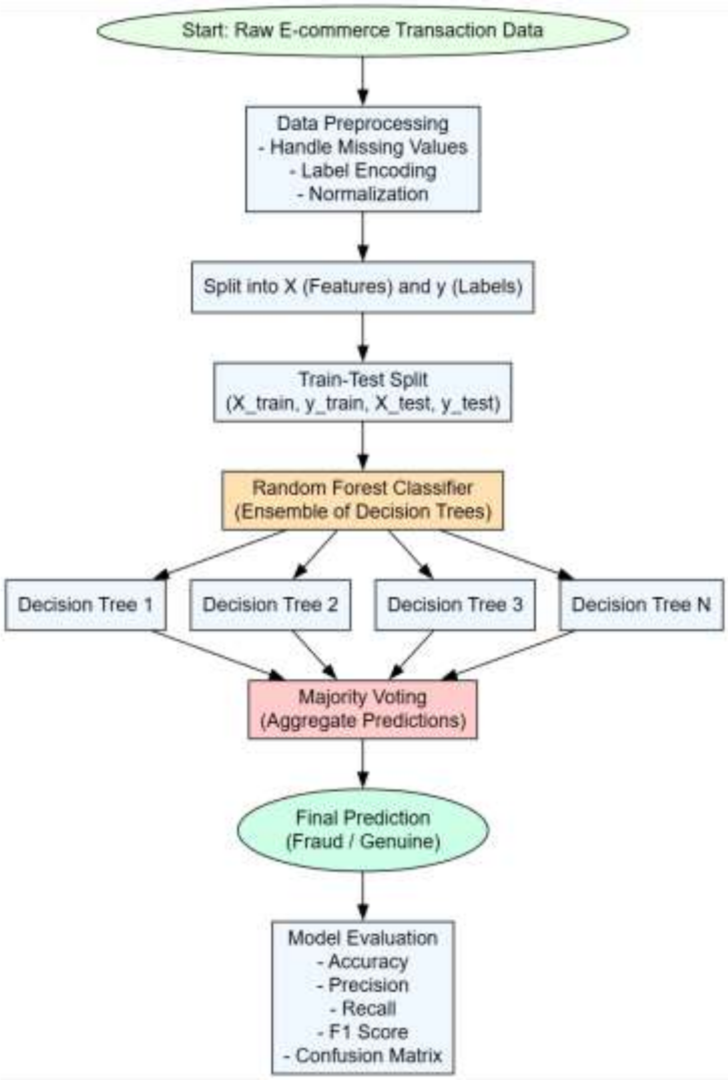

Figure 3: workflow of RFC

4. RESULTS

Figure 4 AI-based financial fraud detection web application is designed for clarity and functionality, featuring key elements that support real-time monitoring and analysis of transactions. At the top, the application title "Financial Fraud Detection" clearly communicates its core purpose. Navigation options such as "Home" and "User Login" provide easy access to various sections, with login functionality ensuring secure access to user-specific dashboards and features. Once logged in, users can leverage the system's real-time transaction monitoring capabilities, where financial data is continuously analyzed for suspicious patterns using advanced AI algorithms. These algorithms detect potentially fraudulent transactions by applying both predefined rules and machine-learned behavioral patterns, enabling swift identification and response to threats.

Figure 4: Homepage

Figure 5: Login Page

The figure 5 shows login form, accessible via the "Login" button on the homepage, presents a straightforward and secure interface for returning users. It consists of two primary input fields: "Username" and "Password." The "Username" field prompts users to enter the unique username they created during the signup process. The "Password" field, appropriately masked for security, requires users to input the corresponding password associated with their account. Below these fields, a prominent "login" button allows users to submit their credentials for verification. Upon successful authentication, users are granted access to their personalized accounts and the platform's mental health support services.

Figure 6: Upload and processing the data

The Figure 6 shows a portion of a web application interface for AI-based financial fraud detection. The user interface of the AI-based financial fraud detection system features a structured and workflow-driven design, beginning with a prominent header, “AI-Based Financial Fraud Detection in Real-Time Transactions,” that clearly communicates the application's objective. The tabbed navigation includes sections such as “Load & Process Dataset,” “Process Mining,” and “Run ML Algorithm,” reflecting a step-by-step approach to data handling and model execution, with the “Load & Process Dataset” tab currently active. Within this module, users can upload datasets for analysis via a “Browse Dataset” button, which opens a file selection dialog, allowing users to navigate their local system—illustrated by the selection of a file named "fraud_transaction.csv." The selected file path is displayed for confirmation, and a “Submit” button likely initiates the loading and preprocessing of the dataset, marking the first step in building and running the fraud detection model.

Figure 7: Pre-processed Data

The Figure 7 shows the data preview section of a dataset loader module in a financial fraud detection application. It confirms that a dataset has been loaded and provides a glimpse into the data's structure, including transaction IDs, fraud labels, amounts, and other anonymized features. The large number of columns suggests a rich dataset with many potential factors used for fraud detection.

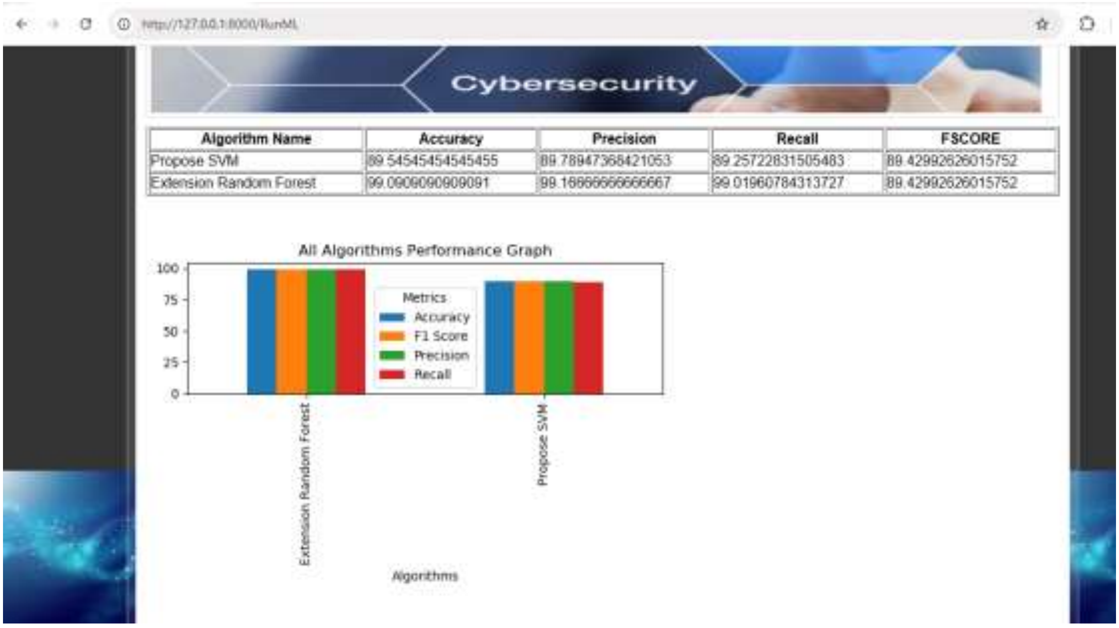

Figure 8: Comparison of Existing SVM and proposed RFC

The Figure 8 shows a comparison of the performance of two machine learning algorithms: Propose

Figure 9: upload Detect Fraud Dataset

The Figure 9 shows the data loading phase of a financial fraud detection application. Users can select a dataset file, which will then be loaded and processed. The application will use process mining and machine learning algorithms to detect fraudulent transactions, and the "Detect Fraud Module" allows for testing the model on a separate dataset. The interface is designed to guide users through the process of data loading, model training, and evaluation.

Figure 10: Predicted Outcomes

The Figure 10 shows the data is formatted in a way that's difficult to read directly, possibly representing anonymized or encoded transaction details. There are several rows visible, each with a long string of characters (likely representing numerical and/or categorical data). The rightmost column shows the outcome of the fraud detection process for each test case. The visible results are either Fraud or Normal.

5. CONCLUSION

The integration of Artificial Intelligence (AI) into financial fraud detection systems has significantly enhanced the ability of institutions to identify and prevent fraudulent activities. Traditional rule-based systems, while effective to a degree, often struggle to keep pace with the evolving tactics of fraudsters. AI-driven models, particularly those utilizing machine learning algorithms, offer a dynamic and adaptive approach to fraud detection. These systems can process vast amounts of transaction data in real-time, identifying complex patterns and anomalies that indicate fraudulent behavior. This capability not only improves detection rates but also reduces false positives, thereby enhancing operational efficiency and customer satisfaction. The financial sector has witnessed a substantial shift towards AI-powered solutions, with the AI in fraud detection market projected to grow at a compound annual growth rate (CAGR) of 24.5%, reaching approximately USD 108.3 billion by 2033. This growth underscores the increasing reliance on AI technologies to safeguard financial transactions. Institutions adopting these advanced systems benefit from real-time analysis, scalability, and the ability to adapt to new fraud patterns without extensive manual intervention. Moreover, AI's capacity to learn from historical data enables continuous improvement in detection capabilities, making it a formidable tool against sophisticated fraud schemes. However, the implementation of AI in fraud detection is not without challenges. Concerns regarding data privacy, the need for large datasets to train models effectively, and the potential for algorithmic biases must be addressed to ensure the ethical and effective use of AI. Additionally, the rapid advancement of AI technologies necessitates ongoing investment in infrastructure and talent to maintain and enhance system capabilities. In conclusion, AI has revolutionized financial fraud detection, offering unprecedented accuracy and efficiency in identifying fraudulent activities. As financial transactions continue to grow in volume and complexity, the role of AI in ensuring security and integrity becomes increasingly vital. By embracing AI-driven fraud detection systems, financial institutions can better protect themselves and their customers from the ever-evolving threat of fraud.

REFERENCES

[1] Hilal, W.; Gadsden, S.A.; Yawney, J. Financial Fraud: A Review of Anomaly Detection Techniques and Recent Advances. Expert Syst. Appl. 2021, 193, 116429.

[2] Ashtiani, M.N.; Raahemi, B. Intelligent Fraud Detection in Financial Statements Using Machine Learning and Data Mining: A Systematic Literature Review. IEEE Access 2021, 10, 72504–72525.

[3] Albashrawi, M. Detecting Financial Fraud Using Data Mining Techniques: A Decade Review from 2004 to 2015. J. Data Sci. 2016, 14, 553–570.

[4] Choi, D.; Lee, K. An Artificial Intelligence Approach to Financial Fraud Detection under IoT Environment: A Survey and Implementation. Secur. Commun. Netw. 2018, 2018, 1–15.

[5] Ngai, E.W.T.; Hu, Y.; Wong, Y.H.; Chen, Y.; Sun, X. The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. Decis. Support Syst. 2021, 50, 559–569.

[6] Ryman-Tubb, N.F.; Krause, P.; Garn, W. How Artificial Intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark. Eng. Appl. Artif. Intell. 2019, 76, 130–157.

[7] Al-Hashedi, K.G.; Magalingam, P. Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019. Comput. Sci. Rev. 2021, 40, 100402.

[8] Chaquet-ulldemolins, J.; Moral-rubio, S.; Muñoz-romero, S. On the Black-Box Challenge for Fraud Detection Using Machine Learning (II): Nonlinear Analysis through Interpretable Autoencoders. Appl. Sci. 2022, 12, 3856.

[9] Da'U, A.; Salim, N. Recommendation system based on deep learning methods: A systematic review and new directions. Artif. Intell. Rev. 2019, 53, 2709–2748.

[10] Zeng, Y.; Tang, J. RLC-GNN: An Improved Deep Architecture for Spatial-Based Graph Neural Network with Application to Fraud Detection. Appl. Sci. 2021, 11, 5656.

[11] Delamaire, L.; Hussein, A.; John, P. Credit card fraud and detection techniques: A review. Banks Bank Syst. 2019, 4, 57–68.

[12] Zhang, D.; Zhou, L. Discovering Golden Nuggets: Data Mining in Financial Application. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. 2021, 34, 513–522.

[13] Raj, S.B.E.; Portia, A.A. Analysis on credit card fraud detection methods. In Proceedings of the 2011 International Conference on Computer, Communication and Electrical Technology (ICCCET), Tirunelveli, India, 18–19 March 2021; pp. 152–156.

[14] Phua, C.; Lee, V.; Smith, K.; Gayler, R. A Comprehensive Survey of Data Mining-based Fraud Detection Research. arXiv 2020, arXiv:1009.6119.

[15] West, J.; Bhattacharya, M. Intelligent financial fraud detection: A comprehensive review. Comput. Secur. 2019, 57, 47–66.

[16] Abdallah, A.; Maarof, M.A.; Zainal, A. Fraud detection system: A survey. J. Netw. Comput. Appl. 2016, 68, 90–113.