Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 16, Issue. 1: 2025
ISSN : **1906-9685**

Machine Learning Techniques for E-Commerce Fraud Detection: A Systematic Literature Review on Voting Classifiers

#1B.AMARNATHREDDY, #2 SK.SHAHIRUN

#1 Assistant Professor #2 M.C.A Scholar

Department of Master of Computer Applications,

Qis College of Engineering and Technology

Abstract: The exponential growth of e-commerce, further accelerated by the COVID-19 pandemic, has led to a surge in digital fraud, creating significant financial and security challenges. Addressing these risks requires the development of advanced and reliable fraud detection systems, often constrained by limited real-world data. This work utilizes the Fraudulent E-Commerce Transaction Data dataset to enhance fraud detection capabilities. To address class imbalances in the dataset, Synthetic Minority Oversampling Technique (SMOTE) is applied, ensuring equitable representation of fraudulent and legitimate transactions. Multiple machine learning models were evaluated to identify the most effective approach for accurate fraud detection. Among the methods tested, the Voting Classifier, which integrates Bagging with Random Forest and Boosted Decision Trees, achieved the highest accuracy of 97.2%, demonstrating its superiority in detecting fraudulent activities. The results highlight the importance of combining ensemble techniques and advanced sampling strategies to improve predictive performance in e-commerce fraud detection, offering a robust solution to combat emerging threats.

"Index Terms *–E-commerce; fraud detection; Machine Learning (ML); systematic review; organized retail f*

1. INTRODUCTION

The COVID-19 pandemic has significantly accelerated the shift towards online communication and e-commerce, with more people relying on digital platforms for daily activities such as work, school, shopping, doctor's appointments, and entertainment [1]. E-commerce platforms like Amazon, eBay, and the Facebook Marketplace have experienced substantial growth, largely driven by reduced mobility and the fear of contracting the virus. This surge in online activities has brought about a corresponding increase in cybercrimes and fraud, with fraudsters exploiting the expanded digital footprint to commit crimes [2]. As more people engage with digital platforms, the global economy faces billions of dollars in losses each year due to cybercrime, undermining both public safety and financial stability [3].

Fraud and cybercrime encompass a wide range of criminal activities, including extortion, blackmail, phishing, malware attacks, fraudulent transactions on e-commerce platforms, romance scams, and tech

support scams [2]. Other pervasive forms of fraud in the digital era include credit card theft, money laundering, and fraudulent financial transactions, which pose serious risks to businesses and individuals alike [2], [4]. These illicit activities not only harm the financial health of victims but also damage businesses' reputations and cause significant psychological distress.

According to a recent analysis by Juniper Research, losses from fraudulent online payments are growing at an alarming rate of 18 percent annually, highlighting the urgent need for robust fraud detection and prevention measures [5]. Despite ongoing efforts, current strategies often struggle to keep pace with increasingly sophisticated fraudsters who continuously adapt their methods to exploit vulnerabilities in e-commerce platforms [6]. The lack of practical data, coupled with the reluctance of businesses to share sensitive information to protect platform security, further complicates the development of effective fraud prevention systems. In this context, while fraud prevention aims to avert the occurrence of these criminal activities, detection systems remain essential for identifying fraud as soon as it occurs [7], [8].

2. RELATED WORK

The increasing prevalence of e-commerce platforms and the shift towards online transactions have simultaneously led to a significant rise in cybercrimes and fraudulent activities. As highlighted by Ali et al. [9], financial fraud detection has garnered substantial attention due to the growing threat posed by malicious actors in the online space. Machine learning (ML) techniques, in particular, have emerged as a valuable tool in detecting fraudulent activities, offering improved accuracy and adaptability compared to traditional methods. A systematic review of these techniques reveals that ML-based approaches can enhance the effectiveness of fraud detection systems by automating the identification of suspicious patterns in vast datasets, which would otherwise be difficult to manage manually.

In the context of e-commerce, fraud detection and prevention are of paramount importance. Rodrigues et al. [10] explore various ML models and their applications in preventing fraudulent transactions on e-commerce platforms. They emphasize that fraud detection in online retail systems faces unique challenges, including large transaction volumes, rapid transaction times, and the ability of fraudsters to continuously adapt their methods. Traditional fraud detection methods often rely on rule-based systems, which may not be sufficient to address these challenges. By contrast, machine learning models can learn from historical data, adapt to new fraud patterns, and improve detection accuracy over time. Rodrigues et al. suggest that hybrid models, combining different machine learning techniques, offer promising results for detecting fraud in e-commerce systems.

In addition to the general ML-based fraud detection systems, credit card fraud detection has been a particular focus area. Xournals [11] offers a comprehensive review of various credit card fraud detection techniques in e-commerce. One of the critical challenges in credit card fraud detection is the imbalanced nature of the dataset, where fraudulent transactions represent a small percentage of the total transactions. As a result, detecting fraudulent transactions becomes a challenging task for traditional classifiers. Machine learning models, such as Random Forest, Support Vector Machines (SVM), and deep learning approaches, have shown considerable promise in improving the detection rates by efficiently handling imbalanced datasets

and identifying subtle fraud patterns. Additionally, ensemble models, which combine multiple algorithms to improve predictive performance, have gained attention in recent years due to their ability to boost accuracy and reduce the risk of overfitting.

Moreover, the integration of blockchain technology with machine learning for fraud detection has been explored as a novel approach. Pranto et al. [12] discuss the potential of combining blockchain with machine learning techniques to create a more secure and transparent fraud detection system. Blockchain provides a decentralized and immutable ledger, which can significantly enhance the transparency of transactions, thereby making fraudulent activities easier to detect. By applying ML algorithms to the data stored on the blockchain, organizations can identify fraudulent patterns more effectively. This hybrid approach is particularly useful for e-commerce platforms that deal with large volumes of transactions and require robust, secure systems to combat fraud.

Festa and Vorobyev [13] propose a hybrid ML framework for e-commerce fraud detection that combines decision trees, neural networks, and clustering techniques. Their framework aims to address the limitations of existing fraud detection models by integrating multiple machine learning techniques into a single system. The proposed approach is designed to improve detection accuracy while reducing false positives, a common issue in fraud detection systems. By employing ensemble methods, the model is able to better capture complex relationships between features and improve its ability to detect various types of fraudulent activities. This hybrid framework demonstrates that a combination of diverse ML techniques can offer significant advantages in the detection of online fraud.

Feature selection plays a crucial role in improving the performance of fraud detection models. In the study by Ileberi et al. [14], the authors focus on credit card fraud detection using genetic algorithms (GA) for feature selection. Feature selection is an essential step in the development of machine learning models, as it helps reduce the dimensionality of the data and ensures that the model is focused on the most relevant attributes. By using genetic algorithms to select features, the authors were able to enhance the performance of their fraud detection system and improve its predictive accuracy. This approach demonstrates the importance of choosing the right features in building effective fraud detection models.

In addition to feature selection, various machine learning algorithms have been applied to fraud detection models. Nasr et al. [15] present a proposed fraud detection model based on e-payment attributes, specifically for Egyptian e-payment gateways. Their model employs a combination of data mining and machine learning techniques to identify fraudulent activities in real-time transactions. The study demonstrates the effectiveness of combining various approaches to detect fraud in the e-commerce space, particularly in regions where digital payment systems are rapidly growing but may lack adequate fraud detection mechanisms.

Lim and Ahn [16] explore fraud detection techniques in the context of peer-to-peer (P2P) platforms, such as C2C (consumer-to-consumer) marketplaces. Their research emphasizes the importance of contextual information, such as transaction descriptions and user behaviors, in identifying fraudulent activities. By utilizing machine learning techniques such as Doc2Vec for textual data representation, the authors were able to improve the detection of fraudulent transactions in

the P2P environment. This approach highlights the potential of natural language processing (NLP) and unsupervised learning methods in enhancing fraud detection systems, particularly in more complex and decentralized transaction environments.

3. MATERIALS AND METHODS

We proposes a system that aims to develop an efficient e-commerce fraud detection mechanism using machine learning techniques. The system will leverage the Fraudulent E-Commerce Transaction Data [25] dataset, with a focus on addressing class imbalances through the application of Synthetic Minority Oversampling Technique [22] (SMOTE). A range of machine learning algorithms will be employed to build and evaluate various fraud detection models, including Logistic Regression, Random Forest, Decision Tree, Naive Bayes, SVM, ANN-MLP, KNN, [24] XGBoost, CatBoost, [22] AdaBoost, and Gradient Boosting. Additionally, an ensemble approach using a Voting Classifier, which combines Bagging with Random Forest and Boosted Decision Trees, will be explored to enhance model robustness. The goal is to identify the most effective model for detecting fraudulent transactions while maintaining high accuracy, precision, and recall. By integrating these algorithms and advanced sampling techniques, the system aims to provide a comprehensive solution for real-time fraud detection in e-commerce platforms.

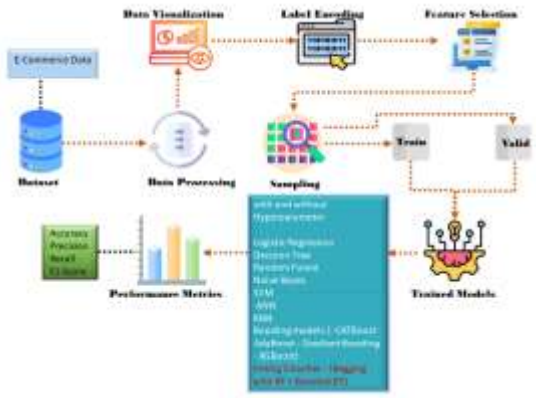

Fig.1 Proposed Architecture

The image (Fig.1) depicts a typical machine learning workflow for an e-commerce dataset. It starts with data visualization, followed by label encoding and feature selection. The dataset is then preprocessed and split into training and validation sets. Various machine learning models, including logistic regression, decision trees, random forests, and neural networks, are trained on the training data. The models are then evaluated using performance metrics like accuracy, precision, recall, and F1-score. The best-performing model is selected for further use.

i) Dataset Collection:

The dataset [25] "Fraudulent_E-Commerce_Transaction_Data_" contains 23,634 entries and 14 columns (Fig.2) initially. It includes categorical features such as 'Transaction ID', 'Customer ID', 'Transaction Amount', 'Payment Method', 'Product Category', 'Quantity', 'Customer Age', 'Customer Location', 'Device Used', 'IP Address', 'Shipping Address', 'Billing Address', 'Is Fraudulent', 'Account Age Days', and 'Transaction Hour'.

Fig.2 Dataset Collection Table

After preprocessing, null and duplicate entries were removed. Columns that were deemed less relevant for fraud detection, such as 'Transaction ID', 'Customer ID', 'Transaction Date', 'Customer Location', 'IP Address', 'Shipping Address', and 'Billing Address', were dropped. This resulted in a final dataset with 9 columns. These include transactional details like 'Transaction Amount', 'Payment Method', 'Product Category', 'Quantity', 'Customer Age', 'Device Used', 'Account Age Days', 'Transaction Hour', and the target variable 'Is Fraudulent'.

ii) Pre-Processing:

Data pre-processing is crucial to prepare the dataset for machine learning models. It includes cleaning, transforming, and balancing data to improve the accuracy and efficiency of predictive models.

a) Data Processing: In this step, duplicate data entries were identified and removed to ensure dataset integrity. Irrelevant and redundant columns were dropped to streamline the dataset and reduce noise. Additionally, missing or inconsistent values were handled to ensure a clean, reliable dataset for analysis. This processing ensures the data is ready for further modeling and evaluation.

b) Data Visualization: Data visualization is essential for understanding the patterns, trends, and relationships within the dataset. Various visual techniques like histograms, bar charts, and scatter plots are used to explore distributions, correlations, and potential outliers. This helps in identifying key features, guiding the selection of the most relevant attributes for model training. Visualization enhances data-driven decision-making by making complex information more accessible and understandable.

c) Label Encoding: Label encoding is a method used to convert categorical string values into numerical representations. This step is necessary as machine learning models require numerical input. By converting categories into integer labels, the data becomes suitable for algorithmic processing, while maintaining the information about different classes. Label encoding ensures that models can handle categorical variables effectively without losing the integrity of the data's structure.

d) Oversampling: Oversampling is used to address class imbalance in the dataset, where one class (e.g., fraudulent transactions) is underrepresented compared to the other. [22] SMOTE (Synthetic Minority Over-sampling Technique) generates synthetic samples of the minority class to balance the dataset. This process helps improve the model's ability to learn patterns in the minority class, ensuring it does not bias towards the majority class and can make accurate predictions for both classes.

iii) Training &Testing:

The dataset is split into training and testing sets in an 80:20 ratio, with 80% allocated for training the model and 20% reserved for testing its performance. This split ensures that the model is trained on a substantial portion of the data, allowing it to learn patterns and make predictions effectively. The test set provides an unbiased evaluation of the model's generalization ability, helping assess its performance on unseen data. This split is crucial for preventing overfitting and ensuring reliable results.

iv) Algorithms:

Logistic Regression: A statistical method for binary classification that models the probability of a certain class (fraudulent or non-fraudulent) based on input features. [17] It is widely used for fraud

detection due to its simplicity, interpretability, and effectiveness on linearly separable data.

Random Forest: An ensemble method that constructs multiple decision trees and combines their outputs to improve predictive accuracy. [18] Random Forest is used for fraud detection by providing robust predictions, handling high-dimensional data, and reducing overfitting compared to individual decision trees.

Decision Tree: A supervised learning algorithm that splits data into subsets based on feature values, leading to a tree structure for prediction. In fraud detection, [19] Decision Trees are used to identify which features most influence the classification of transactions as fraudulent.

Naive Bayes: A probabilistic classifier based on Bayes' [20] theorem, assuming independence between features. It is used for fraud detection by calculating the likelihood of fraud based on feature values, providing a fast and efficient solution, especially for high-dimensional datasets.

SVM (Support Vector Machine): A machine learning algorithm that finds the optimal hyperplane separating data into distinct classes. [21] SVM is applied to fraud detection to classify transactions by identifying the boundaries between legitimate and fraudulent activities, even in complex, high-dimensional datasets.

ANN-MLP (Artificial Neural Network - Multi-layer Perceptron): A type of neural network that uses multiple layers of neurons to model complex patterns in data. In fraud detection, MLP is used to capture non-linear relationships between features, identifying subtle patterns associated with fraudulent transactions.

KNN (K-Nearest Neighbors): A simple, instance-based learning algorithm that classifies a transaction based on its proximity to the nearest labeled data points. [23] KNN is effective for fraud detection by classifying transactions based on similarity to known fraudulent or non-fraudulent cases.

XGBoost: An efficient implementation of gradient boosting that constructs an ensemble of decision trees. [24] XGBoost is used for fraud detection due to its high accuracy, ability to handle imbalanced datasets, and speed in training large datasets while optimizing predictive performance.

CatBoost: A gradient boosting algorithm that is highly efficient with categorical features. [21] CatBoost is applied in fraud detection to improve predictive accuracy by handling categorical variables more effectively, capturing complex patterns that traditional models may struggle to learn.

AdaBoost: A boosting algorithm that combines weak classifiers to create a strong classifier by focusing on errors made in previous iterations. AdaBoost [22] is used in fraud detection to improve performance by iteratively correcting misclassified transactions, enhancing classification accuracy.

Gradient Boosting: A boosting method that builds decision trees sequentially, where each tree corrects the errors of its predecessor. [21] Gradient Boosting is employed in fraud detection to enhance the model's ability to classify complex, subtle fraudulent behaviors by iteratively improving predictions.

Voting Classifier (Bagging with RF + Boosted DT): An ensemble method that combines the outputs of multiple classifiers, typically using

Bagging (Random Forest) and Boosting (Boosted Decision Trees). The Voting Classifier aggregates predictions to achieve a higher accuracy, offering a powerful solution for complex fraud detection tasks.

4. RESULTS & DISCUSSION

Accuracy: The accuracy of a test is its ability to differentiate the patient and healthy cases correctly. To estimate the accuracy of a test, we should calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as:

$$Accuracy = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} (1)$$

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

$$Precision = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} (2)$$

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

$$Recall = \frac{\text{TP}}{\text{TP} + \text{FN}} (3)$$

F1-Score: F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

$$F1\ Score = 2 * \frac{Recall\ \text{X}\ Precision}{Recall + Precision} * 100 (1)$$

Fig. 3 Dash Board

The Fig. 3 shows the user dashboard of an e-commerce fraud detection system. It has a welcoming message and an illustration of people working with data. There is also a "Signup" button.

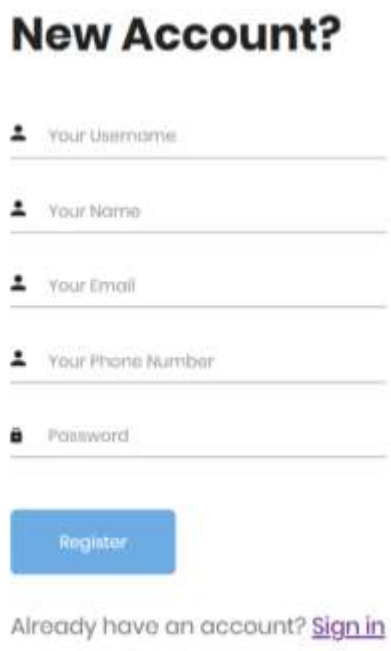

Fig. 4 Register Page

The Fig. 4 shows a user registration form. It requires a username, name, email, phone number, and password. It also includes a "Register" button and a link to "Sign in" for existing users.

Fig. 5 Login page

The Fig. 5 shows a login page with the message "Welcome Back." The username field is pre-filled with "admin." It also has a password field and a "Log In" button. There is also an option to "Remember me" and a "Forgot Password" link. Additionally, users can sign up for a new account.

Fig. 6 Main page

The Fig. 6 shows the main page of a dashboard with the "Prediction" tab selected. This suggests that the user is likely interested in viewing or analyzing predictions generated by the system.

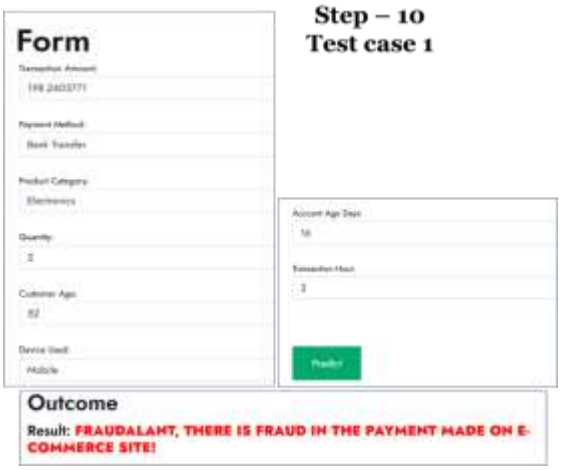

Fig. 7 Test case – 1

The Fig. 7 shows a form for detecting e-commerce fraud. It collects data like transaction amount, payment method, product category, quantity, customer age, and device used. After inputting data, the form predicts that "FRAUDALANT, THERE IS FRAUD IN THE PAYMENT MADE ON E-COMMERCE SITE!"

Fig. 8 Test case – 2

The Fig. 8 shows a form for detecting e-commerce fraud. It collects data like transaction amount, payment method, product category, quantity, customer age, and device used. After inputting data, the form predicts that "NON-FRAUDALANT, THERE IS NO FRAUD IN THE PAYMENT MADE ON E-COMMERCE SITE!"

5. CONCLUSION

In conclusion, the proposed system effectively addresses the growing challenge of e-commerce fraud detection by employing a range of machine learning algorithms and advanced techniques. The use of the Fraudulent E-Commerce Transaction Data dataset and the application of SMOTE for oversampling ensured a balanced representation of fraudulent and legitimate transactions. Among the various models tested, the Voting Classifier, which combines Bagging with Random Forest and Boosted Decision Trees, emerged as the highest-performing model, achieving an impressive accuracy of 97.2%. This model demonstrated exceptional capability in accurately distinguishing between fraudulent and non-fraudulent transactions, outperforming other algorithms in terms of accuracy, precision, and recall. The results underscore the effectiveness of ensemble methods, specifically the combination of Bagging and Boosting techniques, in improving predictive performance in fraud detection tasks. By leveraging these machine learning techniques, the system offers a robust solution for identifying fraudulent activities in e-commerce, contributing significantly to the security and financial stability of online platforms.

Future work will focus on further enhancing the fraud detection system by exploring more sophisticated machine learning and deep learning models, such as recurrent neural networks (RNN) and transformer-based architectures. Additionally, feature engineering techniques and hyperparameter optimization methods like grid search and Bayesian optimization will be employed to improve model performance. Ensemble learning can also be expanded by integrating other robust classifiers, aiming to build an even more accurate and adaptive fraud detection system for e-commerce platforms.

REFERENCES

[1] S. Monteith, M. Bauer, M. AIda, J. Geddes, P. C. Whybrow and T. Glenn, "Increasing cybercrime since the pandemic: Concerns for psychiatry", Curr. Psychiatry Rep., vol. 23, no. 4, pp. 18, 2021.

[2] S. Kodate, R. Chiba, S. Kimura and N. Masuda, "Detecting problematic transactions in a consumer-to-consumer e-commerce network", Appl. Netw. Sci., vol. 5, no. 1, pp. 90, 2020.

[3] R. Samani and G. Davis, McAfee mobile threat report, 2019, [online] Available: https://www.mcafee.com/enterprise/en-us/assets/reports/rp-mobile-threat-report-2019.pdf.

[4] E. W. T. Ngai, Y. Hu, Y. H. Wong, Y. Chen and X. Sun, "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature" in Decis. Support Syst., vol. 50, no. 3, pp. 559-569, 2011.

[5] "Online payment fraud: Market forecasts emerging threats & segment analysis 2022–2027", Sam Smith and Juniper Research, 2024, [online] Available: https://www.juniperresearch.com/press/losses-online-payment-fraud-exceed-362-billion/.

[6] A. Abdallah, M. A. Maarof and A. Zainal, "Fraud detection system: A survey", J. Netw. Comput. Appl., vol. 68, pp. 90-113, 2016.

[7] R. J. Bolton and D. J. Hand, "Statistical fraud detection: A review", Statistical Science, vol. 17, no. 3, pp. 235-255, 2002.

[8] C. Phua, V. Lee, K. Smith and R. Gayler, "A comprehensive survey of data mining-based fraud detection research", arXiv preprint, 2010.

[9] A. Ali, S. Abd Razak, S. H. Othman, T. A. E. Eisa, A. Al-Dhaqm, M. Nasser, et al., "Financial fraud detection based on machine learning: A systematic literature review", Appl. Sci., vol. 12, no. 19, pp. 9637, 2022.

[10] V. Rodrigues, L. Policarpo and D. E. da Silveira, Fraud detection and prevention in e-commerce: A systematic literature review, 2022, [online] Available: https://www.sciencedirect.com/science/article/pii/S1567422322000904?casa_token=UOjgVT_FXuwAAAAA:YgIpy5PUX5dEdF_dJ2NdlHz-664Vr32oHJPDq_ZbevxtOazQ38tP_I-PVDtKsCBFXXu_6-Ri6Q.

[11] I. Xournals, A review of credit card fraud detection techniques in e-commerce, 2022, [online] Available: https://www.academia.edu/39529497/A_review_of_Credit_card_Fraud_Detection_techniques_in_e_commerce.

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

[13] Y. Y. Festa and I. A. Vorobyev, "A hybrid machine learning framework for e-commerce fraud detection", Model Assist. Stat. Appl., vol. 17, no. 1, pp. 41-49, 2022.

[14] E. Ileberi, Y. Sun and Z. Wang, "A machine learning based credit card fraud detection using the GA algorithm for feature selection", J. Big Data, vol. 9, no. 1, pp. 24, 2022.

[15] M. H. Nasr, M. H. Farrag and M. M. Nasr, "A proposed fraud detection model based on e-Payments attributes a case study in Egyptian e-Payment gateway", Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 5, pp. 179-186, 2022.

[16] D. H. Lim and H. Ahn, "A study on fraud detection in the C2C used trade market using Doc2vec", J. Korea Soc. Comput. Inform., vol. 27, no. 3, pp. 173-182, 2022.

[17] G. Sasikala, M. Laavanya, B. Sathyasri, C. Supraja, V. Mahalakshmi, S. S. S. Mole, J. Mulerikkal, S. Chidambaranathan, C. Arvind, K. Srihari et al., "An innovative sensing machine learning technique to detect credit card frauds in wireless communications" in Wirel. Commun. Mob. Comput., vol. 2022, pp. 2439205, 2022.

[18] P. Verma and P. Tyagi, "Analysis of supervised machine learning algorithms in the context of fraud detection", ECS Trans., vol. 107, no. 1, pp. 7189-7200, 2022.

[19] A. Aziz and H. Ghous, "Fraudulent transactions detection in credit card by using data mining methods: A review", Int. J. Sci. Prog. Res., vol. 79, no. 1, pp. 31-48, 2021.

[20] K. S. Lim, L. H. Lee and Y. W. Sim, "A review of machine learning algorithms for fraud detection in credit card transaction", Int. J. Comput. Sci. Netw. Secur., vol. 21, no. 9, pp. 31-40, 2021.

[21] P. Gamini, S. T. Yerramsetti, G. D. Darapu, V. K. Pentakoti and P. R. Vegesena, "A review on the performance analysis of supervised and unsupervised algorithms in credit card fraud detection", Int. J. Res. Eng. Sci. Manag., vol. 4, no. 8, pp. 23-26, 2021.

[22] E. Ileberi, Y. Sun and Z. Wang, "Performance evaluation of machine learning methods for credit

card fraud detection using SMOTE and AdaBoost", IEEE Access, vol. 9, pp. 165286-165294, 2021.

[23] T. Pourhabibi, K. L. Ong, B. H. Kam and Y. L. Boo, "Fraud detection: A systematic literature review of graph-based anomaly detection approaches", Decis. Support Syst., vol. 133, pp. 113303, 2020.

[24] S. Lei, K. Xu, Y. Huang and X. Sha, "An Xgboost based system for financial fraud detection", E3S Web Conf., vol. 214, pp. 02042, 2020.

[25] Shriyash Jagtap, "Fraudulent E-Commerce Transactions dataset", Available at: https://www.kaggle.com/datasets/shriyashjagtap/fraudulent-e-commerce-transactions

Authors:

Mr. B. Amarnath Reddy is an Assistant Professor in the Department of Master of Computer Applications at QIS College of Engineering and Technology, Ongole, Andhra Pradesh. He earned his M.Tech from Vellore Institute of Technology(VIT), Vellore. His research interests include Machine Learning, Programming Languages. He is committed to advancing research and fostering innovation while mentoring students to excel in both academic and professional pursuits.