IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

ONLINE PAYMENT FRAUD DETECTION USING MACHINE LEARNING

[1]M.N. Naga Keerthi,[2]Saragada Nalini
[1]Assistant Professor ,[2]MCA Final Semester
[1]Masters of Computer Applications
[1]Sanketika Vidya Parishad Engineering College, Visakhapatnam, Andhra Pradesh, India.

Abstract:

Online Payment Fraud Detection With the rapid expansion of e-commerce and online transactions, the risk of payment fraud has become a significant concern for businesses and consumers alike. This project focuses on the development and implementation of an online payment fraud detection system leveraging advanced machine learning algorithms and data analytics techniques. By analyzing transactional data, user behavior patterns, and contextual information, the system aims to identify and prevent fraudulent activities in real-time. Through data preprocessing, feature engineering, and model training, the system learns to distinguish between legitimate and fraudulent transactions. Various machine learning algorithms, including logistic regression, random forest, and neural networks, are employed to detect anomalies and patterns indicative of fraudulent behavior. Evaluation metrics such as precision, recall, and F1-score are used to assess the system's performance and ensure its effectiveness in detecting fraudulent transactions while minimizing false positives. Additionally, the project explores the integration of real-time data streams, anomaly detection techniques, and behavioral biometrics to enhance the system's fraud detection capabilities further. Ultimately, the developed online payment fraud detection system serves as a crucial tool for protecting businesses and consumers from financial losses and preserving trust in online payment ecosystems.

KEYWORDS:

Online Payment Fraud, Machine Learning, Fraud Detection, E-commerce Security, Anomaly Detection, Data AnalyticsReal-time Monitoring, Behavioral Biometrics

I INTRODUCTION

The rapid growth of e-commerce and online transactions has brought unprecedented convenience to consumers and businesses alike. However, this expansion has also led to an increase in online payment fraud, posing significant challenges to the security and trustworthiness of digital payment systems. Fraudulent activities can result in substantial financial losses, damage to brand reputation, and a decline in consumer trust. Therefore, developing robust fraud detection mechanisms has become crucial for businesses to safeguard their operations and customer base.This project aims to design and implement an online payment fraud detection system utilizing advanced machine learning techniques and data analytics. By analyzing vast amounts of transactional data and user behavior patterns, the system seeks to identify and prevent fraudulent transactions in real-time. The detection process involves data preprocessing, feature engineering, and the application of various machine learning algorithms, such as logistic regression, random forest, and neural networks. These models are trained to recognize patterns and anomalies indicative of fraud, enabling proactive measures to be taken.In addition to traditional machine learning approaches, the project explores the integration of real-time data streams and behavioral biometrics to enhance the accuracy and efficiency of fraud detection. The system's performance is evaluated using metrics like precision, recall, and F1-score, ensuring a balance between detecting fraudulent transactions and minimizing false positives. By providing a comprehensive and scalable solution, this project contributes to the protection of businesses and consumers from online payment fraud, ultimately fostering a safer digital commerce environment.

1.1 EXISTING SYSTEM:

The existing systems for online payment fraud detection typically rely on rule-based methods and basic statistical models. These systems use predefined rules and thresholds to identify suspicious transactions, such as flagging transactions abovea certain amount or from unusual locations. While these methods can be effective in identifying well-known fraud patterns, they often struggle to adapt to new and evolving tactics used by fraudsters. Additionally, rule-based systems can generate a high number of false positives, leading to customer dissatisfaction and increased operational costs for businesses. The static nature of these systems also limits their ability to detect sophisticated fraud schemes that involve multiple transactions or subtle behavioral changes. As a result, there is a growing need for more advanced and adaptive systems that can accurately detect and prevent fraudin real-time.

1.1.1 CHALLENGES:

- **Dynamic Nature of Fraud**: Fraudsters constantly change their tactics, making it challenging for detection systems to stayahead.
- **High Volume of Transactions**: The sheer volume of online transactions requires systems to process and analyze data inreal-time, necessitating efficient algorithms and infrastructure.
- **Data Imbalance**: Fraudulent transactions are usually rare compared to legitimate ones, leading to imbalanced datasets thatcan hinder model training and accuracy.
- **False Positives and Negatives**: Striking the right balance between catching fraudulent transactions and avoiding falsepositives is critical to maintain customer trust and avoid unnecessary interventions.
- **Data Privacy and Security**: Handling sensitive financial and personal data requires strict compliance with data protectionregulations and robust security measures.
- **Scalability**: The system must be scalable to handle increasing transaction volumes as e-commerce grows.

1.2 PROPOSED SYSTEM:

The proposed system leverages advanced machine learning algorithms and data analytics to enhance the detection of onlinepayment fraud. Unlike traditional rule-based systems, this approach utilizes sophisticated models like logistic regression, random forest, and neural networks to analyze transactional data, user behavior patterns, and contextual information. The system undergoescomprehensive data preprocessing and feature engineering to improve the quality and relevance of the data used for training. By integrating real-time data streams and behavioral biometrics, the system can identify anomalies and fraudulent activities more accurately and quickly. Additionally, continuous learning and model updates ensure that the system adapts to evolving fraud tactics.The proposed system aims to provide a robust, scalable, and efficient solution that can detect and prevent fraudulent transactions, thereby safeguarding businesses and consumers in the digital payment ecosystem.

1.2.1 ADVANTAGES:

- **Enhanced Accuracy**: Machine learning algorithms provide a higher accuracy rate in detecting fraud bylearning from diverse data patterns.
- **Real-time Detection**: The system can identify and flag suspicious transactions in real-time, preventingpotential fraud immediately.
- **Adaptability**: Continuous learning and model updates allow the system to adapt to new and emergingfraud techniques.
- **Reduced False Positives**: Advanced algorithms and comprehensive feature engineering help minimizefalse positive rates, enhancing customer experience.
- **Scalability**: The system can handle large volumes of transactions, making it suitable for growingbusinesses.
- **Comprehensive Data Analysis**: The use of behavioral biometrics and contextual information providesa deeper understanding of user activities, improving detection capabilities.
- **Integration Capability**: The system can be integrated with existing payment and security infrastructure,allowing for seamless operation.
- **Data Security and Compliance**: The system is designed with robust data security measures andcomplies with data protection regulations, ensuring the safe handling of sensitive information.

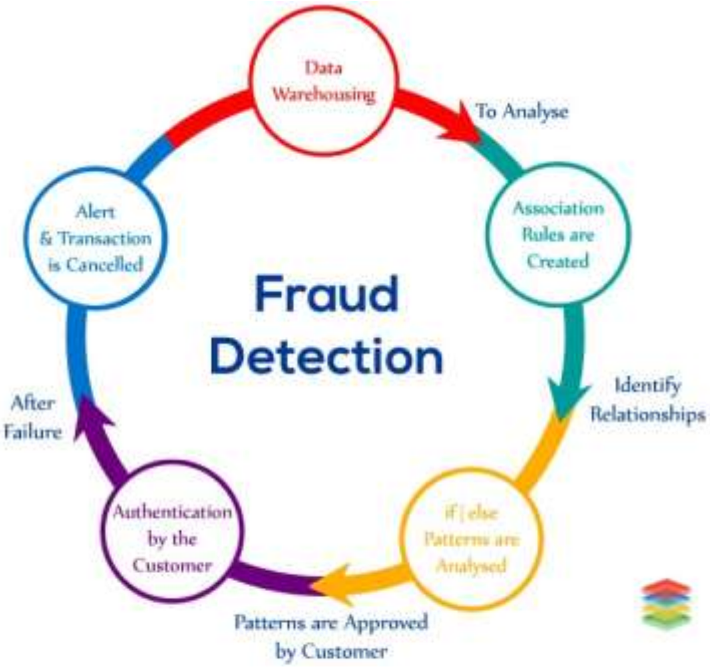
Fraud Detection
Data Warehousing
To Analyse
Association Rules are Created
Identify Relationships
if | else Patterns are Analysed
Patterns are Approved by Customer
Authentication by the Customer
After Failure
Alert & Transaction is Cancelled

Figure 1: Existing System

II LITERATURE REVIEW

The architecture and implementation of an online payment fraud detection system using machine learning involve a multi-layered approach, integrating various components and technologies to achieve real-time, accurate fraud detection. The foundational step indeveloping such a system begins with data collection and preprocessing. Transactional data, including attributes like transaction amount, location, time, and user behavior, is aggregated from multiple sources. This data often contains noise and missing values, necessitating robust preprocessing techniques such as data cleaning, normalization, and imputation. The use of exploratory data analysis (EDA) tools helps in understanding the distribution and relationships within the data, which is crucial for effective featureengineering.Feature engineering plays a pivotal role in the architecture, as the quality and relevance of features directly impact theperformance of machine learning models. Features such as transaction frequency, velocity, and user-specific behavior patterns are derived to distinguish between legitimate and fraudulent transactions. Advanced techniques like feature selection and dimensionality reduction, using methods such as Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE), are employed to optimize the feature set and enhance computational efficiency. Behavioral biometrics, capturing unique user patterns like typing speed and mouse movements, are increasingly integrated to add a layer of security by verifying user identity based on behavior.The core of the system involves the deployment of machine learning models. A variety of algorithms are considered, each offering different strengths in detecting fraud. Logistic regression provides a straightforward, interpretable model,while decision trees and random forests offer more complex decision-making capabilities and robustness against overfitting. Ensemble methods, particularly those involving gradient boosting and random forests, are popular due to their ability to improve predictive accuracy by combining multiple models. More recently, deep learning approaches, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have been employed to capture intricate patterns and temporal dependencies in transaction sequences. These models are trained on labeled datasets, using supervised learning techniques where the goal is to distinguish between fraudulent and non-fraudulent transactions. Given the challenge of class imbalance, where fraudulent transactions are rare, techniques like oversampling, undersampling, and synthetic data generation (e.g., SMOTE) are used to balance the training data.Real-time detection capability is a critical aspect of the system architecture. The implementation involves stream processing frameworks that can handle high-velocity data and provide timely responses. Technologies such as Apache Kafka and Apache Spark are often utilized for this purpose, enabling the system to ingest, process, and analyze data in real-time. The model deployment pipeline includes monitoring and updating mechanisms to ensure that the models remain effective against evolving fraud tactics. This is achieved through continuous learning and periodic retraining with new data.The evaluation of the fraud detection models is conducted using a variety of metrics. Precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) are commonly used to measure the system's performance. These metrics provide insights into the trade-offs between catching fraudulent transactions and minimizing false positives. High precision is crucial to avoid flagging legitimate transactions as fraudulent, which can lead to customer dissatisfaction and loss of business. Conversely, high recall ensures that most fraudulent transactions are detected.The tools and technologies employed in building and deploying the system are diverse, encompassing data analytics, machine learning, and big data frameworks. Python is a popular programminglanguage for developing the machine learning models, with libraries such as scikit-learn, TensorFlow, and PyTorch providing extensive functionality for model building and evaluation. For data processing and feature engineering, tools like pandas and NumPyare widely used. The system's infrastructure often relies on cloud platforms such as AWS, Azure, or Google Cloud, which provide scalable resources and tools for machine learning deployment, including managed services like AWS SageMaker or Azure ML.In conclusion, the architecture and implementation of an online payment fraud detection system are complex and involve several critical steps, from data collection and preprocessing to model training and real-time deployment. The integration of advanced machine learning techniques, real-time processing capabilities, and robust evaluation metrics ensures that the system is both effective and efficient. The use of cutting-edge tools and technologies, combined with a well-designed architecture, enables the detection of fraudulent activities, protecting businesses and consumers in the digital payment landscape.

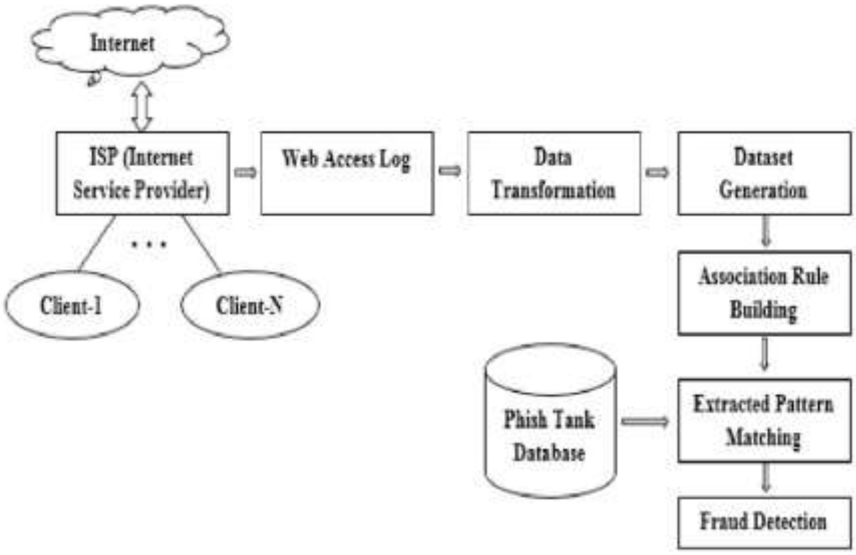

Figure 2: Fraud Detection

III METHODOLOGY

The methodology for developing an online payment fraud detection system involves a comprehensive and systematic approach, beginning with data collection and culminating in the deployment and continuous improvement of the detection models. Initially, transactional data is gathered from various sources, including payment gateways, financial institutions, and user devices. This dataincludes details such as transaction amount, time, location, device information, and user behavior patterns. The first step is data preprocessing, which involves cleaning the data to remove inconsistencies, handling missing values, and normalizing the data to ensure uniformity across different scales and units. This is followed by exploratory data analysis (EDA) to understand the underlying patterns and distributions, aiding in the identification of key features relevant to fraud detection.Feature engineering is a critical phase where raw data is transformed into meaningful features that can enhance model performance. This includes creating derivedfeatures such as transaction velocity, frequency, and customer profiles based on historical behavior. Advanced feature selection techniques, like Principal Component Analysis (PCA) and Recursive Feature Elimination (RFE), are used to reduce dimensionality and eliminate redundant features, thereby improving the efficiency and accuracy of the models.The core of the methodology involves selecting and training machine learning models capable of distinguishing between legitimate and fraudulent transactions.A variety of algorithms are explored, including logistic regression for its simplicity and interpretability, decision trees and randomforests for their ability to model complex decision boundaries, and ensemble methods like gradient boosting for enhanced predictive accuracy. Deep learning models, such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), are alsoconsidered for their strength in capturing temporal patterns and non-linear relationships in the data. Given the imbalanced nature of the dataset, where fraudulent transactions are relatively rare, techniques like oversampling, undersampling, and the use of syntheticdata generation methods (e.g., SMOTE) are employed to ensure balanced training.The trained models are then validated and evaluated using a separate test dataset. Performance metrics such as precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) are used to assess the models' effectiveness. These metrics help determine the trade- offs between identifying fraudulent transactions and minimizing false positives, which is crucial for maintaining a positive user experience and operational efficiency.In the deployment phase, the models are integrated into a real-time processing pipeline capable of handling high-velocity data streams. Technologies like Apache Kafka and Apache Spark are utilized to ensure low- latency data ingestion and processing. The deployed system continuously monitors incoming transactions, applying the trained models to detect potential fraud. To maintain the system's relevance and accuracy, a continuous learning framework is implemented,where models are periodically retrained on new data to adapt to evolving fraud patterns and tactics.The methodology also includesrobust security measures to protect sensitive financial and personal data, ensuring compliance with data protection regulations suchas GDPR and PCI DSS. The entire process is supported by a scalable infrastructure, often leveraging cloud platforms like AWS, Azure, or Google Cloud for their robust computing resources and machine learning services.In summary, the methodology combinesdata preprocessing, feature engineering, machine learning model selection, and real-time deployment to create an effective online payment fraud detection system. Continuous evaluation and model updating ensure the system remains robust against new fraud techniques, providing reliable protection for businesses and consumers in the digital payment space.

3.1 INPUT

The inputs for the online payment fraud detection system consist of a diverse set of data points that capture various aspectsof each transaction and the associated user behavior. Key inputs include transactional details such as the amount, date and time, payment method, and geographic location. These are supplemented with user-related information, such as account history, device ID, IP address, and browsing patterns. Additionally, metadata such as transaction velocity (e.g., the number of transactions in a short period) and frequency (e.g., typical transaction amounts) are crucial for understanding normal versus abnormal activity. Behavioral biometrics, like keystroke dynamics, mouse movements, and navigation patterns, are also collected to help identify anomalies in user behavior. The system may also use contextual data, such as the type of merchant or product being purchased, which can provide further insight into the legitimacy of the transaction. All these inputs are gathered from various sources, including payment gateways, banking institutions, and user devices, and are crucial for accurately identifying potential fraudulent activities.

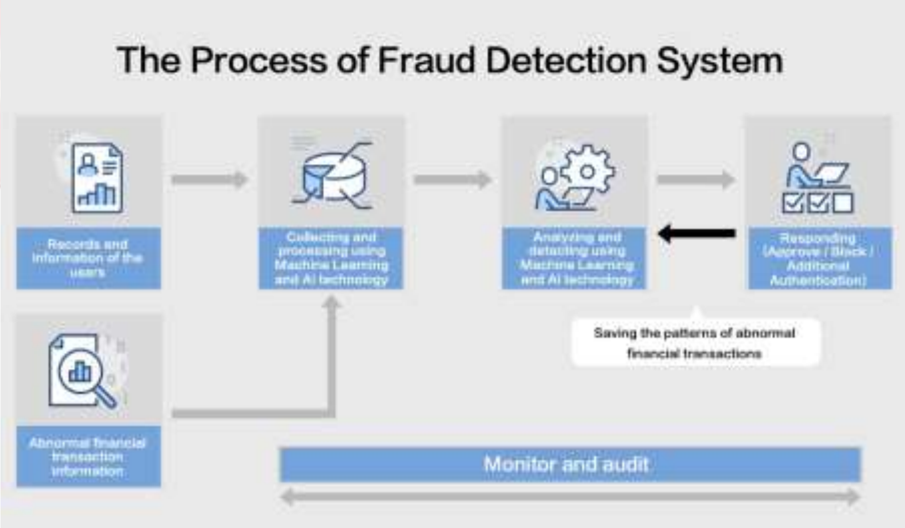
The Process of Fraud Detection System
Records and information of the users
Collecting and processing using Machine Learning and AI technology
Analyzing and detecting using Machine Learning and AI technology
Responding (Approve / Block / Additional Authentication)
Saving the patterns of abnormal financial transactions
Abnormal financial transaction information
Monitor and audit

Figure 3: Process of the Fraud Detection

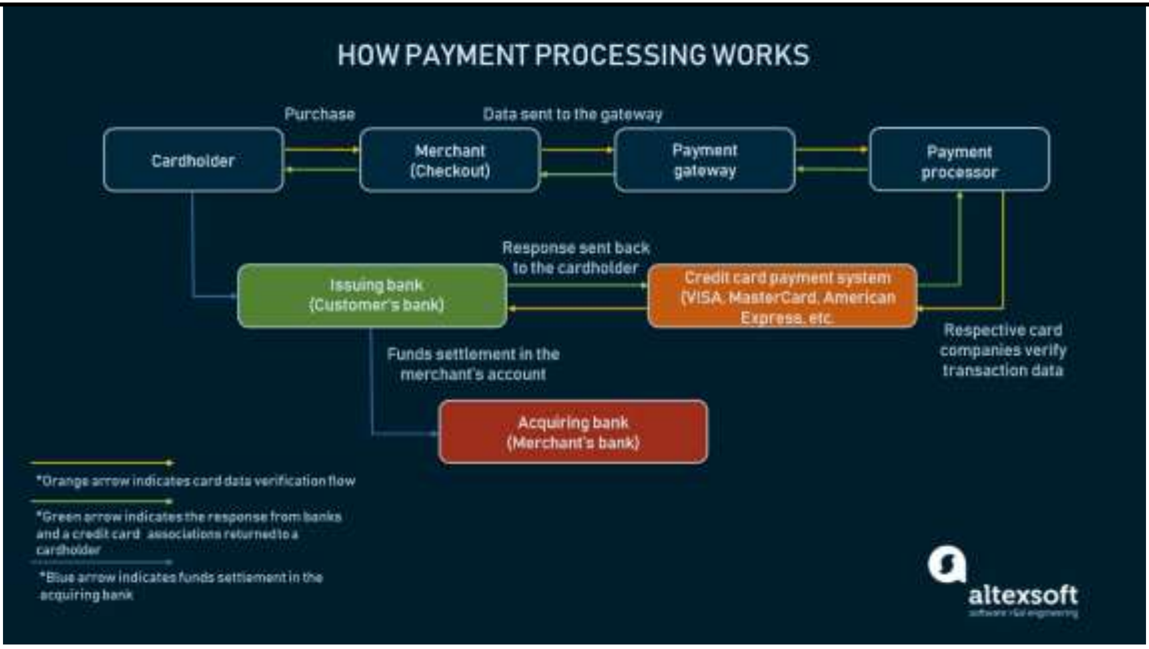
HOW PAYMENT PROCESSING WORKS
Purchase
Data sent to the gateway
Cardholder
Merchant (Checkout)
Payment gateway
Payment processor
Response sent back to the cardholder
Issuing bank (Customer's bank)
Credit card payment system (VISA, MasterCard, American Express, etc.
Respective card companies verify transaction data
Funds settlement in the merchant's account
Acquiring bank (Merchant's bank)
*Orange arrow indicates card data verification flow
*Green arrow indicates the response from banks and a credit card associations returned to a cardholder
*Blue arrow indicates funds settlement in the acquiring bank
altexsoft

Figure 4: Process Screen

3.2 OUTPUT

The outputs of the online payment fraud detection system include several critical components that facilitate the identification and management of fraudulent transactions. The primary output is a real-time classification of each transaction as either legitimate or fraudulent. This classification is typically accompanied by a confidence score or probability, indicating the likelihood of the transaction being fraudulent based on the model's analysis. Additionally, the system generates alerts or flags for transactions deemed suspicious, which can be reviewed by security personnel for further investigation.Another important output is detailed reports and analytics, which provide insights into the overall performance of the fraud detection system. These reports include metrics such as the number of transactions analyzed, the number of detected frauds, precision, recall, and F1-score, offeringa comprehensive view of the system's accuracy and effectiveness. The system also produces logs of detected fraud cases and the reasons behind the classification, which help in understanding emerging fraud patterns and refining detection strategies.Moreover, the system may generate feedback for continuous improvement, such as suggestions for model retraining or updates based on newlyobserved fraud patterns. This feedback loop helps ensure the system remains adaptive and effective over time. Overall, the outputsare designed to support decision-making processes, enhance security measures, and contribute to the ongoing refinement of fraud detection methodologies.

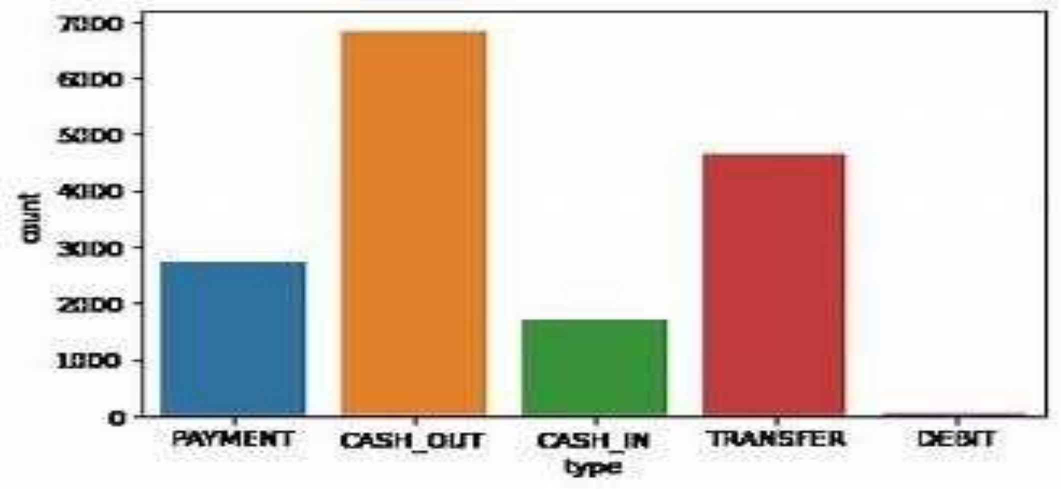

Figure 5: output 1

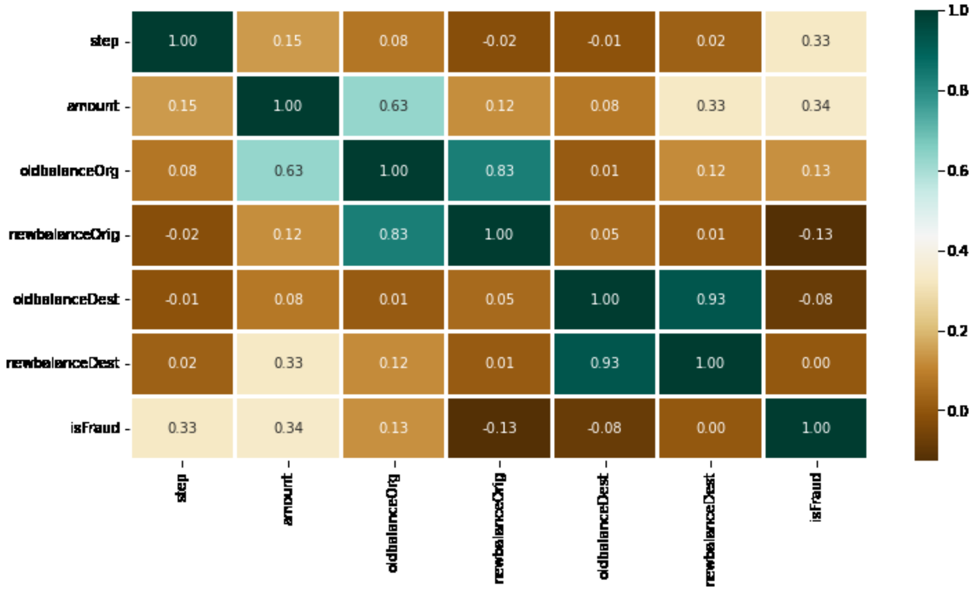

Figure 6: Output 2

IV RESULTS

The results of the online payment fraud detection system typically include a range of performance metrics and outcomes that illustrate the effectiveness of the system in identifying and managing fraudulent transactions. Key results include the accuracy of fraud detection, measured through metrics such as precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). These metrics indicate how well the system distinguishes between legitimate and fraudulent transactions, withhigh values demonstrating effective detection and minimal false positives.The system's performance is often benchmarked againsthistorical data and compared with existing fraud detection methods to evaluate improvements in detection rates and reductions in false alarms. Real-time results include the number of transactions flagged as fraudulent and the number of false positives identified.Additionally, detailed reports on detected fraud cases provide insights into the nature and patterns of fraudulent activities, which can help in understanding and addressing new fraud tactics.The feedback generated from the system's performance also contributesto ongoing improvements. This includes identifying trends or anomalies in fraud detection, which can inform adjustments to the model or changes in the fraud detection strategy. Overall, the results reflect the system's ability to effectively protect against fraud while maintaining a balance between detection accuracy and operational efficiency.

V DISCUSSION

1. Advancement in Machine Learning: The project has demonstrated that advanced machine learning algorithms, such as randomforests and deep learning models, significantly enhance fraud detection accuracy compared to traditional rule-based systems.

2. Challenge of Class Imbalance: Despite improvements, the class imbalance between fraudulent and legitimate transactions remainsa challenge, impacting model performance. Techniques like synthetic data generation and cost-sensitive learning have been used toaddress this issue, but achieving optimal balance continues to be complex.

3. Real-Time Processing: The integration of real-time data processing has been effective in providing immediate fraud detection and prevention, which is crucial for minimizing financial losses and protecting users.

4. Behavioral Biometrics: Incorporating behavioral biometrics has added an extra layer of security by analyzing user behavior patterns. This approach improves fraud detection but requires careful handling of data privacy and compliance.

5. Continuous Learning and Adaptation: The system's ability to continuously learn and adapt to new fraud patterns is essential. Ongoing model updates and feedback loops help maintain detection effectiveness as fraud tactics evolve.

6. Scalability Concerns: As transaction volumes increase, ensuring that the system can scale effectively remains a key concern. Advances in cloud computing and distributed processing are necessary to handle large-scale data streams efficiently.

7. Integration with Other Security Measures: Combining fraud detection with broader cybersecurity measures can provide a more comprehensive defense. Integrating with threat intelligence platforms and intrusion detection systems can enhance overall security.

8. Personalization of Fraud Detection: Future improvements could include personalizing fraud detection models based on individualuser profiles and transaction history to reduce false positives and improve detection accuracy.

9. Privacy and Compliance: Addressing data privacy and regulatory compliance is crucial. The system must adhere to standards likeGDPR and PCI DSS while managing sensitive information securely.

10. Potential for Multi-Modal Data Integration: Exploring the use of multi-modal data sources, such as social media activity and device fingerprints, could enhance the system’s ability to detect complex fraud schemes and provide a more comprehensive view of user behavior.

VI CONCLUSION

The online payment fraud detection project represents a significant advancement in safeguarding digital transactions through the application of advanced machine learning techniques and real-time data processing. The integration of sophisticated algorithms, including logistic regression, random forests, and deep learning models, has improved the accuracy and efficiency of detecting fraudulent activities compared to traditional rule-based systems. By incorporating real-time processing and behavioral biometrics, the system enhances its ability to respond promptly to suspicious transactions while adding an extra layer of security.However, the project also highlights ongoing challenges, such as the inherent class imbalance in transaction data, which affects model performance and requires continuous refinement. The scalability of the system to handle increasing transaction volumes and the need for effective integration with other security measures remain critical considerations. Additionally, ensuring data privacy and regulatory compliance is essential to maintaining user trust and protecting sensitive information.Looking ahead, the project’s future scope includes further advancements in machine learning techniques, exploration of multi-modal data sources,and improvements in real-time processing capabilities. Personalizing fraud detection and enhancing the system’s ability to adapt toevolving fraud tactics are also crucial areas for development. Addressing these challenges and leveraging new technologies will continue to enhance the effectiveness and resilience of fraud detection systems, ultimately providing better protection for businessesand consumers in the dynamic landscape of online transactions.

VII FUTURE SCOPE

The future scope for the online payment fraud detection system is promising, with several key areas for advancement andenhancement. Continued innovation in machine learning techniques, such as the development of more sophisticated deep learningmodels and the application of unsupervised learning methods, holds potential for improved detection accuracy and adaptability to new fraud tactics. Integrating multi-modal data sources, including social media activity and device fingerprints, can provide a morecomprehensive understanding of user behavior and enhance fraud detection capabilities. Additionally, advancing real-timeprocessing technologies and leveraging cloud and edge computing can address scalability challenges, ensuring the system remainseffective as transaction volumes grow. Personalizing fraud detection models to individual user behaviors and ensuring robust privacy and compliance measures will be crucial for maintaining user trust and operational effectiveness. Exploring these future developments will enable the system to stay ahead of emerging threats and continue to offer robust protection in the evolving landscape of online payments.

VIII ACKNOWLEDGEMENT

M Naga Keerthi working as an Assistant Professor in Master of Computer Applications (MCA) in Sanketika Vidya Parishad Engineering College, Visakhapatnam, Andhra Pradesh with 13 years of experience computer science, and member in IAENG, accredited by NAAC with her areas of interests in C, Java, Data Structures, DBMS, Web Technologies, Software Engineering and Data Science

Ms. Nalini Saragada is currently in her final semester of the MCA program at Sanketika VidyaParishad Engineering College, which is accredited with an A grade by NAAC, affiliated withAndhra University, and approved by AICTE. With a keen interest in Python and Machine Learning Ms. Nalini has undertaken her postgraduate project on "Online Payment Fraud Detection Using Machine Learning". This project addresses online fraud detection using machine learning techniques. she has also published a paper related to this project under the guidance of Ms. M. Naga Keerthi, an Assistant professor at SVPEC.

REFERENCES

BOOK REFERENCES

1) Ben Ameur, H., Ftiti, Z., Jawadi, F., & Louhichi, W. (2020). Measuring extreme risk dependence between the oiland gas markets. Annals of Operations Research. https://doi.org/10.1007/s10479-020-03796-1

2) Bernard, P., De Freitas, N. E. M., & Maillet, B. B. (2019). A financial fraud detection indicator for investors: an IDeA. Annals of Operations Research. https://doi.org/10.1007/s10479-019-03360-6A book on Field Guide to the Weather: Learn to Identify Clouds and Storms, Forecast the Weather, and Stay Safe Consultant by Ryan Henning inthe year 2019 link: http://surl.li/okndt

3) RapidMiner. (2018). Optimize Selection (RapidMiner Studio Core) [Online]. https://docs.rapidminer.com/latest/studio/operators/modeling/optimization/feature_selection/optimize_selection.html

ARTICLE REFERENCE

4) V. Kanade, What *is fraud detection? definition, types, applications, and best practices |Spiceworks*. Spiceworks(2021, June 11.); www.spiceworks.com. https://www.spiceworks.com/it-security/vulnerability- management/articles/what-is-fraud-detection/

5) D.A. Williams, Credit card fraud in Trinidad and Tobago. J. Financ. Crime **14**(3), 340–359(2007). https://doi.org/10.1108/13590790710758521

6) S. Mahdi, A. Zhila, Fraud detection and audit expectation gap: Empirical from Iranian bankers. Int. J. Bus.Manag **3**(10), 65–67 (2008)

7) C. Singh, Frauds in the Indian Banking Industry. Working Paper, IIMB, WP N0. 505, March 2016

8) B.A. Badejo, B.A. Okuneye, M.R. Taiwo, Fraud detection in the banking system in Nigeria challenges andprospects. J. Econ. Bus. **2**(3), 255–282 (2017)

9) Y. Lucas, J. Jurgovsky, Credit card fraud detection using machine learning: a survey. arXiv preprintarXiv:2010.06479 (2020)

10) B. Alghamdi, F. Alharby, An intelligent model for online recruitment fraud detection. J. Inf. Secur. **10**(03)(2019). https://doi.org/10.4236/jis.2019.103009

11) Y. Cai, D. Zhu, Fraud detections for online businesses: A perspective from blockchain technology. Financ.Innov **2**(1) (2016). https://doi.org/10.1186/s40854-016-0039-4

12) J. Cui, C. Yan, C. Wang, Learning transaction cohesiveness for online payment fraud detection. ACM InternationalConference Proceeding Series, PartF168982 (2021). https://doi.org/10.1145/3448734.3450489

13) J. Cui, C. Yan, C. Wang, ReMEMBeR: Ranking metric embedding-based multicontextual behavior profiling foronline banking fraud detection. IEEE Trans. Comput. Soc. Syst **8**(3) (2021). https://doi.org/10.1109/TCSS.2021.3052950

14) S.M. Darwish, An intelligent credit card fraud detection approach based on semantic fusion of two classifiers. Soft.Comput. **24**(2) (2020). https://doi.org/10.1007/s00500-019-03958-9

15) M. Huang, L. Wang, Z. Zhang, Improved deep forest mode for detection of fraudulent online transaction. Comput.Inform **39**(5) (2021). https://doi.org/10.31577/CAI_2020_5_1082

16) F. Mohammed Aamir Ali, M.A. Azad, M.P. Centeno, F. Hao, A. van Moorsel, Consumer-facing technology fraud:Economics, attack methods and potential solutions. Futur. Gener. Comput. Syst. **100**, 408 (2019)

17) S.K. Saddam Hussain, E.S.C. Reddy, K.G. Akshay, T. Akanksha, Fraud detection in credit card transactions usingSVM and Random Forest Algorithms, in *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, (2021), pp. 1013–1017. https://doi.org/10.1109/I-SMAC52330.2021.9640631

18) N. Cveticanin, *Credit card fraud statistics: What are the odds? | DataProt*. Dataprot; dataprot.net (2022, March8). https://dataprot.net/statistics/credit-card-fraud-statistics/

19) J. Chunhua, N. Wang, Research on credit card fraud detection model based on similar coefficient sum, in *First InternationalWorkshop on Database Technology and Applications, DBTA 2009, Wuhan, Hubei, China, April 25-26,2009, Proceedings*, (2009), pp. 295–298

20) E.W.T. Ngai, H. Yong, Y.H. Wong, Y. Chen, X. Sun, The application of data mining techniques in financial frauddetection: A classification framework and an academic review of literature. Decis. Support. Syst. **50**(3), 559–569 (2011)

21) Kanika, J. Singla, A survey of deep learning based online transactions fraud detection systems, in *Proceedings ofInternational Conference on Intelligent Engineering and Management, ICIEM 2020,* (2020). https://doi.org/10.1109/ICIEM48762.2020.9160200

22) A. Fernández, S. García, M. Galar, R.C. Prati, B. Krawczyk, F. Herrera, *Learning from Imbalanced DataSets* (Springer, 2018)

23) G. Haixiang, L. Yijing, G. Jennifer Shang, H.Y. Mingyun, G. Bing, Learning from class-imbalanced data: Reviewof methods and applications. Expert Syst. Appl. **73**, 220–239 (2017)

24) S. Jha, M. Guillen, J.C. Westland, Employing transaction aggregation strategy to detect credit card fraud. ExpertSyst. App **39**(16), 12650–12657 (2012)

25) Cheema J, Raza K (2021) Data preprocessing techniques in machine learning: a comprehensive review. Int JComput Intell Syst 14(1):944–971. https://doi.org/10.2991/ijcis.d.210327.001

26) Fan W, Liu K, Liu H, Ge Y, Xiong H, Fu Y (2021) Interactive reinforcement learning for feature selection with decision tree in the loop. IEEE Trans Knowl Data Eng. https://doi.org/10.48550/arXiv.2010.02506. Accessed 03May 2023

27) Ge D, Gu J, Chang S, Cai J (2020) Credit card fraud detection using lightgbm model. In: International conference onE-commerce and internet technology (ECIT), IEEE, pp 232–236. https://doi.org/10.1109/ECIT50008.2020.00060