World Journal of Advanced Engineering Technology and Sciences

eISSN: 2582-8266

Cross Ref DOI: 10.30574/wjaets

Journal homepage: https://wjaets.com/

(REVIEW ARTICLE)

Check for updates

Anomaly detection in financial transactions

Ganesh SP *, Aniruddha Nagesh Salvankar, Shawn Glanal Saldanha, Varun MC and Maryjo M George

Department of Artificial Intelligence and Machine Learning, Mangalore Institute of Technology & Engineering, India.

World Journal of Advanced Engineering Technology and Sciences, 2025, 15(02), 2120-2127

Publication history: Received on 09 April 2025; revised on 16 May 2025; accepted on 19 May 2025

Article DOI: https://doi.org/10.30574/wjaets.2025.15.2.0789

Abstract

Fraud detection in e-commerce has grown in importance as the volume of online transactions continues to rise. The rise in fraudulent behavior has led to significant financial losses and a decline in customer trust. This study explores the application of machine learning algorithms to identify fraudulent transactions, with a focus on anomaly detection methods. We examine many classification models, including XGBoost, Decision Tree, Random Forest, Bernoulli Naïve Bayes, and Logistic Regression, using a publicly available e-commerce fraud dataset. A range of performance criteria, such as the confusion matrix, F1-score, recall, accuracy, and precision, are used to assess the models. Random Forest achieved the highest accuracy (96.51%) of all the models tested, followed by XGBoost (95.22%) and Decision Tree (94.38%). With Optuna, Random Forest's accuracy was hyperparameter tuned to 97.08%. The results demonstrate the effectiveness of machine learning in detecting fraudulent transactions, with Random Forest emerging as the most dependable model. In addition to providing insights into improving fraud detection systems for e-commerce platforms, this research has the potential to inform future efforts aimed at improving model performance and real-time detection capabilities.

Keywords: Anomaly Detection; Financial Transactions; E-commerce Fraud; Random Forest; Optuna; Hyperparameter Tuning; Classification

1. Introduction

Fraud detection is a crucial concern in the fast-developing world of e-commerce. Unfortunately, the growing volume of online transactions has coincided with an increase in fraudulent activity, resulting in significant financial losses for firms and degradation of customer trust. The ability to reliably and swiftly detect fraudulent transactions is critical for ensuring the integrity of online marketplaces and protecting both merchants and customers. This study investigates the use of machine learning approaches to tackle this important issue, with an emphasis on anomaly detection in e-commerce financial transactions.

With the increased adoption of online shopping and digital payment methods, fraudsters have changed their strategies to exploit flaws in these systems. These strategies include everything from stolen credit card information and account takeovers to sophisticated phishing operations and malware attacks. Compared to typical brick-and-mortar purchases, online transactions frequently lack face-to-face verification, making them more vulnerable to fraud. This involves the creation of robust automated systems that can detect irregularities in real time. Undetected fraud can have serious consequences, including cash chargebacks, reputational damage, and a loss of customer confidence.

The goal of this research is to create and test a high-performance anomaly detection system for e-commerce financial transactions that uses machine learning. We use a huge dataset of real-world transactions to examine numerous algorithms, including Logistic Regression, Bernoulli Naive Bayes, Decision Tree, Random Forest, and XGBoost. We also look into the influence of hyperparameter adjustment using

* Corresponding author: Ganesh SP

Optuna to improve the performance of the best-performing model. Section 2 includes a review of related work. Section 3 outlines the proposed workflow, which includes data preprocessing, feature engineering, and model training. Section 4 presents the results and analysis. Finally, Section 5 summarizes the article and discusses future research directions.

2. Literature review

The use of machine learning techniques including Random Forest, Logistic Regression, SVMs, Decision Trees, and Naive Bayes for fraud detection has been highlighted in recent publications; ensemble and hybrid approaches have shown promise [1, 4]. Although the accuracy and applicability of these methods vary depending on the dataset, they are preferred. Even though real-world application presents obstacles, federated learning in conjunction with ANN improves fraud detection while maintaining privacy [2]. In e-commerce, supervised learning outperforms conventional rule-based techniques, and forthcoming developments might incorporate deep learning and reinforcement learning for increased precision [3, 5].

Furthermore, when it comes to detecting fraudulent credit card transactions, logistic regression performs better than Naïve Bayes and K-nearest neighbor, particularly when under-sampled for unbalanced datasets [6]. By using fuzzy clustering, the SBS system enhances under-sampling [7]. Despite the widespread use of random forests and artificial neural networks in e-commerce, little is known about bot and reseller fraud [8]. The increase in cybercrime since the epidemic has had an effect on mental health, underscoring the importance of cybersecurity awareness [9]. Fraudulent users are successfully identified by network analysis using random forest classifiers, indicating useful applications in online fraud detection [10]. These research highlight how crucial it is to modify detection techniques to account for emerging forms of fraud and how mental health specialists may support cybersecurity education.

Due to data scarcity in industries like banking and healthcare, recent data mining research have proposed methods like the GBAD framework for graph-based anomaly detection, emphasizing the focus on online social networks [11]. The need for different data sources is highlighted by challenges in credit card fraud detection, such as imbalanced datasets, dataset size, and the inaccessibility of real datasets [12]. In terms of accuracy and adaptability, machine learning—especially supervised learning—offers a sophisticated method of detecting fraud, outperforming conventional rule-based systems [13]. But since there isn't a single, universally applicable answer, models must be continuously assessed and adjusted to accommodate changing fraud strategies. Additionally, e-commerce tasks like product suggestion, customer care, and commerce forecasting are increasingly being handled by learning-based techniques such graph neural networks, deep learning, and reinforcement learning, showcasing their adaptability and potential for innovation [14, 15].

Several studies have focused on certain methods for spotting fraudulent activity on online platforms and credit card transactions, building on the use of machine learning in fraud detection. For example, studies have revealed that Sybil accounts and bot farms are frequently used to create phony evaluations on e-commerce websites, underscoring the need for more potent detection techniques [16]. While support vector machines and logistic regression have shown promise in managing unbalanced datasets, stacking classifiers and neural networks have emerged as the best performing supervised and unsupervised algorithms in credit fraud detection [17, 20]. Additionally, it has been suggested that cooperative methods, including those that make use of blockchain technology, can improve fraud detection systems by allowing e-commerce companies to jointly develop strong machine learning models while maintaining privacy and encouraging involvement [18]. Furthermore, sophisticated frameworks that link system performance to operational risks have been developed to enhance banks anti-fraud systems by integrating AI models such as neural networks and decision trees [19]. Together, these studies highlight how crucial it is to combine cutting-edge technologies with cooperative tactics in order to successfully fight fraud in a variety of contexts.

3. Proposed workflow

Data collection, preprocessing, model design, training, selection, and hyperparameter tuning are all covered in this section's discussion of the suggested system.

3.1. Data acquisition

The data is sourced from a public dataset on Kaggle. The dataset includes two CSV files for training and testing.

https://www.kaggle.com/datasets/aryanrastogi7767/ecommerce-fraud-data.

3.2. Preprocessing

Data preprocessing refines the dataset by reducing noise and enhancing its relevance for modeling tasks. The following steps are taken:

- **Date Conversion and Feature Extraction:** The Transaction Date column is converted from object to date time format. Additional features like Transaction Day, Day of Week, and Transaction Month are developed.
- **Handling Customer Age:** Invalid values in the Customer Age field is corrected. Values between -9 and 8 are replaced with the mean, and extremely negative values are converted to absolute equivalents.
- **Address Matching:** A new feature, Is Address Match, is introduced to indicate whether the shipping and billing addresses are the same.
- **Irrelevant Feature Removal:** Columns including Transaction ID, Customer ID, Customer Location, IP Address, and address-related features are removed because they are redundant or irrelevant to fraud detection.
- **Data Optimization:** Integer and float columns are downcast to reduce memory consumption.

3.3. Model architecture

The model architecture is depicted in Fig.1.

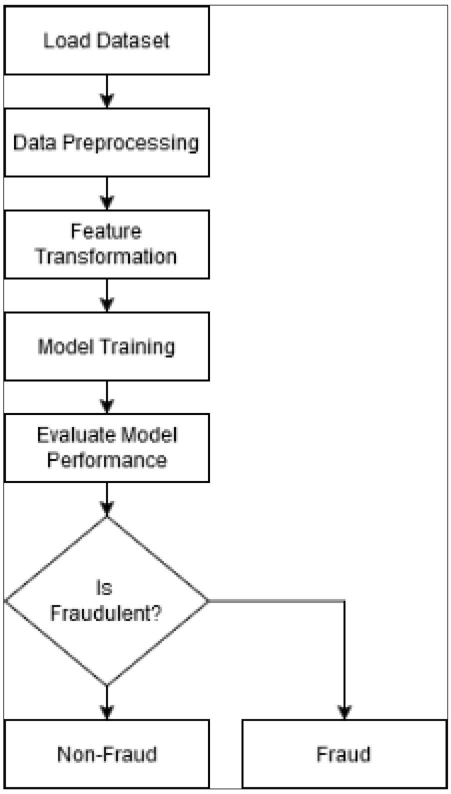

Figure 1 Model Architecture

3.4. Model building and training

The model development process includes evaluating multiple algorithms and fine-tuning the preprocessing pipeline:

- **Transformations:** Numerical columns are standardized, but categorical variables are one-hot encoded. A Column Transformer is used to carry out these changes.
- **Candidate Models:** Several classifiers are tested, including: Logistic Regression, Bernoulli Naïve Bayes, Decision Tree, Random Forest and XGBoost.
- **Pipeline Implementation:** Each model is integrated into a pipeline that includes preprocessing steps to ensure consistency and streamline the training process.

3.5. Model selection and hyperparameter tuning

Random Forest emerges as the best-performing model in terms of accuracy and robustness. The following optimization strategies are used:

- **Hyperparameter Tuning with Optuna:** Optuna's optimization framework is used to tune parameters including learning_rate, n_estimators, max_depth, min_child_weight, gamma, subsample, colsample_bytree, and reg_alpha.
- **Cross-Validation:** Cross-validation is used during the optimization process to ensure that the model is stable and generalizable across different subsets of the training data.
- **Best Parameters:** After 70 trials, the optimal hyperparameters are identified which significantly improves the model's predictive accuracy.

3.6. User interface

Using Flask, the system's web-based interface was created to make interacting with the fraud detection model easier. It allows users to view predictions, enter transaction data, and comprehend possible signs of fraud. These are the components that make up the interface:

- **Homepage:** The homepage welcomes users to the system, guides them to the transaction form for fraud detection, and offers a straightforward and user-friendly starting point.
- **Transaction Form:** The transaction date, amount, product category, payment method, IP address, shipping and billing addresses, account age, device used, quantity, and customer information (such as age and location) are all captured in an easy-to-use form. The form guarantees that all necessary fields are filled out in order to preserve data consistency for precise forecasts.
- **Prediction Results:** The results page shows the fraud probability expressed as a percentage as well as whether the transaction was reported as fraudulent or legitimate. An LLM generates a comprehensive explanation for flagged transactions that helps users understand the prediction by highlighting possible fraud indicators like mismatched addresses or unusual transaction amounts.

4. Results and Analysis

Accuracy, precision, recall, F1-score, and confusion matrix were used to assess the models. Below is a detailed breakdown of each model's classification results.

4.1. Logistic regression

Accuracy: 0.6963

4.1.1. Confusion Matrix

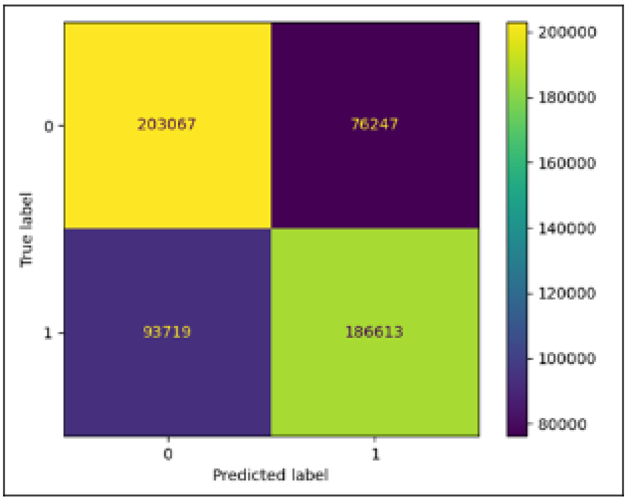

Figure 2 Confusion Matrix for Logistic regression

The accuracy of Logistic Regression was 69.63%. With 203,067 true negatives and 186,613 true positives, the confusion matrix shows a reasonably balanced performance. However, there are more false positives (76,247) than false negatives (93,719). Class 0 (non-fraud) had precision and recall of 0.68 and 0.73, respectively, while class 1 (fraud) had precision and recall of 0.71 and 0.67. As a result, the F1-scores for classes 0 and 1 were 0.70 and 0.69, respectively. This implies that there is a trade-off between precision and recall, and that Logistic Regression has some difficulty in predicting fraud cases.

4.2. Bernoulli naïve bayes (nb)

Accuracy: 0.6451

4.2.1. Confusion Matrix

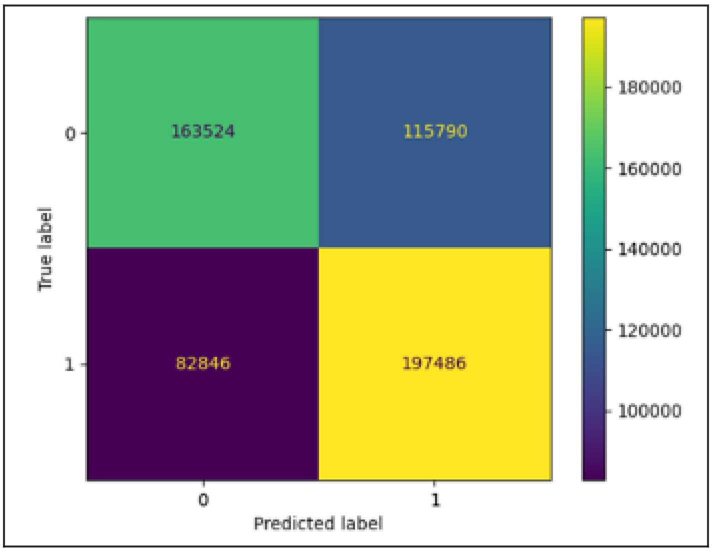

Figure 3 Confusion Matrix for Bernoulli naïve bayes

The accuracy of the Bernoulli Naïve Bayes classifier was 64.51%. There are a considerable number of false positives (115,790) and false negatives (82,846) in the model's confusion matrix, which also shows 163,524 true negatives and 197,486 true positives. Class 1 (fraud) had precision and recall of 0.63 and 0.70, respectively, while class 0 (non-fraud) had precision and recall of 0.66 and 0.59. This model performs poorly in precision, often classifying legitimate transactions as fraudulent, but it has a respectable recall for fraud cases (F1-score of 0.62 for class 0 and 0.67 for class 1).

4.3. Decision tree

Accuracy: 0.9438

4.3.1. Confusion Matrix

The Decision Tree classifier produced a remarkable 94.38% accuracy rate. False positives (17,436) and false negatives (14,059) are extremely rare, while the confusion matrix displays 261,878 true negatives and 266,273 true positives. It performs exceptionally well in classification, with a precision of 0.95 and recall of 0.94 for class 0 (non-fraud) and 0.94 and 0.95 for class 1 (fraud), respectively. Both classes' F1-scores of 0.94 show a balanced performance, indicating that Decision Tree successfully distinguishes between fraudulent and non-fraud transactions with few misclassifications.

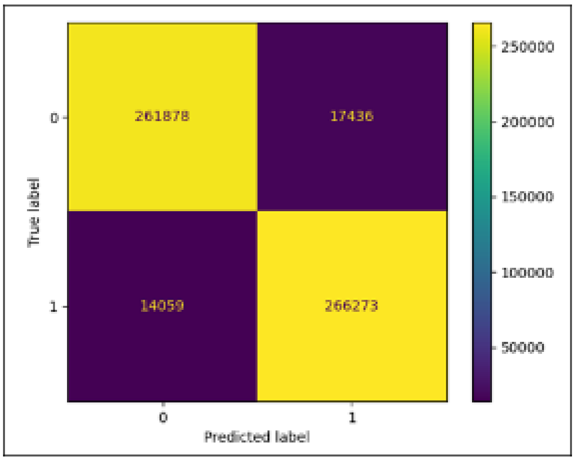

Figure 4 Confusion Matrix for Decision Tree

4.4. Random forest

Accuracy: 0.9651

4.4.1. Confusion Matrix

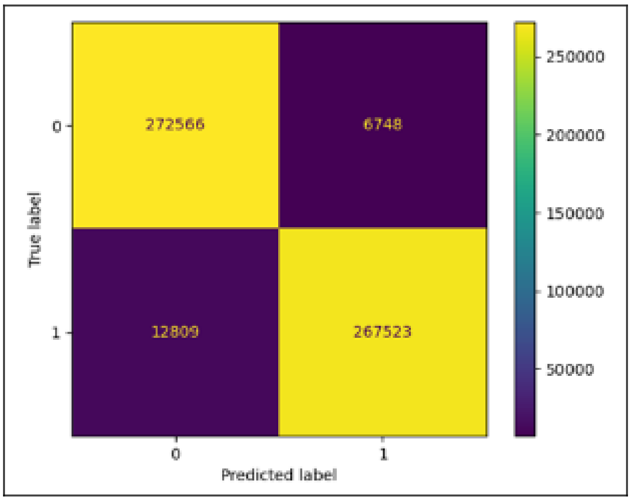

Figure 5 Confusion Matrix for Random Forest

Random Forest was the most accurate at 96.51%. In the confusion matrix, there are only 6,748 false positives and 12,809 false negatives, whereas there are 272,566 real negatives and 267,523 true positives. This classifier demonstrated exceptional classification abilities with a precision of 0.96 and recall of 0.98 for class 0 and a precision of 0.98 and recall of 0.95 for class 1. With F1-scores of 0.97 and 0.96, it demonstrates an exceptional capacity to detect both fraudulent and non-fraudulent transactions while lowering false positives and false negatives.

4.5. Xgboost

Accuracy: 0.9522

Confusion Matrix

XGBoost yielded a 95.22% accuracy rate. There are 274,238 true negatives and 258,644 true positives, along with 5,076 false positives and 21,688 false negatives, according to the confusion matrix. Class 1 (fraud) had precision and recall of 0.92 and 0.98, respectively, while class 0 (non-fraud) had precision and recall of 0.93 and 0.98. With strong precision for fraud detection and somewhat higher false negatives for fraud cases compared to non-fraud cases, XGBoost performs nearly equally well in identifying both classes, according to the F1-scores of 0.95 and 0.95.

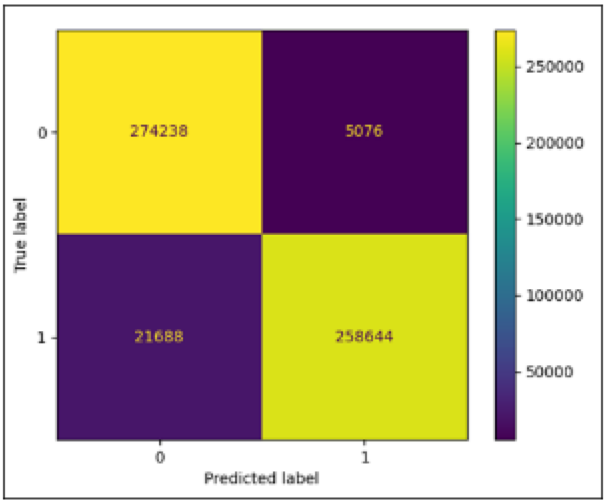

Figure 6 Confusion Matrix for Xgboost

4.6. Optuna hyperparameter tuning

Optuna was used to tune hyperparameters for the Random Forest model. After 70 trials, the optimal combination of hyperparameters was discovered, resulting in an increase in accuracy from 0.9651 to 0.9708. The ideal hyperparameters comprised a learning rate of 0.2381, 439 estimators, a maximum depth of 6, and other parameters such as subsample and reg_alpha, all of which contributed to improved model performance in detecting fraudulent transactions.

5. Conclusion

With a focus on anomaly detection, this study examined the application of machine learning techniques to detect fraudulent activity in e-commerce transactions. The study began with an overview of the growing concern about online fraud and the need for robust detection tools. Using a publicly available Kaggle dataset, the data gathering process was then described, along with the preparation steps taken to get the data ready for modeling. Several classification algorithms, such as Logistic Regression, Bernoulli Naïve Bayes, Decision Tree, Random Forest, and XGBoost, were evaluated using accuracy, precision, recall, and F1-score. Random Forest was the most accurate of these, and its performance was enhanced by modifying its hyperparameters with Optuna. Users may submit transaction data and receive thorough explanations and fraud forecasts once the finished system was integrated into a Flask-based web interface. The findings demonstrate that machine learning models, specifically Random Forest and XGBoost, can effectively identify fraudulent transactions, offering a practical tool for boosting the security and credibility of e-commerce platforms.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Najem, S. M., & Kadeem, S. M. (2021). A survey on fraud detection techniques in e-commerce. Tech-Knowledge, 1(1), 33-47.

[2] Bin Sulaiman, R., Schetinin, V., & Sant, P. (2022). Review of machine learning approach on credit card fraud detection. Human-Centric Intelligent Systems, 2(1), 55-68.

[3] Reddy, S. R. B., Kanagala, P., Ravichandran, P., Pulimamidi, R., Sivarambabu, P. V., & Polireddi, N. S. A. (2024). Effective fraud detection in e-commerce: Leveraging machine learning and big data analytics. Measurement: Sensors, 33, 101138.

[4] Islam, M. A., Uddin, M. A., Aryal, S., & Stea, G. (2023). An ensemble learning approach for anomaly detection in credit card data with imbalanced and overlapped classes. Journal of Information Security and Applications, 78, 103618.

[5] Bello, H. O., Ige, A. B., & Ameyaw, M. N. (2024). Adaptive machine learning models: concepts for real-time financial fraud prevention in dynamic environments. World Journal of Advanced Engineering Technology and Sciences, 12(02), 021-034.

[6] Itoo, F., Meenakshi, & Singh, S. (2021). Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. International Journal of Information Technology, 13(4), 1503-1511.

[7] Ahmad, H., Kasasbeh, B., Aldabaybah, B., & Rawashdeh, E. (2023). Class balancing framework for credit card fraud detection based on clustering and similarity-based selection (SBS). International Journal of Information Technology, 15(1), 325-333.

[8] Mutemi, A., & Bacao, F. (2024). E-Commerce Fraud Detection Based on Machine Learning Techniques: Systematic Literature Review. Big Data Mining and Analytics, 7(2), 419-444.

[9] Monteith, S., Bauer, M., Alda, M., Geddes, J., Whybrow, P. C., & Glenn, T. (2021). Increasing cybercrime since the pandemic: Concerns for psychiatry. Current psychiatry reports, 23, 1-9.

[10] Kodate, S., Chiba, R., Kimura, S., & Masuda, N. (2020). Detecting problematic transactions in a consumer-to-consumer e-commerce network. Applied Network Science, 5(1), 90.

[11] Pourhabibi, T., Ong, K. L., Kam, B. H., & Boo, Y. L. (2020). Fraud detection: A systematic literature review of graph-based anomaly detection approaches. Decision Support Systems, 133, 113303.

[12] Aziz, A., & Ghous, H. (2021). Fraudulent transactions detection in credit card by using data mining methods: A review. Int. J. Sci. Prog. Res., 79(1), 31-48.

[13] Lim, K. S., Lee, L. H., & Sim, Y. W. (2021). A review of machine learning algorithms for fraud detection in credit card transaction. International Journal of Computer Science & Network Security, 21(9), 31-40.

[14] Policarpo, L. M., da Silveira, D. E., da Rosa Righi, R., Stoffel, R. A., da Costa, C. A., Barbosa, J. L. V., ... & Arcot, T. (2021). Machine learning through the lens of e-commerce initiatives: An up-to-date systematic literature review. Computer Science Review, 41, 100414.

[15] Yin, S., & Luo, X. (2021, November). A review of learning-based E-commerce. In 2021 16th International Conference on Intelligent Systems and Knowledge Engineering (ISKE) (pp. 483-490). IEEE.

[16] Paul, H., & Nikolaev, A. (2021). Fake review detection on online E-commerce platforms: a systematic literature review. Data Mining and Knowledge Discovery, 35(5), 1830-1881.

[17] Gamini, P., Yerramsetti, S. T., Darapu, G. D., Pentakoti, V. K., & Vegesena, P. R. (2021). A review on the performance analysis of supervised and unsupervised algorithms in credit card fraud detection. International Journal of Research in Engineering, Science and Management, 4(8), 23-26.

[18] Pranto, T. H., Hasib, K. T. A. M., Rahman, T., Haque, A. B., Islam, A. N., & Rahman, R. M. (2022). Blockchain and machine learning for fraud detection: A privacy-preserving and adaptive incentive based approach. IEEE Access, 10, 87115-87134.

[19] Festa, Y. Y., & Vorobyev, I. A. (2022). A hybrid machine learning framework for e-commerce fraud detection. Model Assisted Statistics and Applications, 17(1), 41-49.

[20] Verma, P., & Tyagi, P. (2022). Analysis of supervised machine learning algorithms in the context of fraud detection. ECS Transactions, 107(1), 7189.