Detecting E-Commerce Fraud in Real Time with Machine learning

Mrs.Sowmya V[2] , Manoj Kengalagutti[1]

[2]Assistant Professor, Department of MCA, RVITM, Bengaluru
[1] Student,4th Semester MCA, Department of MCA, RVITM, Benagluru

Abstract

The rapid rise of online transactions spearheaded by e-commerce have restructured the global marketplace. Online transactions represent the highest level of convenience and efficiency, but have also posed a continuous threat of fraudulent activities that impact both customers and companies, and it is important to recognize that the rapid growth of e-commerce presents a problem of complexity predisposed to technological solutions to assist with amelioration. In this paper, we detail a systematic approach to engaging in fraud detection across many participants in an e-commerce transaction. The plan is based on building and validating several machine learning models to examine complex transactional data that considers multi-faceted dimensions. The classification machine learning models created include Random Forests, Support Vector Machines (SVM), Naive Bayes, Logistic Regression, and Gradient Boosting classifiers.Following the modeling, the classification schemes were unified in an experiment to examine each models' accuracy in predicting between fraudulent and legitimate transactions. Our experience supports the precision regarding the random forest classifier (97.06% accurately classified). In this experiment we possessed a sizable dataset of transactional attributes, customer device and IP address details. The dataset was preprocessed utilizing techniques like standard scaling and one-hot encoding. Evaluation of models considered measures of Recall, accuracy, precision and F1 score on each models' predictive abilities. Finally, we integrated the Random Forests classifier into a web application based on Flask

***Keywords*—E-commerce, Fraud Detection, Class Imbalance, Random Forest, SVM, logical regression, gradient boosting, navies bayes, flask app, machine learning**

I INTRODUCTION

A. Background

The environment surrounding electronic commerce (e-commerce) has changed to be radically different and has created new and unprecedented levels of convenience and accessibility for consumers around the world. The way businesses operate has been drastically altered by the digital transformation of commerce. operate, allowed organizations to interact globally and possesses created it easy for millions of businesses (large and small) to possess a presence on the Internet. While this progress has received immense focus and praise, it has a looming dark feature that influences and impacts businesses and consumers alike, the increased risk of fraud.1 E-commerce fraud is defined as a fraudulent transaction that encompasses numerous deceptive strategies to accomplish the task. Major abuses include credit card fraud, identity theft, account takeovers, and different types of credit cards and other account fraud while in cyberspace rapid escalation for security considerations.

B. Objectives

The primary objective of the study was to design the online transactions has led to a staggering increase in digital fraud and a consequential arms race between fraudsters and fraud detection, thereby necessitating the need for strong cyber security regimes and anti-fraud solutions that are proactive, continuously learning, and not just reactive. Traditional fraud-detection modes and systems are subject to rigid restrictions that impact their effectiveness, as they often depend on human-aided review processes based on pre-existing rules. These fraud-detection systems are reactionary, struggle with high false-positive rates, and embed time-consuming processes and modelling. They simply cannot adapt to changes in fraud schemes. In addition, the ease of detecting fraud in a physical transactional environment relies on easily observable physical cues, while the rise of online

C. LITERATURE REVIEW

Mutemi & Bacao (2024) carried out a systematic literature review of e-commerce fraud detection methods, employing real and semi-synthetic datasets to test a multiview graph clustering algorithm. They used a heterogeneous information graph to model the e-commerce network, generating user-similarity graphs with varying metapaths and converting them to embeddings for clustering. The study concluded that the proposed method successfully detected fraudulent activities and outperformed the standard methods.

Tang (2023) tested an automatic framework to use Artificial Neural Networks (ANNs) and Deep Reinforcement Learning (DRL) for e-commerce fraud detection. Using a new dataset from Boyner Group e-commerce and mobile app, the authors compared against decision tree, logistic regression, random forest and extreme gradient boosting algorithms. The abstract did not specifically state their conclusions but the authors discussed the investigation of feature

Srinivasa Raju & Varma (2025) developed a fraud detection system at the merchant-level derived from merchant-level data including merchant ID's and prior fraud data. An implementation of random forests, decision trees, and logistic regression were performed to detect fraud in merchants. Conclusions suggested that these algorithms show some indicating ability to detect fraud in merchants and provide important work including generalizability testing on larger datasets, comparisons across models, and performance.

Kakde et al. (2025) provided a literature review of counterfeit detection in e-commerce using machine learning. They claimed that limitations on real-world implementation presented challenges, proposed practical algorithms to implement, but ultimately did not included details on the dataset, conclusions, or possible future directions. Although there has been some progress made between both studies in that research centers are looking to utilize novel, contemporary machine learning techniques in the area of e-commerce fraud detection, specifically harnessing the power of ensemble models, graph-based clustering, deep reinforcement learning—there remain limitations in dataset information on transparency for comparison, explicit conclusions, and future directions to continue the research in a more exploratory manner.

A. Identified *gaps*

This paper identifies an important issue - the constant danger of fraud associated with e-commerce transactions. Although technologies continue to advance, digital payment systems remain vulnerable to increasingly virulent forms of fraud, such as fraudulent transactions, identity theft, and payment fraud. These fraudulent attempts generate millions of dollars in losses for businesses and, more importantly, consumers, based on the kinds of activities, undermine the trust (which is essential for the sustained Development of the digital market. The development of a fraud is the root of the fraud problem detection system that is robust and fluid, which is capable of identifying fraudulent transactions correctly in a vast amount of complex data in multi-participant e-commerce contexts. This complexity makes it hard for humans to make correct decisions about fraud detection in multi-participant e-commerce contexts. It is difficult for organizations to identify fraud in multi-participant e-commerce due to this complexity. There is not just a customer and a merchant, like an online retailer, but also payment processors. All of this exists through a dynamic ecosystem that is ever changing through the Internet of Things and new expectations in the transaction that it is seamless and trustworthy. It occurs in a constantly changing conditions all combine together through complex interactions, and fraud detection is a complex set of complex interactions. The many participants in the multi-participant e-commerce ecosystem cannot be properly modeled using conventional rule-based fraud detection systems – despite their advantages. When a complex system has many participants that converge and layer together using complex businesses and as the interactions become increasingly complex, the traditional, if not outdated, rule-based systems are inadequate for modeling the complexities of the interactions. Fraud detection needs a more sophisticated capability of pattern recognition through machine learning safe in the understanding of recognizing not so obvious relationships in high-dimensional data. It is very important to address this problem, because it can be used to provide greater protection to businesses effects of fraud in their prepared reliance on transaction reliability and trustworthiness. The main focus of this study is to develop and be able to evaluate a fraud detection.

we created a framework to allow multi-participant e-commerce transactions in the use of various machine. metrics offer a more refined and accurate analysis of how well a model can accurately identify the rare instance (fraudulent transactions) with less false positives.

I. METHODOLOGY

A. Approach

The design of the e-commerce machine learning model for fraud detection Python, and Flask will involve a stepwise approach. This approach involves first collecting and pre-processing the dataset which includes transactional, behavioral, and demographic characteristics of customers. Once collected and pre-processed, the data must be cleaned, normalized, and features must be extracted from the data, which could potentially represent the presence of fraud, such as unusual purchase amounts, device anomalies, or location mismatches. Once the data is prepared for analysis, machine learning techniques like support vector machine and random forest, Naïve Bayes, Logistic Regression, and Gradient Boosting are applied to train models which can make predictions regarding transaction activity as being fraudulent or originating from a legitimate use. Each model will be evaluated based upon one or multiple accuracy, precision, recall, F1-score, AUC-ROC metrics etc. to assess the robustness of the model for real-time rapid detection. The models selected through one or multiple of these processes will be built into a Flask based web application to allow users to upload transaction data and receive fraud detection specifications in real-time. Overall, the proposed framework is relative and useful due to the scalability of Python for data resources, the reliability of supervised ML algorithms, and the interactivity of Flask for providing mathematical fraud detection results for users across multiple devices/platforms.

B. Data Collection and Preparation

The first step is to download a full dataset from Kaggle containing transactional data that is capable of e- commerce fraud detection. After downloading the raw data, it went through extensive preprocessing to prepare it to be ready and structure it properly for machine learning models. The preprocessing included cleaning the dataset, normalizing the numerical features (e.g., via the Standard Scaler), encoding the categorical features (e.g., via a One Hot Encoder so they could be turned into numerical values usable by the algorithms), etc. A significant step of the data preparation for fraud detection is the consideration of class imbalance and the more significant challenge of imbalanced data. Legitimate users far outnumber fraudulent ones in this type of dataset.1 Although the algorithms we selected are considered ensemble methods, and even though algorithms (like Random Forest and Gradient Boosting) can offer some inherent resistance to class imbalance, in practice area, it is usually "best practice" to apply other class-balancing techniques, such as using Synthetic Minority Over-sampling Technique (SMOTE), or Adaptive Synthetic Sampling (ADASYN), to balance the dataset and avoid bias in the model carrying over to the majority of the class.

Algorithm Selection

A variety of machine learning methods that are especially well suited for tasks involving fraud detection were implementation and comparison.

Random Forest: An ensemble learning method known for its classification performance, and ability to handle high-dimensional data.

Support Vector Machines (SVM): Notable for its ability separate out data, even in high-dimensions, through a hyperplane.

Naïve Bayes: Although the simplest and fastest probabilistic classifier based on Bayes' theorem, Large datasets have demonstrated the effectiveness of Naive Bayes.

Logistic regression: A linear model, frequently applied to binary classification problems, producing probabilistic outcomes.

Gradient Boosting: An ensemble procedure that incrementally combines multiple weak learners.

Although deep reinforcement learning (DRL) and artificial bee colony (ABC) algorithms have been used in prior works and explored as potential approaches to complex fraud detection tasks and optimization, in this project only a comparative

An assessment of the previously developed machine learning algorithms is taken into consideration.

Model Training and Evaluation

The preprocessed dataset is divided into training data (80%) and testing data, (20%) to ensure the model performance evaluation is not biased. Every model algorithm is trained independently on the training data set. The each of the models are then tested on the unseen testing data set with standard evaluation metrics: accuracy, accuracy, memory, and F1-score. The outcomes of the evaluation stage were compared and analyzed to ascertain which efficacy of the algorithm, and which performed better against fraudulent transactions and legitimate transactions.

Model Integration with Flask

For case study and real time fraudulent detection, the top models has been implemented into a web application using Flask. The web application has created endpoints for data input, and the prediction, and, has allowed for businesses to send their transactional data and immediately receive prediction, whether 'legitimate' or 'not'. The web application justified recording the steps completed in the data-preprocessing stage, rationalising why the model was chosen, and detailing the training of the model and the steps to create the web application in Flask. A report documenting all of the steps that were taken, the findings, challenges faced throughout the project process and suggestion for future enhancements will be also be produced. Development of the Fraud Detection System represents a significant improvement in relation to traditional approaches, with the intention of improving both accuracy and efficiency of operation. Existing System Analysis Traditional e-commerce fraud detection systems are generally rule-based, whereby fraud detection relies on either manual review or both fraud alerts with manual review. The systems are typically designed such that a group of pre-defined rules can be generated e.g. transaction amount above $5000 or from high-risk geographical locations, anything that subsequently stimulates fraud alerts and these reviews are conducted manually to confirm. They also react as opposed to be proactive and suffer from acceptance of false positives.

Proposed System Analysis

The proposed solution utilizes the best methods for machine learning to identify fraudulent transactions in e-commerce with much higher accuracy and efficiency than traditional ways. Random Forest, SVM, Naive Bayes, Logistic Regression, and Gradient Boosting are among the methods that the suggested system can employ. to help identify complex and subtle patterns associated with fraud, using historical data. The proposed system uses a trained model integrated into a web application based on Flask to generate predictions in real-time and integrates with other current e-commerce or payment processing platforms. The expected benefits of the proposed system are significant and include: Benefit 1- More Accurate: These machine learning models particularly the Random Forest classifier, can provide higher accuracy in fraud detection, reducing false positives and false negative. Benefit 2 - Real-Time Detection: The integration with a Flask application means fraud will be detected as it occurs, allowing an organization to respond immediately and limit risk and financial exposure. Benefit 3 - Real-Time Benchmarks: The model seamlessly integrates with Flask and handles the necessary input data preprocessing, inherently handling scaling and/or encoding, depending on what is required by the trained model.

Deployment and Testing

The Flask application, incorporating the optimized fraud detection model, is deployed on a suitable server or cloud plaVorm. Following deployment, thorough testing is conducted to verify the application's correct functionality and the accuracy of its predictions in a live environment. Furthermore, the application's performance and scalability are validated under various simulated and real-world scenarios to confirm its reliability and efficiency in handling the dynamic demands of e-commerce operations

I. System Architecture and working design

The fraud detection project has been carefully planned out of the system architecture

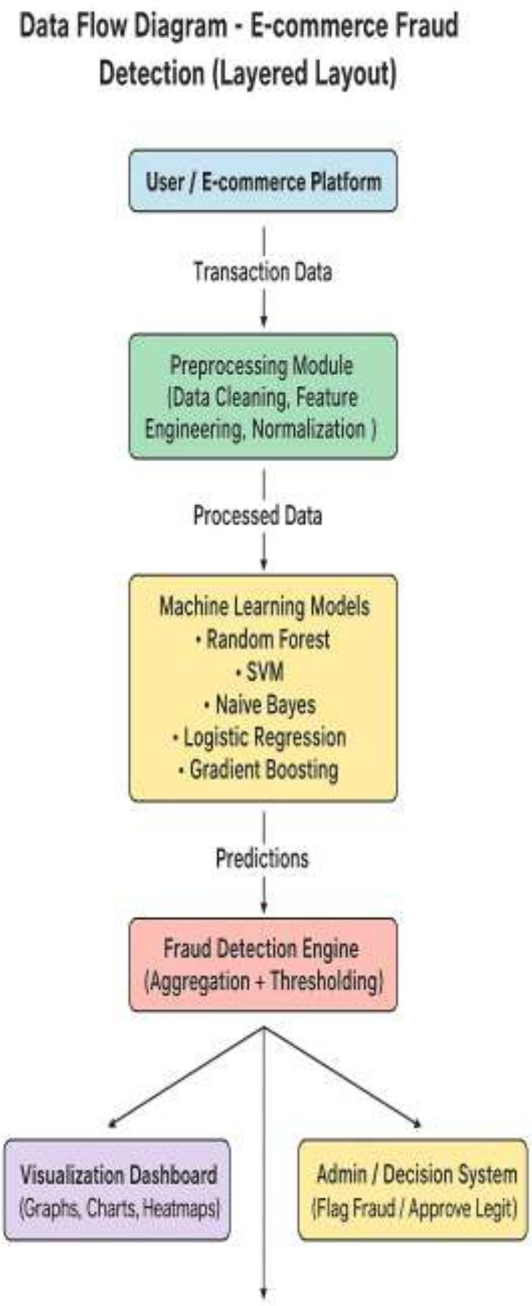
Data Collection: The data process collection was straightforward. This involved downloading a complete data collection from Kaggle that included various features to detect fraud.

Data Preparation: The data set will undergo a data preparation step. During this process the data will be evaluated, for example, checking for null values and dealing with them appropriately. In addition, categorical variables will be transformed into a format suitable to work with.

Feature Selection: its crucial to select relevant and meaningful features to help improve model training performance and to improve accuracies in predicting anomalies.

Training Machine Learning Models: The features selected above will be employed to teach a number of different machine learning models.

Real Time PredictionsThe machine learning models will automatically classify incoming transactional data after they have been trained. activity as an anomaly (fraud) or provide assurance where fraud is not present. This will allow for real-time identification and preventative action against fraudulent behaviour. Integration with Flask web application to facilitate interactive access to the final data set, and offer interaction.


Data Flow Diagram - E-commerce Fraud Detection (Layered Layout)
User / E-commerce Platform
Transaction Data
Preprocessing Module (Data Cleaning, Feature Engineering, Normalization)
Processed Data
Machine Learning Models
• Random Forest
• SVM
• Naive Bayes
• Logistic Regression
• Gradient Boosting
Predictions
Fraud Detection Engine (Aggregation + Thresholding)
Visualization Dashboard (Graphs, Charts, Heatmaps)
Admin / Decision System (Flag Fraud / Approve Legit)

The project architecture for the fraud detection project has been planned and discussed in depth

1.Raw JSON Transaction - It all starts with obtaining the raw transaction in JSON format. It should be noted that there are both numerical (i.e., transaction amount, time difference) and categorical (i.e., payment method, merchant type)

2 Preprocessing: - Upon obtaining the data, we will separate the incoming data into a numerical and categorical, for separation of dedicated preprocessing.

3.Numerical → Standard Scaler: - The numerical features were transformed into a standard approach that had a mean of 0 and standard deviation of 1, this means that the model cannot become biased away from features.

4.Categorical → One Hot Encoder: A process for converting categorical features into binary vectors (0/1 formats), where there is one vector for every unique.

5.Feature Union: - The numerical and categorical features were compiled into one union dataset.

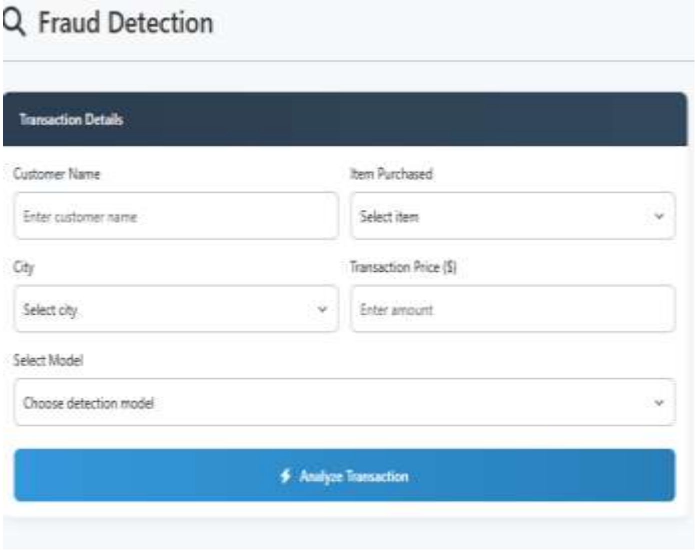
6.Random Forest Classifier:- The union data set will now be input into a Random Forest Classifier, which will use multiple decision trees to vote whether or not the transaction is fraudulent.

7.Output Fraud/Not-Fraud: -The model will output a binary;
-Fraud→ Triggers alert or mitigation actions.
- Not Fraud → Transaction proceeds without intervention.

The efficiency of The models for machine learning and the benefits of the overall design, including ability to scale and detect in real-time, are dependent upon the availability of sufficient computing power. the minimum hardware and software specifications required to operate the fraud detection system.

Results and discussions

The empirical assessment of the proposed e-commerce fraud detection system involved extensive trials and analysis of a Several types of machine learning models on a publicly accessible data The dataset used for this project was sourced from publicly available data on Kaggle consisting of credit card transactions made by cardholders residing in Europe in September 2013. The dataset contains 284,807 transactions from the period of two-days and contains 492 instances of fraudulent digital transactions. Since only 0.172% of the transactions in this dataset were fraudulent, the significant class imbalance must be mentioned. The dataset has input variables, all primarily in awesome forms, that were transformed from what we call Principal Components Analysis (PCA) (V1 to V28). The dataset was not made public since it lacked any more contextual information or the original variables. due to confidentiality, and therefore PCA was performed and only the transformed values were released. The 'Time' attribute describes the time in seconds elapsed between a transaction and the original transaction within this datasetThe transaction's monetary value is shown by the Amount element. The result binary classification is denoted by the Class property. represents the binary output class (target) for an instance. where '1' relates to a digital fraudulent transaction, and '0' represents a legitimate transaction.



Fraud Detection

Transaction Details

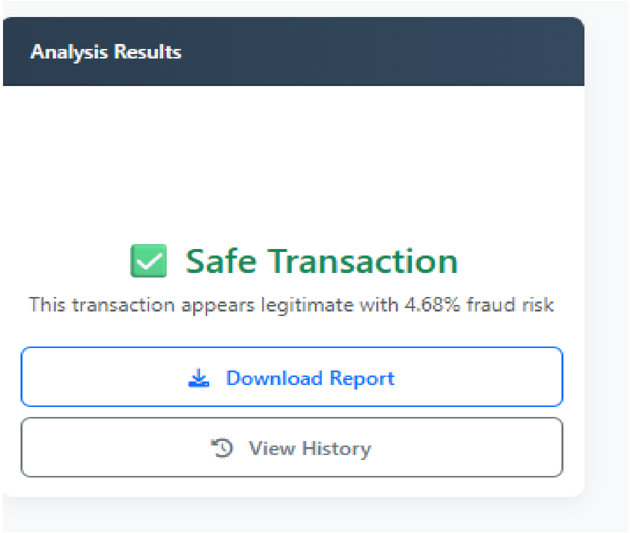
Customer Name: Enter customer name

Item Purchased: Select item

City: Select city

Transaction Price ($): Enter amount

Select Model: Choose detection model



Analysis Results

Safe Transaction

This transaction appears legitimate with 4.68% fraud risk

Fig 4.5 shows the fraud details to input in this dashboard

Fig 4.5 shows the analysis of fraud distribution

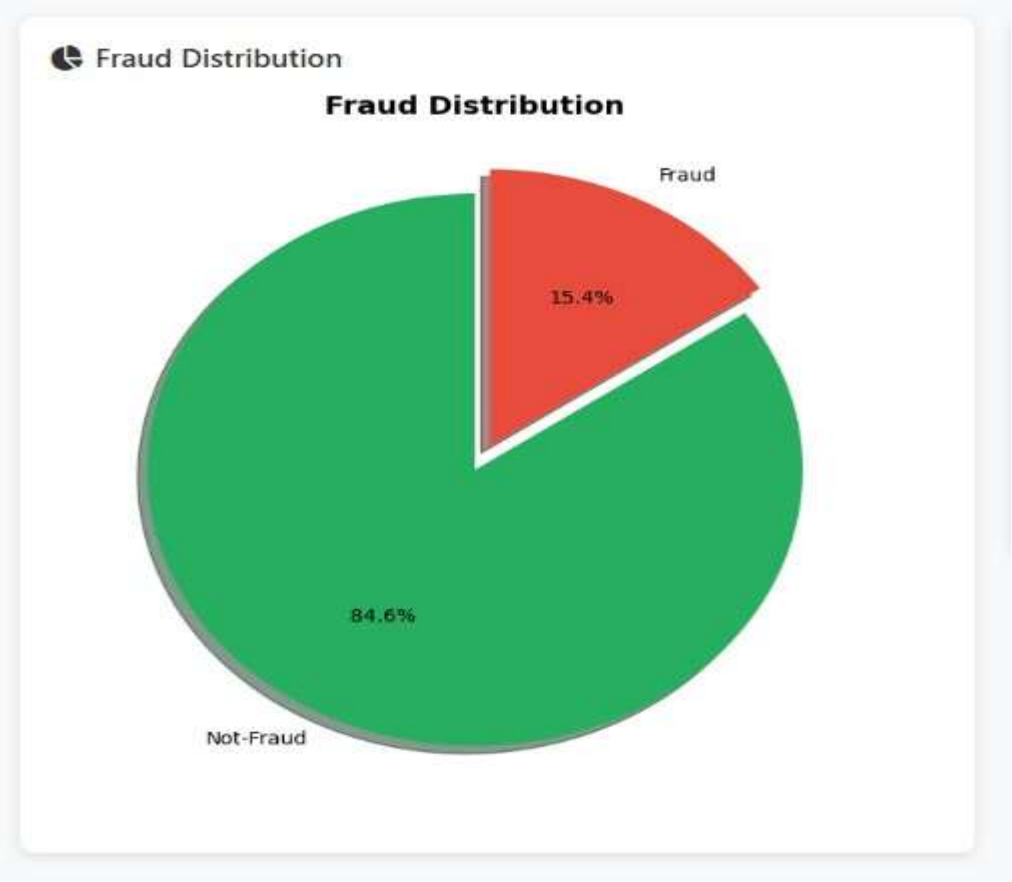
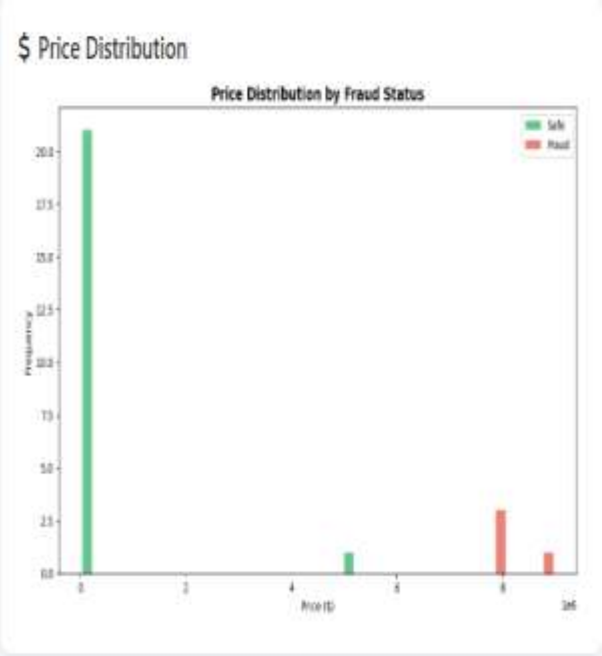


Fig 4.5 shows the status of fraud risk safe or unsafe



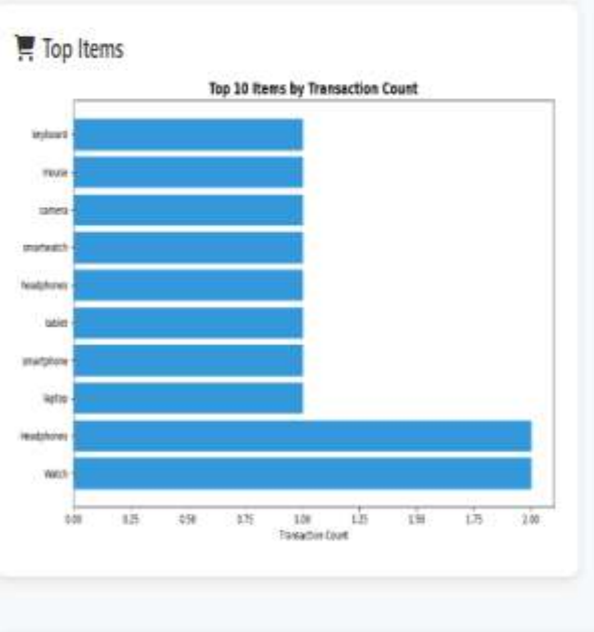


Fig 4.5 shows the price distribution of fraud status

Fig 4.6 shows the top cites fraud transactions

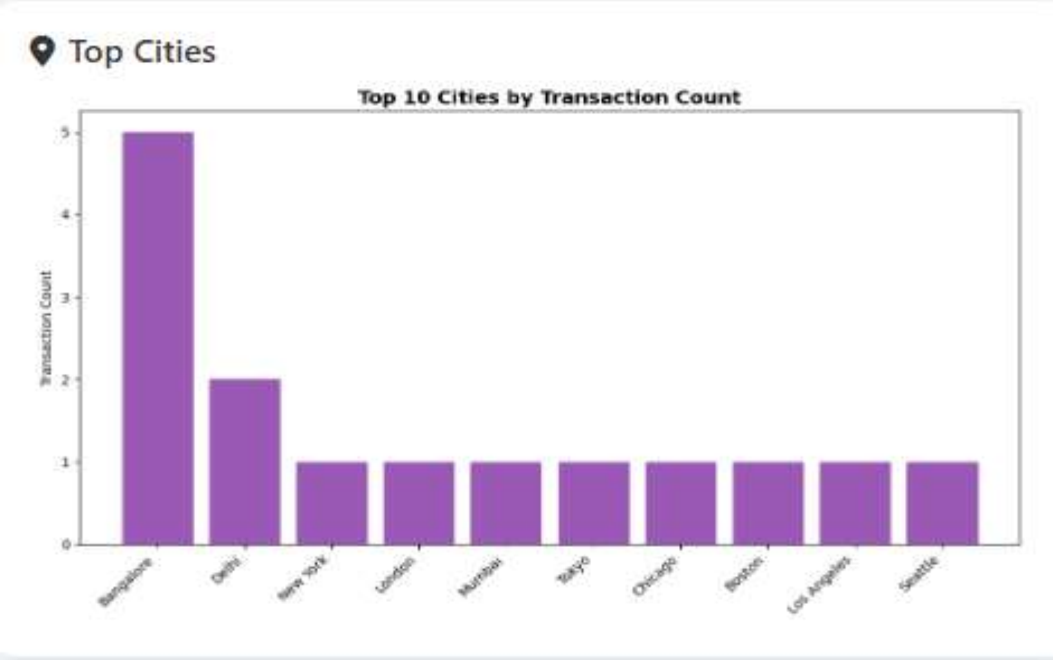
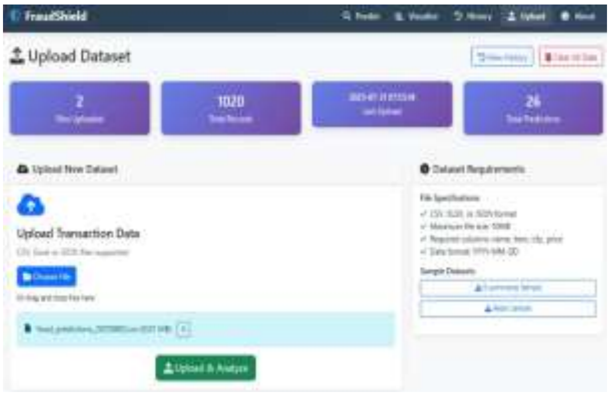


Fig 4.7 shows the top items transaction count



Upload Dataset

2 Files Uploaded

1020 Total Records

2025-07-21 07:55:[illegible] Last Upload

26 Total Predictions

Upload New Dataset

Upload Transaction Data

CSV, Excel or JSON files supported

Or drag and drop files here

[illegible]

Dataset Requirements

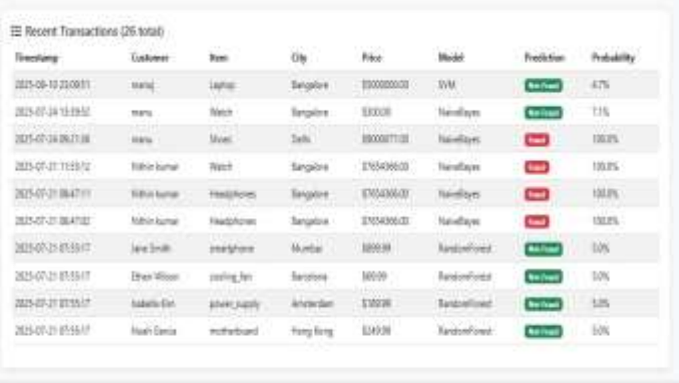
File Specifications

- ✓ CSV, XLSX, or JSON format
- ✓ Maximum file size: 50MB
- ✓ Required columns: name, item, city, price
- ✓ Date format: YYYY-MM-DD

Sample Datasets

- E-commerce Sample
- [illegible]

Fig 4.8 shows the uploading of dataset details



☰ Recent Transactions (26 total)

Timestamp	Customer	Item	City	Price	Model	Prediction	Probability
2025-08-10 23:09:11	manoj	Laptop	Bangalore	[illegible]	SVM	Not Fraud	47%
2025-07-24 15:59:52	manu	Watch	Bangalore	[illegible]	NaiveBayes	Not Fraud	7.1%
2025-07-24 09:21:38	manu	Mixer	Delhi	[illegible]	NaiveBayes	Fraud	100.0%
2025-07-21 11:59:12	Nithin kumar	Watch	Bangalore	[illegible]	NaiveBayes	Fraud	100.0%
2025-07-21 08:47:11	Nithin kumar	Headphones	Bangalore	[illegible]	NaiveBayes	Fraud	100.0%
2025-07-21 08:47:02	Nithin kumar	Headphones	Bangalore	[illegible]	NaiveBayes	Fraud	100.0%
2025-07-21 07:55:17	Jane Smith	smartphone	Mumbai	$899.99	RandomForest	Not Fraud	0.0%
2025-07-21 07:55:17	Ethan Wilson	cooling_fan	Barcelona	$69.99	RandomForest	Not Fraud	0.0%
2025-07-21 07:55:17	Isabella Kim	power_supply	Amsterdam	$139.99	RandomForest	Not Fraud	5.0%
2025-07-21 07:55:17	Noah Garcia	motherboard	Hong Kong	$249.99	RandomForest	Not Fraud	0.0%

Fig 4.9 shows the recent history of fraud and not fraud details

Model Performance Evaluation

Relevant measures, including accuracy, precision, recall, and F1-score, were used to assess the performance of the machine learning models that were taken into consideration and put into practice. These metrics are especially useful to assess performance on imbalanced datasets since accuracy could be misleading. During the project evaluation of the selected machine learning algorithms, accuracy rates were recorded as follows:

Random Forest: 97.06%, Support Vector Machines (SVM): 96%, Naive Bayes: 91%, Logistic Regression: 94%, Gradient Boosting: 81%,The Random Forest classifier produced the most accurate model in this study with a high accuracy rate of 97.06%. It is useful to look at performance from other related studies, even if they used a different dataset or had slight differences within methodologies, to provide context to these results. For example, in another study dealing with automatic fraud detection in e-commerce transaction, they provided the performance metrics reported for other common machine learning algorithms.

VII Future scope enhancements

Although we have made considerable progress in machine learning in fraud detection, we still face multiple inherent challenges, and addressing them will underpin future enhancements in this evolving landscape. Inherent Challenges in Fraud Detection

Data Quality Issues: The potential for machine learning models to effectively deliver is dependent upon the data that is used in their training. Fraud datasets suffer from the following: Noise: Irrelevant or erroneous data can wrongfully alter predictions, which will decrease stem accuracy;

Imbalance: Legitimate transactions will occur which will bias the distribution...Future enhancements to e-commerce systems for detecting fraud should concentrate on defining the extent of these constraints to gain as much performance, flexibility, and real- world value from systems.

Dataset Expansion and Diversification: Expanding and diversifying datasets with more recent and varied transactional data is crucial for enhancing model robustness and adaptability to evolving fraud A significant area for growth involves incorporating multimodal data, such as images and text, beyond traditional numerical and categorical features mined from transaction histories.[1] This suggests that the next frontier in fraud detection involves richer data representations and more complex models that can capture subtle, non-obvious patterns, moving beyond purely tabular transaction information to integrate insights from various digital footprints. The system capacity to identify might be further enhanced by including complex deep learning architectures , such as LSTM networks and graph neural networks or improved anomaly detection approaches and complex fraud behaviors in real-time, pushing accuracy beyond current thresholds. Feedback Loop Mechanisms: Implementing a continuous feedback loop mechanism, based on real-time user interactions and transaction outcomes, could continuously refine model precision and cut down on false positives by allowing the system to learn from its own predictions and real-world results.with Emerging Technologies: Exploring the integration of blockchain technology for immutable transaction records could enhance transparency and tamper-proof auditing, while leveraging AI- driven chatbots for real-time customer

Verification could enhance user experience and improve fraud prevention. Researching Under-Researched Fraud Types: There is a knowledge gap for certain, widespread types of fraud. For example, reseller fraud (product flipping/scalping), is economically relevant but under-explored in the ML literature. There should be more emphasis on developing and using machine learning techniques for these under-researched types, as well as new types of fraud risk (triangulation fraud and bot fraud in particular). Defenses for Adversarial Attacks: Further research and continued development of reliable defenses is needed in order to protect fraud detection from new adversarial attacks that seek to manipulate the output of their models. Enhanced Interpretability: Future work can enhance interpretability by developing more interpretable complex models, or utilizing explainable AI methods to enhance understanding of model decisions, thereby enhancing trust and support of human analysts 1 Sandbox Environments for Real-World Data: To alleviate the problem of data scarcity, seeking collaborations and developing safe sandbox environments to allow researchers to work with anonymized or synthetic real-world fraud data can significantly advance the research field. Ongoing Surveillance and Retraining: Because fraud is not static, implementing sufficient and sound statistical monitoring and policies to retrain models using new data, is critical for accuracy.

IX Conclusion

This project has effectively tackled the vital task of detecting e-commerce fraud in a meaningful way, by showing the potential worth of various machine learning algorithms to address this issue in substantial ways. The overall process and approach to completing this project involved the thoughtful and systematic process of data collection and preparation, as well as careful consideration around which algorithms to apply for the task, to then train and evaluate models, and then finally working to implement a real-time web application. Achievements include having evaluated three predictive models and demonstrating successful outcomes in terms of accuracy rates across all models, with Random Forest. classifying the specific test set with an accuracy of 97.06%, ensuring proper fraud detection. The other major contribution of the project was a deployed Flask application, which greatly simplified model deployment, now allowing organizations to complete real-time fraud prevention. Again, these are noteworthy accomplishments from previous fraud detection accuracy and implementation differentials with rule-based detection methods. Success from this project is not only a technical achievement, it also demonstrated the usefulness of machine learning in practical implementations that can protect financial ecosystems and rebuild consumer trust in online marketplaces. Long term, the e-commerce fraud detection domain will not stop evolving with new modes of fraud so continued development and research in this area is necessary. It is likely other more advanced deep learning solutions, like LSTM networks and Graph Neural Networks, will provide more definitive advancement of detection in identifying subtle and complex fraud techniques.1 Scalability and real time processing will have to greatly improve to provide better adaptability to a statistically inflating number of transactions and changing fraud techniques.1 We will also need to work towards detecting less understood fraud types such as reseller fraud and prove the model's explainability while producing strategies to adapt to adversarial attacks.1 The project results obviously made transactions more secure today, but they also provide a foundation for developing more complex and adaptable fraud avoidance systems with greater sophistication that will be able to further.

XI. REFERENCES

Detection: Potential and Applications," 2022.

C. Bell and F. Evans, "Fraud Detection and Prevention Systems: Design and Implementation," 2020.

N. Peterson and T. Cooper, "The Role of Natural Language Processing in Fraud Detection: A Review," 2021.

H. Evans and A. Adams, "Cybersecurity and Fraud Prevention in E-commerce: A Comprehensive Study," 2023.

L. Walker and I. Scott, "Unsupervised and Semi-supervised Learning for Fraud Detection: A Survey," 2022.

K. Garcia and H. White, "Online Grocery Shopping Fraud Preventions: A Comprehensive Review," 2019.

I. Moore and L. Perry, "Fraud Prevention Experiments in Banking Systems: Lessons Learned," 2021.

A. Roberts and J. Adams, "Data Mining Approaches to Fraud Detection in E-commerce: A Review," 2020.

— and J. Davis, "Data Privacy and Fraud Detection: A Survey," 2022.

R. Clark and L. Brown, "Detecting Cybercrime and Fraud in E-commerce: A Multi-disciplinary Review," 2019.

W. Hall and M. Miller, "Machine Learning for Fraud Detection in E-commerce: A Systematic Review," 2021.

S. Bennett and P. Clark, "Ethical Considerations Surrounding the Use of Machine Learning for Fraud Prevention: A Review," 2022.

F. Johnson and L. Taylor, "Blockchain Technology and Fraud Detection in E-commerce: A Review," 2022.

H. Brown and R. Moore, "Credible Data Sources for Fraud Detection: A Multiple-case study and future directions," 2022.

A.Mutemi and F.Bacao,"E-Commerce Fraud Detection Based on Machine Learning Techniques: Systematic Literature Review,"