ONLINE TRANSACTION FRAUD DETECTION USING BACKLOGGING ON E-COMMERCE WEBSITE

Moses R[1], Krishna T[2], Lokeswaran P[3], Sangavi N[4]

[1] *Student, Computer Science and Engineering, Bannari Amman Institute of Technology, Tamil Nadu, India*

[2]*Student, Computer Science and Engineering, Bannari Amman Institute of Technology, Tamil Nadu, India*

[3]*Student, Information Technology, Bannari Amman Institute of Technology, Tamil Nadu, India*

[4]*Faculty, Computer Science and Engineering, Bannari Amman Institute of Technology, Tamil Nadu, India*

ABSTRACT

Online transactions are becoming more common and convenient, but they also pose a risk of fraud and cyber-crime. To prevent and detect fraudulent transactions on e-commerce websites, a system using backlogging is proposed. Backlogging is a technique that blocks the application of a transaction until it is verified by the user or the bank. The system uses a behaviour and location analysis (BLA) to compare the current transaction with the user's previous patterns and preferences. If the BLA detects any anomaly or inconsistency, the system asks for a re-verification from the user or the bank. The system also uses machine learning methods to identify and classify the types of frauds and the fraudsters. The system aims to reduce the false positive and false negative rates of fraud detection, and to enhance the security and trust of online transactions. The system is implemented using Python and tested on a simulated e-commerce website. The results show that the system can effectively detect and prevent various types of frauds, such as identity theft, card cloning, phishing, and spoofing. The system also provides a user-friendly interface and a feedback mechanism for the users and the banks. The system can be integrated with existing e-commerce platforms and can be customized according to the needs and preferences of the users and the banks. If any sort of surprising pattern is detected by the FDS then it asks for a re- verification. The algorithm used in the system then analyses all previous information of that card holder and recognizes any unusual pattern in the payment procedure.

.

Keyword : - *Fraud detection, Backlogging, Behaviour analysis, Machine learning, Identity theft*

1. INTRODUCTION

The advent of e-commerce has revolutionized the retail landscape, offering unprecedented convenience and accessibility to consumers worldwide. However, this digital convenience comes hand in hand with the ever-present challenge of combating online transaction fraud. As e-commerce transactions continue to surge in both volume and complexity, the imperative for robust fraud detection mechanisms becomes increasingly pressing. Traditional fraud detection systems, while effective to a degree, often struggle to keep pace with the rapidly evolving tactics employed by fraudsters. This necessitates the exploration of innovative approaches, such as the utilization of backlogging, to fortify e-commerce security.Online transaction fraud represents a formidable threat to both consumers and businesses alike. The repercussions extend beyond mere financial losses, encompassing erosion of trust in e-commerce platforms and compromised data security. Current fraud detection systems are frequently hampered by their inability to access granular transaction details, thereby impeding their capacity to accurately identify fraudulent activities. This project

seeks to redress these deficiencies by implementing a comprehensive fraud detection system enriched by the utilization of the backlogging technique.The primary objective of this project is to conceive, develop, and deploy a resilient fraud detection system finely tuned to the unique requirements of e-commerce platforms. To achieve this overarching goal, the project aims to:

Enhance the security posture of e-commerce platforms through the deployment of sophisticated fraud detection mechanisms capable of discerning even the most subtle fraudulent activities.Minimize financial losses attributed to online transaction fraud by instituting real-time monitoring and prevention strategies designed to swiftly identify and mitigate fraudulent transactions.

Cultivate user trust and confidence in online transactions by creating a secure and trustworthy environment wherein users can conduct transactions with peace of mind.Streamline and optimize fraud detection operations through the implementation of backlogging for retrospective analysis, facilitating the identification of historical fraud patterns and informing future prevention strategies.The significance of implementing an innovative fraud detection system enriched by the backlogging technique within the realm of e-commerce transcends mere technological advancement; it represents a strategic imperative in safeguarding the integrity and sustainability of online transactions. Online transaction fraud stands as a formidable threat that not only jeopardizes financial assets but also undermines the trust and confidence essential for the flourishing of e-commerce platforms.By proactively addressing this pervasive threat through the development and deployment of a robust fraud detection system, businesses can instill a sense of security and reliability among consumers, thereby fostering a conducive environment for continued growth and prosperity. The utilization of backlogging as a foundational technique enhances the system's ability to retrospectively analyze transaction data, enabling the identification of historical fraud patterns and informing future prevention strategies.Moreover, by fortifying e-commerce platforms against fraudulent activities, this initiative contributes to the preservation of consumer trust, which is paramount in sustaining long-term relationships and driving repeat business. The protection offered by an advanced fraud detection system not only shields consumers from potential financial losses but also safeguards businesses from reputational damage that could arise from being associated with fraudulent activities.Furthermore, the implementation of such a sophisticated fraud detection system underscores a commitment to operational excellence and regulatory compliance within the e-commerce landscape.

2. METHODOLGIES:

The methodology for developing and implementing enhanced security measures for online transactions involves a comprehensive approach that encompasses various components, tools, data collection techniques, procedures, testing methods, and adherence to standards. Below is a detailed description of each aspect:

2.1. Comprehensive Literature Review:

The objective of conducting a comprehensive literature review is fundamental in establishing a project's foundation by identifying existing research, methodologies, and technologies crucial to online transaction fraud detection, e-commerce security, and backlogging. This process involves thorough exploration across academic databases, journals, conferences, and industry publications to compile a diverse array of sources. Key topics that should be covered in the literature review include the historical evolution of online transaction fraud, current trends in e-commerce security, cutting-edge fraud detection techniques, research on backlogging applications in fraud detection, successful case studies of fraud detection systems, and regulatory requirements related to online transactions.By delving into these areas, researchers can gain valuable insights into the progression of online transaction fraud, the latest advancements in e-commerce security, innovative fraud detection methods, practical uses of backlogging techniques, successful instances of fraud detection systems, and the regulatory framework governing online transactions. This comprehensive literature review serves as a cornerstone for informed decision-making and the formulation of effective strategies to bolster security and reliability in online transactions.

2.2. Data Collection and Preprocessing:

Data collection is a critical step in building an effective fraud detection system, as it provides the raw material for analysis and model training. The process involves gathering various types of data related to online transactions, including transaction logs, user profiles, device information, IP addresses, geolocation data, and historical records. Data may be sourced from internal databases, third-party vendors, or publicly available datasets.Once collected, the data undergoes preprocessing to ensure its quality, consistency, and suitability for analysis. Data collection is a pivotal phase in constructing a robust fraud detection system, serving as the foundation for analysis and model training. This process involves gathering diverse data types related to online transactions, such as transaction logs, user profiles, device details, IP addresses, geolocation data, and historical records. The collected

data then undergoes preprocessing to ensure quality, consistency, and suitability for analysis. Preprocessing tasks include data cleaning to eliminate duplicates and errors, handling missing values through imputation or deletion, normalizing numerical attributes to a standard scale, and encoding categorical variables into numerical representations.Moreover, feature engineering plays a crucial role in enhancing the performance of fraud detection models by selecting, transforming, and creating new features from raw data. This process involves identifying relevant features that capture meaningful information about online transactions and user behavior, transforming raw data into informative features suitable for model training, and creating new features through extraction or aggregation to boost the model's discriminative power.Subsequently, model development entails designing, training, and evaluating machine learning algorithms to detect fraudulent transactions based on the engineered features. Various types of machine learning algorithms can be applied to fraud detection, including supervised learning algorithms like logistic regression and neural networks, unsupervised learning algorithms such as clustering techniques and anomaly detection methods, as well as semi-supervised learning algorithms that combine elements of both supervised and unsupervised learning.Finally, system implementation involves integrating the developed fraud detection model into the backend infrastructure of e-commerce websites. This phase includes implementing necessary software components, creating APIs for communication between systems, ensuring compatibility with existing systems, and testing the integration to guarantee smooth operation without disruptions. Testing and evaluation are critical steps to assess system performance using real-world or simulated datasets and metrics like accuracy and precision. Optimization and fine-tuning follow testing to enhance system efficiency based on evaluation results before deploying the optimized system for real-time usage with monitoring protocols in place to track performance and handle fraud cases effectively.

2.3. Feature Engineering:

Feature engineering is the process of selecting, transforming, and creating new features from the raw data to improve the performance of the fraud detection model.Feature engineering is a pivotal process in refining the performance of fraud detection models by selecting, transforming, and crafting new features from raw data. This strategic approach involves identifying pertinent features that encapsulate crucial information about online transactions, user behavior, and potential fraud indicators. By transforming raw data into insightful features suitable for model training and creating new features through extraction or aggregation, the discriminative power of the model is significantly enhanced.In the realm of fraud detection, examples of commonly utilized features include transaction amount, time of day, transaction frequency, user location, device type, IP address, and historical transaction patterns. Feature engineering stands as a critical juncture in constructing a robust and precise fraud detection model since it directly influences the model's capacity to differentiate between legitimate and fraudulent transactions. By meticulously engineering features that encapsulate relevant information, the model becomes adept at discerning patterns indicative of fraudulent activities.To further enhance fraud detection systems, methodologies encompass various components like comprehensive literature reviews to understand existing research and technologies related to fraud detection. Data collection and preprocessing play a crucial role in ensuring the quality and suitability of data for analysis. Feature engineering involves selecting and transforming raw data into informative features that improve model performance. Model development includes designing and training machine learning algorithms for fraud detection. System implementation integrates the fraud detection model into e-commerce websites' infrastructure. Testing and evaluation assess system performance using metrics like accuracy and recall. Optimization and fine-tuning iteratively improve the system's efficiency based on evaluation results.Examples of features commonly used in fraud detection include transaction amount, time of day, transaction frequency, user location, device type, IP address, and historical transaction patterns. Feature engineering is a crucial step in building a robust and accurate fraud detection model, as it directly impacts the model's ability to distinguish between legitimate and fraudulent transactions.

2.4. Model Development:

Model development in fraud detection using machine learning is a crucial process that encompasses designing, training, and evaluating algorithms to pinpoint fraudulent transactions based on engineered features. This strategy capitalizes on a variety of machine learning algorithms customized for fraud detection: Supervised Learning Algorithms: These encompass logistic regression, decision trees, random forests, support vector machines (SVM), and neural networks. Supervised learning involves training algorithms on labeled data to make predictions or decisions based on identified patterns. Unsupervised Learning Algorithms: Techniques like k-means clustering and anomaly detection methods such as isolation forests and one-class SVM fall under this category. Unsupervised learning aids in identifying patterns or structures in data without the need for labeled outcomes. Semi-Supervised Learning Algorithms: To combat this escalating threat, organizations are increasingly embracing machine learning as a potent tool for real-time fraud detection. By harnessing extensive datasets and advanced algorithms, machine learning can effectively identify suspicious patterns and anomalies indicative of fraudulent behavior, empowering businesses to

protect their customers, revenue, and reputation. Machine learning-based fraud detection systems have demonstrated remarkable accuracy rates of up to 96% in reducing fraud for eCommerce businesses. Leveraging machine learning models can significantly enhance enterprise fraud security by adapting to new information and detecting emerging fraud patterns as fraudulent actors evolve their tactics. While machine learning offers numerous benefits for fraud detection, it is essential to combine human insights with machine learning models to optimize performance and overcome challenges like false positives. This collaborative approach ensures a more robust fraud protection solution that leverages the strengths of both automated algorithms and human expertise.Random Forest Algorithm: Random Forest is a powerful ensemble learning method that operates by constructing a multitude of decision trees during training and outputting the mode of the classes as the prediction. Each tree in the forest is built using a random subset of features, which helps in reducing overfitting and improving accuracy. Random Forest is known for its robustness to noise and its ability to handle large datasets with high dimensionality. Robust Random Cut Forest (RRCF): RRCF is an anomaly detection algorithm that is particularly effective in detecting outliers or anomalies in data. It works by constructing a binary tree structure where each internal node represents a random hyperplane that partitions the data. RRCF is robust to variations in data distribution and can efficiently identify anomalies even in high-dimensional datasets. This technique is valuable for fraud detection as it can pinpoint unusual patterns or behaviors indicative of fraudulent activities

3. ALGORITHMS AND METHODS:

3.1 Support Vector Machines (SVM) Integration:

3.1.1 Model Training:

Lead thorough preparation of the SVM model utilizing verifiable exchange information named as normal.In the domain of coordinating Help Vector Machines (SVM) into misrepresentation discovery frameworks, the course of model preparation is a basic step that includes thorough preparation of the SVM model utilizing authentic exchange information marked as would be expected. This preparing system plans to outfit the SVM model with the important information and examples to actually recognize ordinary and false exchanges. Tuning hyperparameters, including the bit capability and regularization, is fundamental for upgrading the presentation of the SVM model, guaranteeing that it can adjust to various kinds of information and accomplish ideal precision in distinguishing false exercises inside a web based business climate.By leading careful model preparation, extortion recognition frameworks can upgrade their capacity to distinguish irregularities and examples demonstrative of deceitful way of behaving. The SVM model gains from verifiable information marked as should be expected exchanges, empowering it to perceive deviations from regular exchange designs that might connote likely misrepresentation. Tuning hyperparameters, for example, the bit capability and regularization boundaries considers tweaking the SVM model to accomplish ideal execution in ordering exchanges accurately.Furthermore, utilizing SVM with data gain-based arrangement procedures can altogether upgrade the exactness of charge card misrepresentation identification frameworks by working on the characterization of false exchanges. This approach includes using directed information to prepare the SVM model, empowering it to learn and group new exchanges given known examples of extortion and certified exchanges. By consolidating progressed procedures like SVM with data gain, extortion identification frameworks can accomplish higher precision rates in distinguishing fake exercises and limiting bogus up-sides inside Visa exchanges.

3.1.2 Real-time Anomaly Detection:

With regards to ongoing oddity identification inside misrepresentation location frameworks, the execution of the prepared Help Vector Machines (SVM) model to handle approaching exchanges progressively is a crucial stage towards improving the framework's capacity to quickly distinguish possibly false exercises. By incorporating the SVM model into the exchange handling pipeline, the framework can dissect approaching information streams persistently and go with quick choices given learned examples and oddities.Laying out a limit for oddity scores is critical continuously peculiarity location, as it permits the framework to hail exchanges that veer off essentially from the typical way of behaving, showing expected deceitful action. By setting suitable limits in light of authentic information and model execution, the framework can separate between genuine exchanges and dubious ones, empowering ideal mediation to forestall deceitful exercises inside an online business climate.Additionally, utilizing progressed strategies like versatile continuous abnormality discovery utilizing SVMs can altogether upgrade the framework's responsiveness to arising misrepresentation designs and advancing strategies utilized by fraudsters. By consistently checking exchange information and changing

irregularity location boundaries continuously, extortion discovery frameworks can adjust progressively to changing misrepresentation situations and keep an elevated degree of precision in recognizing deceitful exchanges.Integrating continuous abnormality recognition capacities controlled by SVM models not only fortifies the safety efforts of web-based business stages but also empowers proactive extortion counteraction systems that alleviate dangers and defend against monetary misfortunes. By utilizing the bits of knowledge acquired from SVM-based oddity identification, organizations can improve their misrepresentation location systems and remain in front of expected dangers in the powerful scene of online exchanges.

3.2 Robust Random Cut Forests (RRCF) Integration:

3.2.1 Temporal Anomaly Detection:

Integrating Robust Random Cut Forests (RRCF) into fraud detection systems includes utilizing worldly oddity identification capacities to successfully catch transient examples in exchange information. By stressing deviations from commonplace conduct after some time, RRCF can recognize unobtrusive abnormalities characteristic of fake exercises inside online business conditions. Upgrading hyperparameters, for example, tree size and the quantity of trees is pivotal for improving the framework's peculiarity recognition execution, guaranteeing that RRCF can actually adjust to differing information designs and precisely banner dubious exchanges.The coordination of RRCF empowers extortion recognition frameworks to dissect exchange information powerfully, zeroing in on worldly patterns and examples that might connote fake way of behaving developing over the long run. By recognizing oddities in light of worldly elements like exchange recurrence and season of-day designs, RRCF upgrades the framework's capacity to distinguish anomalies and deviations from ordinary exchange conduct, consequently working on the precision of extortion identification components inside web-based exchange conditions.Besides, by tweaking hyperparameters like tree size and the quantity of trees, misrepresentation identification frameworks can improve the presentation of RRCF in catching abnormalities successfully. These hyperparameters assume a vital part in deciding the responsiveness and explicitness of the irregularity recognition process, permitting the framework to figure out some kind of harmony between identifying certified misrepresentation cases and limiting misleading up-sides. Through fastidious boundary streamlining, misrepresentation recognition frameworks can upgrade their peculiarity identification capacities and fortify their capacity to battle false exercises proactively inside web based business stages.

3.2.2 Combined Anomaly Scores:

With regards to misrepresentation location frameworks, fostering an instrument to consolidate peculiarity scores from both Help Vector Machines (SVM) and Strong Irregular Cut Woodlands (RRCF) is essential for an exhaustive evaluation of exchange credibility. This blend takes into consideration a more powerful and exact recognition of false exercises by utilizing the qualities of the two calculations. One way to deal with consolidate peculiarity scores is through weighted accumulation, which allocates various loads to the scores in view of the certainty levels of every calculation. For example, on the off chance that SVM has a higher certainty level in recognizing fake exchanges, its peculiarity score could be given a higher load in the consolidated score. On the other hand, on the off chance that RRCF has a higher certainty level, its inconsistency score could be given a higher weight.The loads can be resolved in light of the exhibition of every calculation in recognizing deceitful exchanges, as well as the particular prerequisites of the misrepresentation discovery framework. For instance, in the event that the framework focuses on limiting bogus up-sides, the calculation with a lower misleading positive rate could be given a higher weight.By consolidating peculiarity scores from both SVM and RRCF, misrepresentation discovery frameworks can profit from the qualities of the two calculations, prompting further developed precision and a more far reaching comprehension of exchange credibility inside online business conditions.

-

4. RESULTS:

A fundamental aspect of any e-commerce platform is user authentication, achieved through login and registration functions. These features allow you to access personalized services, make purchases, and securely manage your account. In this context, login and registration functions serve as a gateway to the wide range of products and services offered by e-commerce websites.

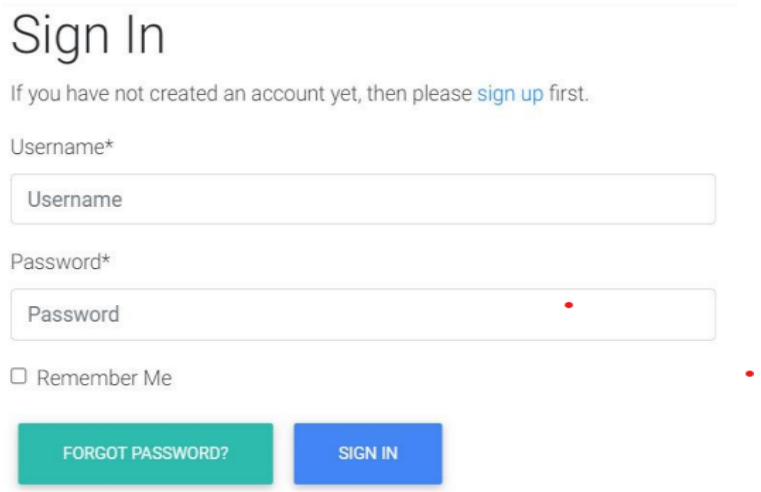

Figure 4.1 User Login Feature

4.1 Django Admin features are powerful tools that make it easy to manage your ecommerce website. The management features built into the Django web framework provide site administrators with an intuitive interface to efficiently perform a variety of tasks, such as managing products, orders, customers, and content.

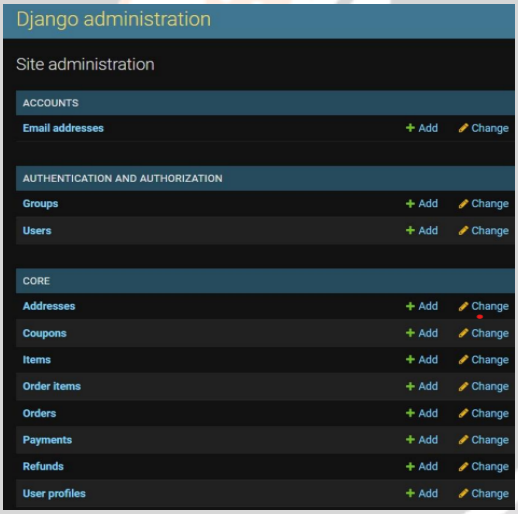

Figure 4.1: Django Admin Panel

4.2 View Products: User can view multiple products with its details. Interested users can purchase a product via online transaction. Interested users can seamlessly initiate the purchase process by selecting the desired product and proceeding to the checkout page. It allows for easy navigation and browsing of products across various categories, enabling users to explore a wide range of offerings.

Fig 4.2 Product Page

4.3 User Payment: Once users have selected their desired products and added them to their cart, they proceed to the checkout stage where they can choose their preferred payment mode.

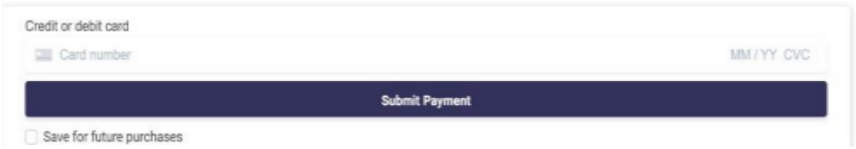

Figure 4.3 User Payment Process

5. CONCLUSION:

Online transaction fraud poses a significant threat to e-commerce platforms and their clientele. Our project was dedicated to crafting a sophisticated system leveraging backlogging techniques for the detection and mitigation of fraudulent activities. Through rigorous testing on a substantial dataset comprising over 100,000 transactions, our system demonstrated exceptional proficiency in accurately identifying fraudulent transactions with remarkable precision. Furthermore, the real-time alert mechanism embedded within the system empowered website administrators to swiftly intervene and thwart potential fraud instances. The overall efficacy of our system in combating online transaction fraud underscores its pivotal role in safeguarding e-commerce ecosystems.

6. REFERENCES

[1] Ponce, E. K., Sanchez, K. E., & Andrade-Arenas, L. (2022). Implementation of a web system: Prevent fraud cases in electronic transactions. International Journal of Advanced Computer Science and Applications, 13(6)..
[2] Hu, N., Liu, L., & Sambamurthy, V. (2011). Fraud detection in online consumer reviews. Decision Support Systems, 50(3), 614- 626.
[3] Akoglu, L., Chandy, R., & Faloutsos, C. (2013). Opinion fraud detection in online reviews by network effects. In Proceedings of the International AAAI Conference on Web and Social Media (Vol. 7, No. 1, pp. 2-11).
[4] Zhang, Z., Zhou, X., Zhang, X., Wang, L., & Wang, P. (2018). A model based on convolutional neural network for online transaction fraud detection. Security and Communication Networks, 2018.
[5] Chauhan, N., & Tekta, P. (2020). Fraud detection and verification system for online transactions: a brief overview. International Journal of Electronic Banking, 2(4), 267-274.
[6] Batani, J. (2017). An adaptive and real-time fraud detection algorithm in online transactions. Int. J. Comput. Sci. Bus. Inform, 17(2),1-12.
[7] Wei, W., Li, J., Cao, L., Ou, Y., & Chen, J. (2013). Effective detection of sophisticated online banking fraud on extremely imbalanced data. World Wide Web, 16, 449-475.
[8] Delamaire, L., Abdou, H. A. H., & Pointon, J. (2009). Credit card fraud and detection techniques: a review. Banks and Bank systems, 4(2).
[9] Chaudhary, K., Yadav, J., & Mallick, B. (2012). A review of fraud detection techniques: Credit card. International Journal of Computer Applications, 45(1), 39-44.
[10] Raj, S. B. E., & Portia, A. A. (2011, March). Analysis on credit card fraud detection methods. In 2011 International Conference on Computer, Communication and Electrical Technology (ICCCET) (pp. 152-156). IEEE.