

# Fraud Cascade: A Multi-Role Transaction Anomaly Framework for Distributed E-commerce Networks

Prof. Ashwini Sangam<sup>1</sup>, Bhagyashree<sup>2</sup>

<sup>1</sup> Professor, Master of Computer Application, VTU, Kalaburagi, Karnataka, India

<sup>2</sup> Student, Master of Computer Application, VTU, Kalaburagi, Karnataka, India

\*\*\*

**ABSTRACT** - E-commerce fraud presents a critical threat to businesses, consumers, and financial systems, demanding adaptive detection strategies beyond static rule-based models. This study introduces Fraud Cascade, a multi-role transaction anomaly framework designed for distributed e-commerce networks involving customers, merchants, and intermediaries. The framework integrates advanced preprocessing, feature engineering, and diverse machine learning classifiers—Random Forest, Gradient Boosting, Logistic Regression, Naïve Bayes, K-Nearest Neighbors, and Support Vector Machines—to predict fraudulent behavior with comparative evaluation. Fraud Cascade employs predictive modeling with probability assignment, dynamic visualization, and structured databases for monitoring and audit trails, enhanced by real-time notifications for stakeholders. Experimental results show ensemble methods achieve superior robustness and accuracy, demonstrating Fraud Cascade's scalability and adaptability in mitigating evolving fraudulent threats across digital marketplaces.

**Keywords:** E-commerce fraud, anomaly detection, machine learning, Random Forest, Gradient Boosting, fraud prediction, distributed networks, real-time alerts.

## 1. INTRODUCTION

E-commerce has revolutionized the global marketplace by providing businesses and consumers with unparalleled convenience, scalability, and accessibility. With millions of transactions occurring daily across retail, travel, finance, and digital services, online platforms have become integral to modern commerce. However, the rapid digitalization of transactions has simultaneously created new vulnerabilities that fraudsters exploit through weaknesses in payment systems, identity verification, and monitoring processes [1]. Beyond financial losses, e-commerce fraud erodes consumer trust, damages brand reputation, and threatens the long-term sustainability of online businesses. Global projections indicate rising costs of fraud, driven by advanced attack vectors such as account takeovers, triangulation fraud, and synthetic identity creation [2]. Consequently, developing intelligent, scalable

Fraud detection systems has become a critical necessity for digital platforms. Traditional systems largely depend on predefined rules and threshold-based alerts, which, while effective against known fraud patterns, fail to adapt to evolving techniques and often produce high false-positive rates, increasing customer dissatisfaction [3]. These limitations have driven research toward machine learning and artificial intelligence solutions capable of dynamically analyzing transaction data, detecting anomalies, and distinguishing fraudulent from legitimate behavior with improved accuracy [4]. To address this challenge, we propose Fraud Cascade: A Multi-Role Transaction Anomaly Framework for Distributed E-commerce Networks, which integrates multiple machine learning classifiers—including Gradient Boosting, Random Forest, Logistic Regression, Naïve Bayes, K-Nearest Neighbors, and Support Vector Machines—to evaluate performance across models [5]. The framework encompasses preprocessing, feature extraction, classification, and dynamic reporting, with built-in real-time alerts and distributed scalability [6][7].

## 2. PROBLEM STATEMENT

E-commerce fraud has become a pressing issue in digital marketplaces, fueled by the increasing scale, velocity, and anonymity of online transactions. Conventional detection systems, often based on static rules or basic statistical models, fail to keep pace with the dynamic and sophisticated strategies employed by fraudsters. These methods typically suffer from high false-positive rates, leading to unnecessary customer friction, and high false-negative rates, allowing fraudulent activities to bypass detection. Moreover, the absence of multi-role transaction analysis involving customers, merchants, and intermediaries reduces detection accuracy. Thus, there is a critical need for scalable, adaptive, and intelligent frameworks to address these challenges.

### 3. OBJECTIVES

The primary objective of this study is to design and implement Fraud Cascade, an advanced fraud detection framework tailored for multi-role e-commerce transactions. The work leverages machine learning algorithms—including Support Vector Machine, Random Forest, Naïve Bayes, Logistic Regression, Gradient Boosting, and K-Nearest Neighbors—to classify transactions as legitimate or fraudulent with high accuracy. A publicly available Kaggle dataset on e-commerce fraud is utilized for model training and validation to ensure reliability and diversity. Another key objective is to deploy the best-performing model into a Flask-based web application, providing real-time fraud detection with a scalable, user-friendly interface for practical adoption.

### 4. METHODOLOGY USED

**1) Data Collection:** The dataset for this study is obtained from Kaggle, comprising real-world e-commerce transactions labeled as either fraudulent or legitimate. It includes features such as transaction amount, timestamp, user behavior, and participant attributes, all of which are critical for identifying fraud patterns. Kaggle datasets are widely recognized for their quality and diversity, making them reliable for machine learning research. This dataset serves as the foundation for model development, ensuring alignment with real-world fraud detection scenarios.

**2) Data Preprocessing:** Raw transactional data may contain missing values, duplicates, or noise that can hinder model accuracy. To address this, preprocessing steps such as imputation of missing values, normalization, and encoding of categorical features are performed. Numerical attributes are scaled to maintain consistency across models sensitive to data distribution, such as K-Nearest Neighbors and Support Vector Machines. These steps produce a clean, balanced, and structured dataset for training.

**3) Feature Extraction:** Feature extraction emphasizes selecting attributes that provide meaningful insights into fraudulent behaviors. Key attributes such as transaction frequency, monetary value, device usage, and participant interactions are analyzed. Statistical transformations and correlation analysis are applied to enhance discriminative power between fraudulent and legitimate samples. This process reduces dimensionality while preserving essential information.

**4) Model Selection:** Multiple classification algorithms—including Random Forest, Support Vector Machine, Logistic Regression, Naïve Bayes, Gradient Boosting, and K-Nearest Neighbors—are employed for experimentation. These models are chosen for their proven effectiveness in fraud detection and ability to handle high-dimensional data. By leveraging diverse paradigms, from ensemble learning to probabilistic methods, the study ensures robust comparison.

**5) Model Training:** The data set is divided into training and testing subsets. Each algorithm is trained on the extracted features, learning patterns that distinguish fraudulent from legitimate transactions. Hyperparameter tuning is applied to optimize performance and minimize over fitting, **ensuring strong generalization to unseen data.**

**6) Model Evaluation:** The trained models are evaluated using metrics such as accuracy, precision, recall, and F1-score. Since false negatives can be particularly costly in fraud detection, recall and precision are prioritized. Confusion matrices are used to visualize model predictions, enabling performance comparison and selection of the most effective algorithm.

**7) Integration with Flask:** The best-performing model is deployed through a Flask-based web application to provide real-time fraud detection. Flask's lightweight and flexible framework supports rapid deployment and scalability. The application incorporates user authentication, input forms for transaction details, and dynamic result visualization, bridging the gap between machine learning models and practical e-commerce usability.

### 5. LITERATURE SURVEY

**Article [1]** 'Effective fraud detection in e-commerce: Leveraging machine learning and big data analytics' by S.R.B. Reddy in 2024: This paper explores how machine learning and artificial intelligence advance fraud prevention in online commerce. It surveys common fraud types such as identity theft and credit card abuse, identifying limitations of rule-based systems like their inability to adapt and susceptibility to false positives and negatives. The authors detail the growing challenge of real-time detection due to increasing data volume and complexity. Various ML approaches are analyzed for their ability to uncover subtle behavioral anomalies. The paper provides a panorama of detection methods, critiques dataset handling, and proposes future directions for evolving fraudster tactics. Big data analytics integration is shown to enhance scalability and system responsiveness to new threats.

**Article [2]** 'E-Commerce Fraud Detection Based on Machine Learning Techniques: Systematic Literature Review' by A. Mutemi, et al. in 2024: Conducting a systematic literature review using PRISMA, the authors examine the effectiveness of machine learning and data mining for e-commerce fraud detection. The survey reviews 101 publications over the past decade, focusing on the application and comparative analysis of ML models on platforms like eBay and Facebook. The review uncovers gaps in feature engineering, imbalanced dataset handling, and interpretability challenges. Artificial neural networks and ensemble methods appear prominently. Discussed with key insights for practitioners and industry stakeholders, along with recommendations for future research directions.

**Article [3]** 'Fraud detection and prevention in e-commerce: A systematic literature review' by V.F. Rodrigues and others in 2022: This study reviews 64 articles published between 2016 and 2022 on fraud detection and prevention in e-commerce. It details the rise of global e-commerce and associated growth in fraud losses, mapping fraud types (e.g., payment, merchant, synthetic identity) and domains affected. Reviewed methods include rule-based, ML classifiers, and hybrid approaches. The paper proposes a taxonomy for e-commerce fraud, summarizing used datasets and evaluation metrics.

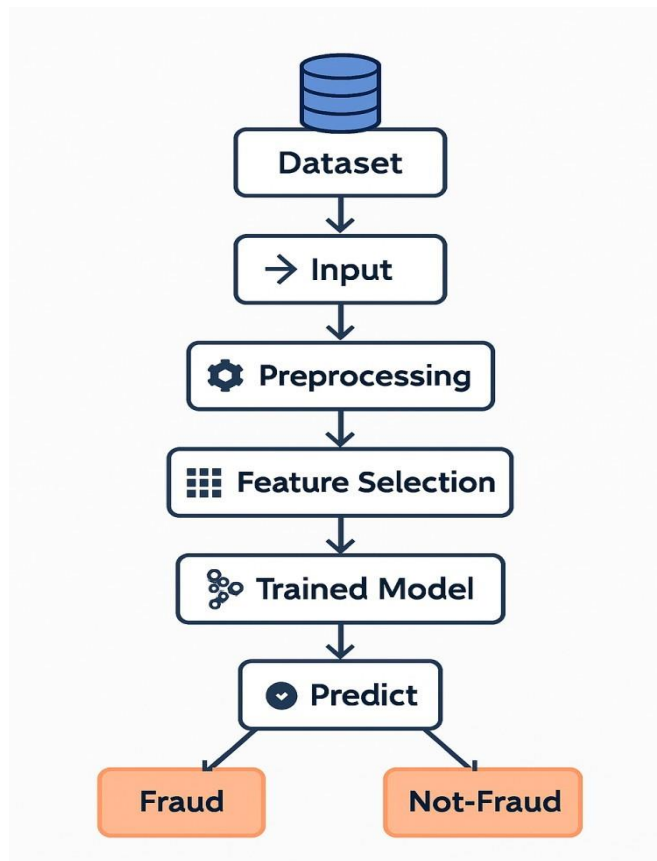
**Article [4]** 'Adaptive Anomaly Detection in Database Transactions Using RL-Based Models' by C. Reddy et al. in 2025: This paper introduces a reinforcement learning (RL)-based system for anomaly detection in complex transactional databases, targeting evolving fraud patterns in finance and e-commerce. Dynamic reward mechanisms reduce false positives and adapt to changing user behavior. The authors discuss imbalanced datasets, embedding transactions, and feedback-driven learning for enhanced accuracy. Performance is evaluated using both public Kaggle and synthetic datasets, with significant gains over traditional models (e.g., RL precision 95.2%, recall 92.4%). The framework's scalability and adaptability are highlighted, along with implications for security management in digital marketplaces.

**Article [5]** 'Enhancing Enterprise Financial Fraud Detection Using Machine Learning' by M.M. Ismail, Mohd Anul Haq in 2024: Focusing on enterprise financial fraud, this paper employs machine learning algorithms and data analytics for accurate anomaly detection. Exploratory data analysis uncovers missing values and tackles class imbalance. Random Forest, Isolation Forest, and Local Outlier Factor algorithms are presented, registering high accuracy (up to 99.9%). The framework emphasizes proactive identification of internal and external fraud, proposes techniques for multi collinearity management, and discusses applicability across business sectors.

**Article [6]** 'Detecting anomalies in block chain transactions using deep learning methods' by M. Hasan et al. in 2024: This research investigates anomaly detection in block chain networks, with direct relevance to e-commerce and financial systems. The authors analyze deep learning and statistical techniques for predicting transaction anomalies, leveraging block chain intelligence and decentralized data structures. Key contributions include unsupervised learning for pattern recognition, improved detection accuracy in real-time transaction flows, and handling vast and heterogeneous data sources.

**Article [7]** 'Year-over-Year Developments in Financial Fraud Detection via Deep Learning: A Systematic Literature Review' by Yisong Chen, Chuqing Zhao, Yixin Xu, Chuanhao Nie in 2025: This paper systematically reviews 57 studies published from 2019 to 2024, highlighting advances in deep learning for financial fraud detection. Convolutional Neural Networks, LSTMs, and transformers are reviewed across credit card, insurance, and audit domains. Major themes include model interpretability, handling data imbalance, and privacy frameworks. Challenges and automation opportunities are discussed alongside block chain integration and PCA for feature reduction. The review pinpoints future avenues for scalable and explainable DL-based fraud system improvement.

## 6. SYSTEMDESIGN



**Figure1: System Architecture of-Commerce Fraud Detection**

The fraud detection system processes transaction data by dividing it into training and testing sets while maintaining the same proportion of fraudulent cases in each subset. Initially, the data undergoes preprocessing, where numerical features such as transaction amounts are scaled, and categorical attributes like device types are encoded into machine-readable form. The system then evaluates multiple machine learning models—including Random Forest, Support Vector Machine, K-Nearest Neighbors, XGBoost, and Gradient Boosting—to determine the most

users enter transaction-specific details required for fraud analysis. The inputs include critical features such as transaction amount, behavioral attributes, device-related information, and other relevant variables used during model training. These parameters are validated and processed before being passed to the trained machine learning model for prediction, ensuring a reliable and effective approach for identifying fraud.

Each model is validated using cross-validation techniques to ensure consistent performance and reliability. The best-performing model is selected based on its ability to detect fraudulent activities accurately while minimizing false positives. Finally, the chosen model is stored and deployed to automatically identify fraud in future transactions.

### 7. SCREENSHOTS

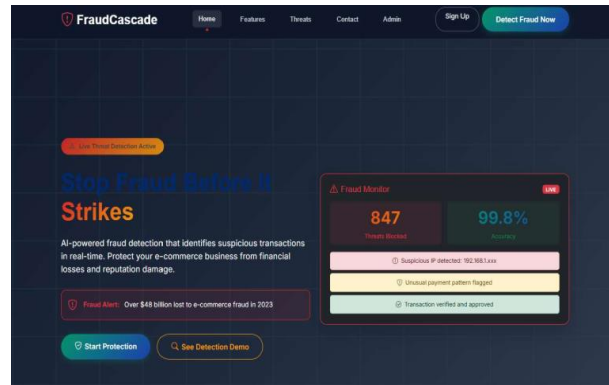


Figure6: Home page

Figure 6 illustrates the home page of the proposed e-commerce fraud detection system developed using the Flask web framework. This page acts as the primary interface for users to access the application. It provides a concise overview of the system functionality and enables seamless navigation to the transaction analysis module. The interface is designed to be intuitive and user-friendly, facilitating efficient interaction with the fraud detection platform.

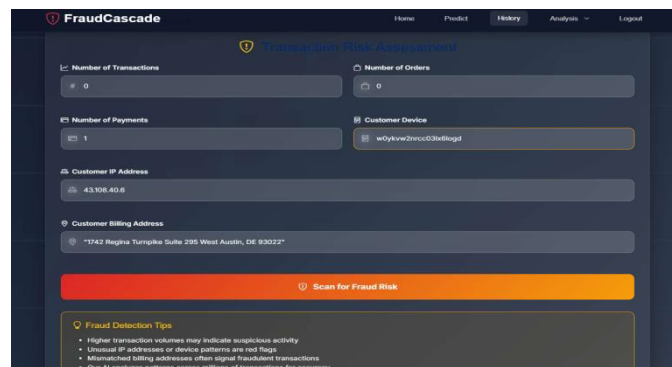


Figure7: Input parameters

real time, demonstrating the system’s effectiveness in providing immediate fraud detection and supporting timely decision-making for e-commerce stakeholders.

### 8. CONCLUSION & FUTURE SCOPE

In this research, an intelligent framework was successfully developed to detect fraudulent transactions in e-commerce environments. Multiple machine learning algorithms—including Random Forest, Support Vector Machine, K-Nearest Neighbors, Gradient Boosting, and XGBoost—were implemented and evaluated on a Kaggle dataset after extensive preprocessing and feature extraction. Comparative analysis highlighted the superior performance of ensemble-based models such as Gradient Boosting and XGBoost, while other algorithms contributed valuable diversity, enhancing system robustness. A dynamic Flask-based web application was deployed to provide real-time fraud prediction, detailed reporting, and secure storage of transactions and prediction logs through an SQL schema, ensuring transparency and auditability.

The system demonstrates high accuracy and efficiency, surpassing traditional rule-based approaches while strengthening transaction security and building user

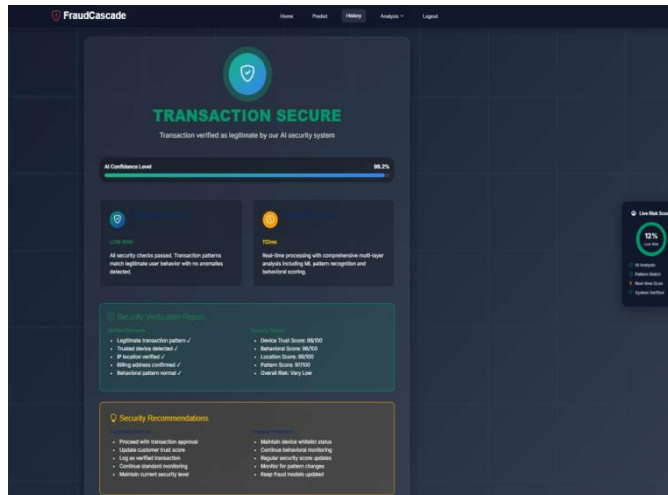


Figure8: Predicated Result

Figure 8 shows the predicted result generated by the system after processing the transaction data. Based on the trained classification model, the transaction is identified as either fraudulent or legitimate. The result is displayed in.

## 9. REFERENCES

- [1] S. R. B. Reddy, "Effective fraud detection in e-commerce: Leveraging machine learning and big data analytics," *ScienceDirect*, 2024.
- [2] A. Mutemi, et al., "E-Commerce Fraud Detection Based on Machine Learning Techniques: Systematic Literature Review," *BDMA*, 2024.
- [3] V. F. Rodrigues, et al., "Fraud detection and prevention in e-commerce: A systematic literature review," *ScienceDirect*, 2022.
- [4] A. Mutemi, et al., "E-Commerce Fraud Detection Based on Machine Learning Techniques: Systematic Literature Review," *SciOpen*, 2024.
- [5] C. Reddy, et al., "Adaptive Anomaly Detection in Database Transactions Using RL-Based Models," *European Journal of Artificial Intelligence*, 2025.
- [6] M. M. Ismail, M. Anul Haq, "Enhancing Enterprise Financial Fraud Detection Using Machine Learning," *Engineering, Technology & Applied Science Research*, 2024.
- [7] M. Hasan, et al., "Detecting anomalies in blockchain transactions using deep learning methods," *ScienceDirect*, 2024.
- [8] Y. Chen, C. Zhao, Y. Xu, C. Nie, "Year-over-Year Developments in Financial Fraud Detection via Deep Learning: A Systematic Literature Review," *arXiv*, 2025.
- [9] S. Kumari, et al., "A Comprehensive Investigation of Anomaly Detection Techniques for Financial Data," *IET Research*, 2024.
- [10] L. Hernandez Aros, et al., "Financial fraud detection through the application of machine learning techniques: A systematic literature review," *Nature*, 2024.