IJNRD.ORG ISSN : 2456-4184

INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG

An International Open Access, Peer-reviewed, Refereed Journal

ONLINE FRAUD DETECION IN BANKING DATA AND TRANSACTIONS USING ML

Jayesh Gulhane
Department of Information Technology, Prof. Ram Meghe Institute of Technology & Research, Amravati, Maharashtra,India

Yash Shinkar
Department of Information Technology, Prof. Ram Meghe Institute of Technology & Research, Amravati, Maharashtra, India

Chaitanya Gorle
Department of Information Technology, Prof. Ram Meghe Institute of Technology & Research, Amravati, Maharashtra, India

Padmaja Tayade
Department of Information Technology, Prof. Ram Meghe Institute of Technology & Research, Amravati, Maharashtra, India

Mentor
Prof. H. D. Misalkar
Associate Professor
Department of Information Technology, Prof. Ram Meghe Institute of Technology & Research, Amravati, Maharashtra, India

Abstract:-The growth of e-commerce has led to a surge in online credit card use and associated fraud. Banks face challenges in detecting fraud, making machine learning crucial for predicting fraudulent transactions. Various machine learning methods, especially the Random Forest algorithm, are used to identify patterns and anomalies indicative of fraud. By leveraging historical data and refining detection methods, banks can distinguish between legitimate and fraudulent activities. This project report focuses on optimizing fraud detection with machine learning, highlighting the importance of data preprocessing and domain-specific attributes for improved accuracy.

Keywords—Fraud Detection, Transaction Analysis, Random Forest Algorithm, Prevention Strategies, Feature Extraction, Credit Card Fraud, Suspicious Transactions, E-commerce Industry.

1. INTRODUCTION

The rapid expansion of the e-commerce sector has revolutionized the way we conduct transactions, offering unparalleled convenience and accessibility. With the increasing prevalence of online purchases using credit cards, however, comes a concurrent rise in fraudulent activities. This trend poses significant challenges for banks in detecting and preventing illicit credit card system activities. As fraudulent schemes evolve in complexity, traditional detection methods struggle to keep pace, prompting the need for advanced solutions.This research paper explores a critical aspect of modern finance: the detection of fraud in credit card transactions. Amidst the limitations of conventional approaches in combating evolving fraudulent tactics, the integration of machine learning emerges as a crucial strategy to fortify financial systems. Machine learning techniques hold the promise of accurate predictions and proactive identification of fraudulent activities, offering financial institutions the means to analyze transactional patterns and enhance predictive capabilities based on past data. In this contest, the study draws inspiration from advancements in machine learning and applies them to a practical scenario within the domain of medical e-commerce. By intricately monitoring user purchase behavior and leveraging machine learning models, the research endeavors to discern between genuine and fraudulent transactions. Through this framework, the aim is to mitigate risks associated with online transactions in the medical field, showcasing the transformative potential of machine learning across diverse domains.

As the financial landscape continues to evolve, the symbiotic relationship between machine learning and fraud detection emerges as a cornerstone in preserving the integrity of digital transactions. This research paper contributes to ongoing discourse by proposing a tangible application that amalgamates technology and healthcare, thereby exemplifying the multifaceted potential of advanced data-driven solutions in modern financial security. Furthermore, the paper outlines specific objectives aimed at developing proactive fraud detection strategies, contributing to technological innovation, integrating machine learning in finance, and providing enhanced security in e-commerce. Through experimentation with real-world data, the efficacy of proposed methods will be evaluated, contributing to the advancement of fraud detection practices in financial and e-commerce sectors.

2. LITERATURE SURVEY

"BLAST-SSAHA Hybridization for Credit Card Fraud Detection" by A. Kundu, S. Panigrahi, S. Sural, and A. K. Majumdar this paper examines the surge in credit card transactions, especially in online shopping, leading to an increase in fraudulent activities. The study emphasizes the importance of efficient fraud detection systems for credit card-issuing banks to curb losses. The authors propose a two-phase sequence alignment method, starting with a profile analyzer (PA) that compares incoming transaction sequences on a credit card to the cardholder's previous spending patterns. Transactions flagged as unusual are then sent to a deviation analyzer (DA) for further alignment with historical fraud patterns [1].

"Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms" by F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan, and M. Ahmed. This paper discusses the growing reliance on credit cards for online transactions due to their convenience. This increased usage has also heightened the risk of credit card fraud, causing significant financial losses for cardholders and financial institutions. The study aims to identify fraud by addressing challenges such as data availability, data imbalance, the evolving nature of fraud, and high false alarm rates. The authors review various machine learning-based methods for credit card fraud detection, including Extreme Learning Method, Decision Tree, Random Forest, Support Vector Machine, Logistic Regression, and XGBoost [2].

"Fraud Detection in Banking Data by Machine Learning Techniques" by S. K. Hashemi, S. L. Mirtaheri, and S. Greco. This paper highlights the expansion of advanced technology and ecommerce services, which has made credit cards a popular payment method and led to a higher volume of banking transactions. This trend, coupled with the substantial rise in fraud, has increased the costs of banking transactions and the importance of detecting fraudulent activities. The authors propose the use of class weight-tuning hyperparameters as a method to manage the weight of both fraudulent and legitimate transactions to improve the effectiveness of fraud detection [3].

"Credit Card Fraud Detection Based on Improved Variational Autoencoder Generative Adversarial Network" by Y. Ding, W. Kang, J. Feng, B. Peng, and A. Yang, this paper proposes the examines the surge in mobile banking and e-commerce, which has led to a significant rise in fraudulent online payments. Despite the common use of machine learning and deep learning for credit card fraud detection, the imbalance in credit card transaction datasets—where fraudulent data is much smaller than normal transaction data—restricts the efficacy of conventional binary classification methods. To address this challenge, researchers often oversample the minority class data and employ ensemble learning classification algorithms, though oversampling presents its own set of drawbacks [4].

"Anomaly Detection in Graphs of Bank Transactions for Anti Money Laundering Applications" by B. Dumitrescu, A. Băltoiu, and S. Budulan, this paper focuses on identifying bank clients involved in suspicious activities related to money laundering by examining the bank's transaction graph. Although the study uses a real dataset with labels, the goal is to achieve significant results not only on the real data but also on random graphs where common anomaly patterns have been introduced. The method aims for both data suitability and robustness. It involves creating new features, particularly those derived from the reduced egonet, which is a subgraph remaining after removing nodes connected to the center by a single edge. Another feature is developed through random walks and acts as an indicator of circular flows [5].

"Credit Card Fraud Detection using Apache Spark Analysis" by N. Sethumadhavan and H. Narayanan AG, this paper examines the persistent threat of credit card fraud within the financial industry. The increasing dependence on internet technology has led to a rise in fraudulent transactions, particularly impacting the banking sector due to the extensive use of credit cards and online banking. Various methods have been proposed to address this challenge, with the paper specifically focusing on identifying fraudulent transactions through the analysis of historical transaction records. The authors employ Apache Spark for this analysis, which allows efficient processing and handling of large datasets for fraud detection [6].

"Credit Card Fraud Identification Based on Unbalanced Data Set and Fusion Model" by D. Li discusses the significant challenges banks face due to the rapid growth in credit card business, particularly the increased threat of credit card fraud. To conduct credit card operations effectively and minimize fraud risks, the paper reviews global and domestic research findings, focusing on balancing model availability, security, and timeliness.In addressing the unbalanced data set, the study employs machine learning methods and proposes three types of anti-balance fraud models. The initial step involves data preprocessing through an undersampling method. Subsequently, models such as Lasso-Logistic and XGBoost are applied to model and predict fraudulent activities. By combining these approaches, the study aims to enhance the detection and prevention of credit card fraud, ultimately providing banks with a more robust strategy to safeguard their credit card business [7].

"E-commerce Merchant Fraud Detection using Machine Learning Approach" by F. Hasan, S. K. Mondal, M. R. Kabir, M. A. Al Mamun, N. S. Rahman, and M. S. Hossen this paper examines the rise of e-commerce as a global phenomenon, noting the parallel increase in fraudulent promotional activities by dishonest marketers seeking to boost sales. These marketers manipulate outcomes through deceptive means, such as fake travel and shopping schemes. The study addresses the issue of deception on major commerce platforms by identifying merchants who have previously engaged in fraud and maintaining a list of such individuals. The authors employ a machine learning approach to train models that can detect whether a given merchant ID is fraudulent or not. This research aims to highlight the importance of safeguarding e-commerce platforms against fraud to protect both merchants and consumers [8].

"Real-time Credit Card Fraud Detection Using Machine Learning" by A. Thennakoon, C. Bhagyani, S. Premadasa, S. Mihiranga, and N. Kuruwitaarachchi, this paper investigates the frequent occurrence of credit card fraud and the consequent financial losses. The growth in online transactions, particularly online credit card transactions, has made fraud detection applications increasingly valuable and sought after by banks and financial institutions. Fraudulent transactions manifest in various forms and can be classified into different categories. The study concentrates on four primary types of fraud in real-world transactions, each of which is addressed using a series of machine learning models. An evaluation process is employed to determine the most effective method for each fraud type, providing a comprehensive guide for choosing the optimal algorithm. This evaluation is supported by relevant performance measures. Additionally, the paper emphasizes real-time credit card fraud detection as a crucial aspect of their research, aiming to provide immediate and efficient identification of fraudulent activities [9].

"Comparative Analysis of Credit Card Fraud Detection using Logistic Regression with Random Forest towards an Increase in Accuracy of Prediction" by M. V. Krishna and J. Praveenchandar, this paper seeks to identify fraud committed using various payment cards, including credit and debit cards. The study experiments with two algorithms—Random Forest and Logistic Regression—to find the most suitable one for detecting fraud. In the methodology, both algorithms are applied using supervised learning to glean insights from historical data [10].

"Ensemble Synthesized Minority Oversampling-Based Generative Adversarial Networks and Random Forest Algorithm for Credit Card Fraud Detection" by F. A. Ghaleb, F. Saeed, M. Al-Sarem, S. N. Qasem, and T. Al-Hadhrami, this paper discusses the severe nature of financial fraud, such as money laundering, which is a serious criminal activity that funds terrorism and other crimes. These illegal operations often involve complex transaction networks, making fraud detection a challenging task. By examining trading and transaction networks along with entity features, suspicious behavior and potential fraud can be identified. The paper introduces CoDetect, a novel framework that leverages both network and feature information to enhance fraud detection performance. This approach aims to overcome the limitation of most existing methods, which typically address either network or feature data separately and miss the opportunity to integrate both types for more effective detection. Experiments conducted on both synthetic and real-world data show the effectiveness of CoDetect in combating financial fraud, with a focus on money laundering [11].

"An Adversary Model of Fraudsters' Behavior to Improve Oversampling in Credit Card Fraud Detection" by D. Lunghi, G. M. Paldino, O. Caelen, and G. Bontempi, this paper explores the challenges of imbalanced learning in credit card fraud detection, focusing on the minority class. The authors introduce an innovative approach, "Adversary-based Oversampling" (ADVO), which views fraudulent behavior as a time-dependent process and models temporal relationships among frauds. The strategy incorporates regression-based oversampling to predict future fraudulent activities and adapts the TimeGAN algorithm for generating synthetic fraud time series. Experiments with real and synthetic data from Worldline S.A. demonstrate ADVO's effectiveness in improving fraud detection performance [12].

"An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine" by A. A. Taha and S. J. Malebary, this paper discusses the increasing use of credit cards for both regular and online purchases due to advancements in e-commerce and communication technologies. This rise in credit card usage has led to a significant increase in fraud, causing substantial financial losses for businesses and consumers. Fraud detection has become critical to ensure the continued use and safety of electronic payments.The authors present an intelligent approach for detecting fraud in credit card transactions using an optimized light gradient boosting machine (OLightGBM) [13].

"Online Payment Fraud Detection Model Using Machine Learning Techniques" by A. A. Almazroi and N. Ayub, this paper discusses the increasing threat of financial fraud in the context of wireless communications. The authors introduce a real-time processing model for financial transaction data using a ResNeXt embedded Gated Recurrent Unit (GRU) model (RXT) to combat financial fraud. The model begins with AI data input and preprocessing, handling data imbalance with SMOTE. It uses an AI ensemble method combining autoencoders and ResNet (EARN) for feature extraction and fine-tunes the RXT model with the Jaya optimization algorithm (RXT-J) for the core classification task. The AI model is evaluated on three financial transaction datasets, consistently outperforming existing algorithms by 10% to 18% across various metrics while maintaining excellent computational efficiency. This research enhances security and efficiency in financial transactions [14].

"E-commerce fraud detection through fraud islands and multilayer machine learning model" by J. Nanduri, Y.-W. Liu, K. Yang, and Y. Jia, this paper specially addresses the challenge of diverse and dynamic fraud patterns in e-commerce transactions. The authors introduce two innovative methods: fraud islands (link analysis) and a multi-layer machine learning model.Fraud islands use link analysis to explore relationships between different fraudulent entities and reveal hidden complex fraud patterns. The multi-layer model handles the varied nature of fraud patterns by incorporating different types of fraud labels from various channels, such as banks' decisions, manual review agents' rejections, and customer chargeback requests. By combining different machine learning models trained with different fraud labels, the study demonstrates significant improvements in the accuracy of fraud detection decisions [15].

3. PROBLEM STATEMENT AND PROPOSED METHODLOGY

In this paper, we proposed online fraud detection system using machine learning algorithm. Fraud can manifest through various means, including theft of credit card details, identity impersonation, or the creation of deceptive websites aimed at soliciting personal information. Financial institutions and businesses are actively engaged in combating fraud by employing sophisticated technologies such as encryption algorithms and machine learning algorithms to detect fraudulent activities. These technologies enable the automatic identification of counterfeit transactions and patterns indicative of fraudulent behavior. Vigilance and prompt reporting of suspicious activities are crucial in safeguarding personal finances and thwarting deceptive schemes. By promoting awareness and proactive engagement in fraud detection efforts, individuals can play a pivotal role in preserving financial security and preventing monetary losses.

In traditional online banking systems, One-Time Passwords (OTPs) are commonly utilized as a security measure to thwart fraudulent transactions. However, if unauthorized individuals obtain access to the OTP, preventing unauthorized transactions becomes significantly challenging. To tackle this issue, we propose a novel approach: a fraud detection system that continuously monitors various transaction details, including OTP authentication time, attempted transactions, client-side IP address, MAC ID, total number of daily transactions, and transaction amounts. The primary objective of our research paper is to develop a robust fraud detection and prevention system by employing machine learning methodologies, with a particular focus on implementing the Random Forest algorithm for fraud detection purposes.

Upon identifying suspicious transactions, our system initiates an identity verification process. This process entails users undergoing a security challenge to authenticate their identity. Successful completion of this verification process confirms the transaction as genuine, allowing it to proceed without interruption. However, failure to authenticate within the specified time frame triggers proactive measures to mitigate potential risks In cases where fraudulent activity is suspected:

Present a comprehensive analysis of the effectiveness and efficiency of the proposed fraud detection and prevention system, showcasing its potential impact in mitigating risk associated with online banking transactions.

Below is an overview of user activities, administrative tasks, system modules and workflow.

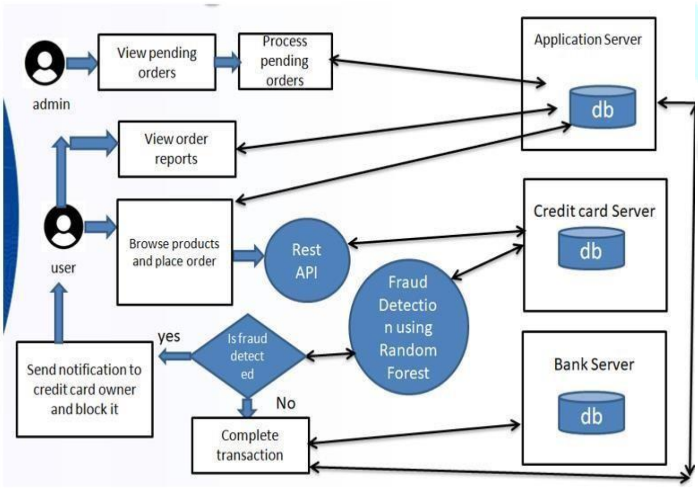


Fig.1:- System Architecture

- **User Activities : Registration**

Introduction:

User registration is fundamental in any online system, establishing a secure identity for individuals within the platform. In the credit card transaction processing system, registration is crucial for ensuring secure access and utilization of system functionalities.

Purpose of Registration:

The primary aim of user registration is to create a secure and personalized environment within the system. By capturing essential user information during registration, unique user identities are established, allowing for tailored user experiences. Additionally, registration facilitates the implementation of robust security measures to protect user data and transactions from unauthorized access and fraudulent activities.

Registration Process:

The registration process typically comprises several steps to gather necessary user information and establish user accounts securely. Users complete a registration form, providing personal details such as name, email address, contact number, and billing information. Account credentials, including a username and password, are generated during this process. The system validates entered data for accuracy and completeness before securely storing it in the database. Upon successful registration, users gain access to system functionalities.

Data Security and Privacy:

Ensuring the security and privacy of user data is paramount during registration. The system employs encryption protocols and secure data transmission methods to safeguard sensitive user information. Stringent privacy policies regulate the collection, storage, and use of user data, ensuring compliance with regulatory standards and protecting user privacy rights. Additional security measures such as multi-factor authentication and CAPTCHA verification may be implemented to prevent unauthorized access to user accounts.

- **Admin Activities :**

Introduction:

Administrators play a critical role in managing and overseeing credit card transaction processing systems. They are provided with dedicated access to an admin panel, facilitating effective management of system functionalities.

Purpose of Admin Panel Access:

The admin panel serves as a centralized hub for administrators to monitor, manage, and analyze various aspects of the system. Through the admin panel, administrators can perform tasks such as viewing pending orders, processing transactions, managing user accounts, and generating reports. This access enables administrators to make informed decisions, address issues promptly, and ensure the smooth operation of the system.

- **Algorithm Used:**

Random Forest :

In the perpetual battle against fraudsters within the financial sector, machine learning algorithms stand as formidable allies, with Random Forest emerging as a key contender due to its efficacy and adaptability. This section provides an in-depth exploration of the Random Forest algorithm, elucidating its inner workings, advantages, and applicability in detecting fraudulent activities within banking transactions.

Random Forest constitutes a supervised machine learning algorithm rooted in ensemble learning principles. Ensemble learning involves amalgamating various algorithms, or iterations of the same algorithm, to construct a more robust prediction model. In the case of Random Forest, multiple decision trees are combined, forming a "forest" of trees, hence the name. This algorithm is versatile, catering to both regression and classification tasks.

Random Forest holds significance as a widely employed machine learning algorithm within the realm of supervised learning. It is adept at handling both Classification and Regression challenges in machine learning. The underlying concept of ensemble learning underpins its efficacy, allowing for the aggregation of multiple classifiers to address intricate problems and enhance model performance.

The following diagram illustrates the functioning of the Random Forest algorithm:

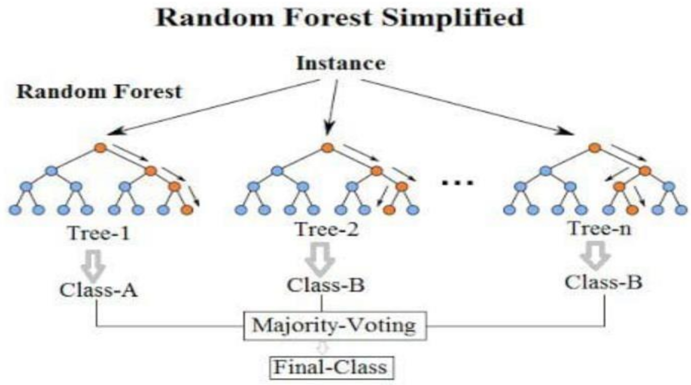


Fig.2 :- Working of the Random Forest Algorithm

Random Forests are trained using the bagging method, also known as Bootstrap Aggregating. This technique involves randomly sampling subsets of the training data, fitting a model to each subset, and aggregating the predictions. By allowing instances to be repeatedly used for training with replacement, bagging reduces the variance of the resulting models. In the context of Random Forests, this method extends to the feature space as well, where subsets of features are sampled randomly, adding diversity and reducing variance at the expense of potentially higher bias. This process, known as "feature bagging," contributes to the robustness of the model. When making predictions with Random Forests, each new data point traverses multiple trees in the ensemble, each grown using random samples of both training data and features. For classification tasks, the mode or most frequent class predicted by individual trees (majority vote) is used for aggregation, while for regression tasks, the average prediction of each tree is employed. Despite its effectiveness, it's essential to cross-validate the Random Forest model to guard against overfitting. Additionally, the algorithm's training stage can be slow due to the need to grow multiple trees, a process that is inherently greedy.

Random Forests are particularly suitable for classification and regression tasks, offering simplicity, flexibility, and high accuracy. Each tree in the forest relies on a random and unbiased dataset, with all trees exhibiting similar dispersion. As the forest grows, it provides an internal, unbiased estimate of the generalization error. The workflow of the Random Forest model involves evaluating various criteria across multiple decision trees. Random datasets are used to train the model, and each decision tree provides likelihood results for classifying transactions as 'fraudulent' or 'legitimate.' By aggregating these results, the final prediction is made. Various transaction parameters, such as card number, location, date, time, IP address, amount, frequency, and age, are evaluated by the fraud detection algorithm to identify variables that aid in splitting the data.

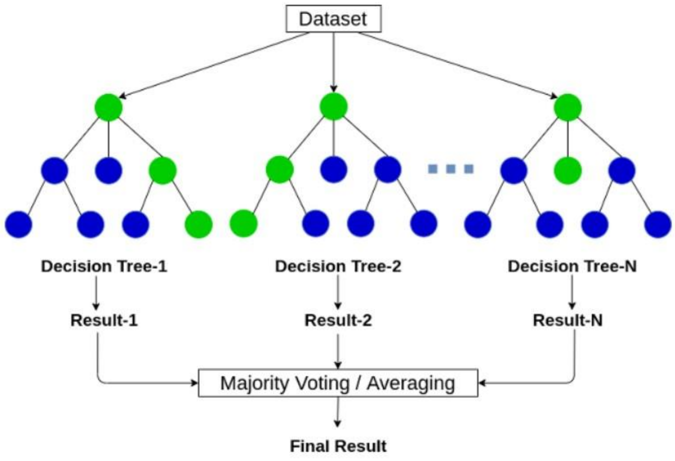


Fig.3:- Random Forest Algorithm

• **WORKING OF RANDOM FOREST :**

The following are the fundamental steps involved in executing the random forest algorithm :

1. Select N random records from the dataset.
2. Construct a decision tree based on these N records.
3. Determine the number of trees desired in your algorithm and repeat steps 1 and 2.
4. For classification problems, each tree in the forest predicts the category to which the new record belongs. Subsequently, the new record is assigned to the category that garners the majority vote.

4. RESULT

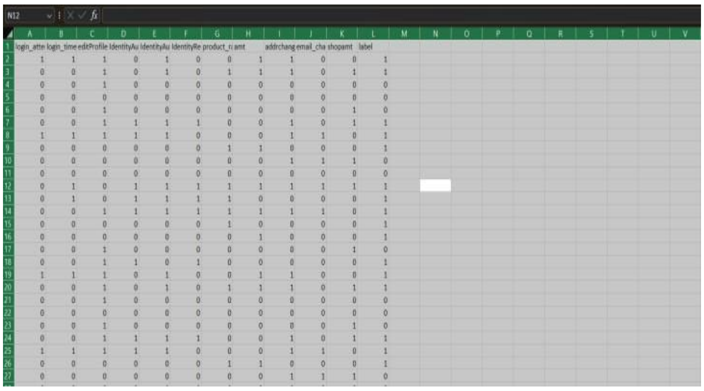
This section presents the outcomes derived from the executed system, demonstrating snapshots to elucidate the functionality and features of our system effectively.

• **Dataset Description:**

The dataset encompasses user activities and transaction details sourced from a banking system, encompassing login attempts, profile modifications, identity verification, transaction initiations, and other pertinent interactions.

It comprises both categorical attributes (e.g., login time, profile edits) and numerical features (e.g., login attempts, transaction amounts).

The label column delineates the classification of each transaction as either fraudulent or legitimate.



	login_atte	login_time	editProfile	IdentityAu	IdentityAu	IdentityRe	product_n	amt	addrchang	email_cha	shopamt	label
2	1	1	1	0	1	0	0	1	1	0	0	1
3	0	0	1	0	1	0	1	1	1	0	1	1
4	0	0	1	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	1	0	0	0	0	0	0	0	1	0
7	0	0	1	1	1	1	0	0	1	0	1	1
8	1	1	1	1	1	0	0	0	1	1	0	1
9	0	0	0	0	0	0	1	1	0	0	0	1
10	0	0	0	0	0	0	0	0	1	1	1	0
11	0	0	0	0	0	0	0	0	0	0	0	0
12	0	1	0	1	1	1	1	1	1	1	1	1
13	0	1	0	1	1	1	1	0	0	0	0	1
14	0	0	1	1	1	1	1	1	1	1	0	1
15	0	0	0	0	0	0	1	0	0	0	0	1
16	0	0	0	0	0	0	0	1	0	0	0	1
17	0	0	1	0	0	0	0	0	0	0	1	0
18	0	0	1	1	0	1	0	0	0	0	0	1
19	1	1	1	0	1	0	0	1	1	0	0	1
20	0	0	1	0	1	0	1	1	1	0	1	1
21	0	0	1	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	1	0	0	0	0	0	0	0	1	0
24	0	0	1	1	1	1	0	0	1	0	1	1
25	1	1	1	1	1	0	0	0	1	1	0	1
26	0	0	0	0	0	0	1	1	0	0	0	1
27	0	0	0	0	0	0	0	0	1	1	1	0

Screenshot 1:-Dataset of User Activities

- **Preprocessing Steps:**

➢ Handling Missing Values : Identify and address any missing values in the dataset. This may involve techniques such as imputation (replacing missing values with calculated estimates) or removal of rows with missing data.

➢ Feature Scaling : Normalize numerical features to ensure they have a similar scale. Common scaling techniques include standardization (mean normalization) or minmax scaling.

➢ Encoding Categorical Variables : Convert categorical variables into a numerical format that machine learning algorithms can interpret. This can be achieved through methods like one-hot encoding, which creates binary columns for each category.

➢ Dealing with Class Imbalance : Address any class imbalance issues in the dataset, where the number of fraudulent transactions may be significantly lower than legitimate transactions. Techniques such as oversampling (creating synthetic instances of minority class) or undersampling (reducing instances of majority class) can be employed to balance the classes.

- **Data Splitting:**

➢ Train-Test Split: Divide the dataset into separate training and testing sets. The training set is used to train the machine learning models, while the testing set is used to evaluate their performance**.**

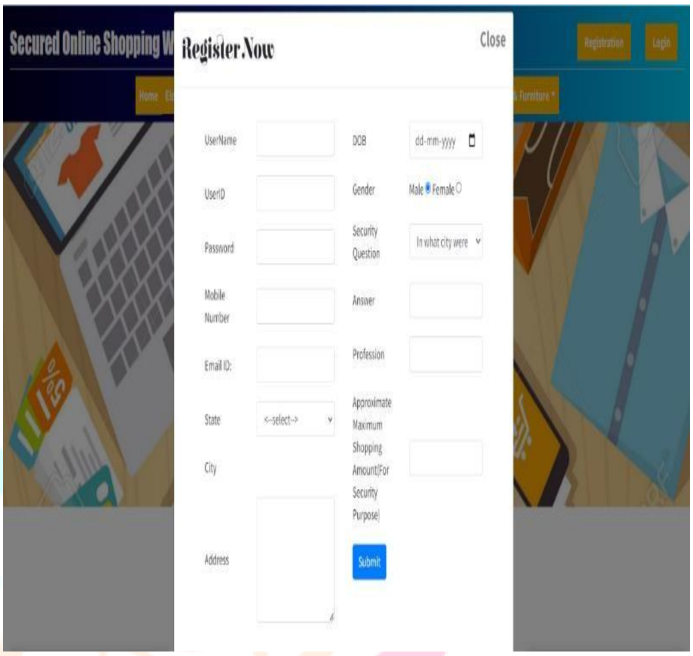
➢ Stratified Sampling: Ensure that the distribution of fraudulent and legitimate transactions is maintained in both the training and testing sets. This helps prevent bias and ensures representative samples for model evaluation.

➢ Cross-Validation (Optional): Consider employing techniques like k-fold cross-validation to further validate model performance. This involves splitting the dataset into multiple folds, training the model on different combinations of training and validation sets, and averaging the results.

- **Experimental Results**

User Activities: Registration: Incorporate a snapshot of the registration process interface or output from your system. Emphasize essential metrics such as registration success rate, average registration time, and any anomalies detected during registration.

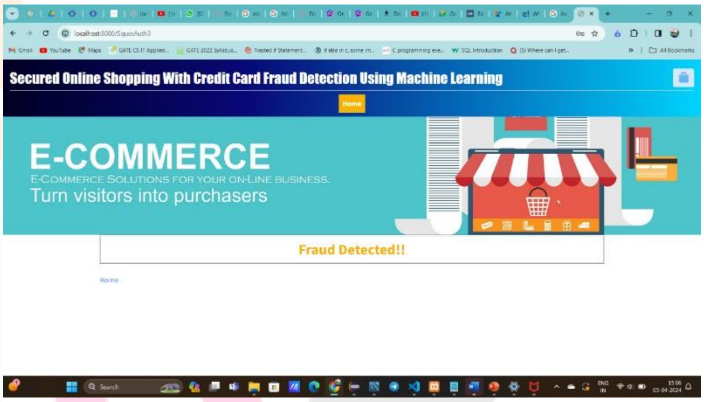
Analysis: Interpret the findings, addressing any obstacles faced during user registration and how the system adeptly manages them. Assess the effectiveness and user-friendliness of the registration process**.**



Screenshot 2:- Registration Process

- **Results Snapshot:**

Below is a snapshot of the outcomes derived from the fraud detection analysis:



Fraud Detected!!

Screenshot 3:- of Fraud Detection

5. CONCLUSION

Our project focuses on the crucial task of fraud detection and prevention in online e-commerce transactions. We have developed an innovative system that integrates robust fraud detection and prevention models. This includes a comprehensive set of attributes to detect fraudulent activities before transactions are completed. Our solution is anchored by an ecommerce web application that meticulously tracks user activities in a centralized database. During transaction processing, we seamlessly transmit behavioral details to the credit card server, where our advanced fraud detection model, leveraging the Random Forest algorithm, discerns whether transactions are legitimate or fraudulent. In summary, our project represents a major advancement in credit card fraud detection. By harnessing machine learning and realtime transaction monitoring, we have established a robust system to protect users from fraudulent activities. Through early detection and rapid intervention, we aim to contribute to a safer and more secure online shopping environment, ultimately minimizing financial losses and providing peace of mind for our customers.

6. REFERENCES

[1] Kundu, A., Panigrahi, S., Sural, S., & Majumdar, A. K. (2009). "BLAST-SSAHA Hybridization for Credit Card Fraud Detection." Published in vol. 6, no. 4, pp. 309-315, Oct.-Dec. 2009

[2] Alarfaj, F. K., Malik, I., Khan, H. U., Almusallam, N., Ramzan, M., & Ahmed, M. (2022). "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms." Published in IEEE Access, vol. 10, pp. 3970039715, 2022.

[3] Hashemi, S. K., Mirtaheri, S. L., & Greco, S. (2023). "Fraud Detection in Banking Data by Machine Learning Techniques."

Published in IEEE Access, vol. 11, pp. 3034-3043, 2023.

[4] Ding, Y., Kang, W., Feng, J., Peng, B., & Yang, A. (2023). "Credit Card Fraud Detection Based on Improved Variational Autoencoder Generative Adversarial Network." Published in

IEEE Access, vol. 11, pp. 83680-83691, 2023.

[5] Dumitrescu, B., Băltoiu, A., & Budulan, Ş. (2022). "Anomaly Detection in Graphs of Bank Transactions for Anti Money Laundering Applications." Published in IEEE Access, vol. 10, pp.

47699-47714, 2022.

[6] A. S., Sethumadhavan, N., & Narayanan AG, H. (2021). "Credit Card Fraud Detection using Apache Spark Analysis." Presented at the 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, pp. 998-100.

[7] D, Li. (2019). "Credit card fraud identification based on unbalanced data set based on fusion model." Presented at the 2019 IEEE 1st International Conference on Civil Aviation Safety and Information Technology (ICCASIT), Kunming, China, pp. 235-239.

[8] Hasan, F., Mondal, S. K., Kabir, M. R., Al Mamun, M. A., Rahman, N. S., & Hossen, M. S. (2022). "E-commerce Merchant Fraud Detection using Machine Learning Approach." Presented at the 2022 7th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, pp. 1123-1127.

[9] Thennakoon, A., Bhagyani, C., Premadasa, S., Mihiranga, S., & Kuruwitaarachchi, N. (2019). "Real-time Credit Card Fraud Detection Using Machine Learning." Presented at the 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, pp. 488-493.

[10] Krishna, M. V., & Praveenchandar, J. (2022). "Comparative Analysis of Credit Card Fraud Detection using Logistic regression with Random Forest towards an Increase in Accuracy of Prediction." Presented at the 2022 International Conference on Edge Computing and Applications (ICECAA), Tamilnadu, India, pp. 1097-1101.

[11] Ghaleb, F. A., Saeed, F., Al-Sarem, M., Qasem, S. N., & Al-

Hadhrami, T. (2023). "Ensemble Synthesized Minority

OversamplingBased Generative Adversarial Networks and Random

Forest Algorithm for Credit Card Fraud Detection." Published in

IEEE Access, vol. 11, pp. 89694-89710, 2023.

Hashemi, S. K., Mirtaheri, S. L., & Greco, S. (2023). "Fraud

Detection in Banking Data by Machine Learning

Techniques."Published in IEEE Access, vol. 11, pp. 3034-

3034,2023.

[12] Lunghi, D., Paldino, G. M., Caelen, O., & Bontempi, G. (2023). "An Adversary Model of Fraudsters' Behavior to Improve Oversampling in Credit Card Fraud Detection." Published in IEEE Access, vol. 11, pp. 136666-136679, 2023.

[13] Taha, A. A., & Malebary, S. J. (2020). "An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine." Published in IEEE Access, vol. 8, pp. 25579-25587, 2020

[14] Almazroi, A. A., & Ayub, N. (2023). "Online Payment Fraud Detection Model Using Machine Learning Techniques." Published in IEEE Access, vol. 11, pp. 137188-137203, 2023.

[15] Nanduri, J., Liu, Y.-W., Yang, K., & Jia, Y. (2020). "Ecommerce fraud detection through fraud islands and multi-layer machine learning model." Presented at the Proc. Future Inf. Commun. Conf., pp. 556-570, 2020