International Journal On Engineering Technology and Sciences – IJETS
ISSN (P):2349-3968, ISSN (O): 2349-3976 Volume XI - Issue III, March - 2024

UPI FRAUD DETECTION USING MACHINE LEARNING

[1]Axlin Jenes J F, [2]Mahendran N, [3]Siluvai Priyan S, [4]Vivek Siddharath, [5]Sajil L
[1,2,3,4]Student, [5]Assistant Professor,
Department of Information Technology,
Loyola Institute of Technology and Science, Nagercoil, Tamilnadu.

Abstract : The project focuses on the development of a machine learning model that can analyze UPI transaction data in real-time to identify fraudulent activities. The primary objective is to create a system that enhances the security of UPI transactions and reduces financial losses due to fraud. With the advancement of technology, today most of the modern commerce is relying upon the online banking and cashless payments. Due to adaption of online payment among businesses, the fraud cases are also increasing which cause financial losses to them. Fraudsters are inventing new techniques to perform fraudulent transaction which seem legitimate. Hence, there is an urgent need to develop fraud detection measures which can deal with these fraudsters on real time basis. Deep learning techniques have the capability to detect these fraudulent transactions efficiently and has a huge scope in fraud detection. However, there are many challenges faced by the researchers in online transactions fraud detection because the datasets are not publicly available due to privacy issue of the financial institutions as customers data is sensitive and it can be misused and the datasets which are available are imbalanced. This paper presents a review of deep learning techniques used for online transactions fraud detection. It also provides the information about datasets used by the researchers and the results achieved by them in their research work.

Index Terms - Fraud Detection**,** Feature Engineering, Cross-Validation, Emerging Technologies.

I. Introduction

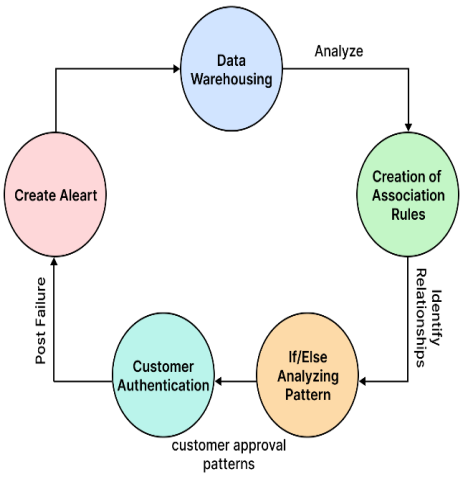
The emergence of online banking has caused a paradigm shift in the current financial transaction landscape, providing individuals and organisations globally with unmatched ease and efficiency. But there are drawbacks to this digital transformation as well, the most significant of which is the growing frequency of fraudulent activity in online transactions. The swift expansion of online banking services has made it easier for bad actors to take advantage of weaknesses, which puts the security and integrity of financial systems at serious risk. Furthermore, the COVID-19 pandemic's start has acted as a trigger, quickening the shift to remote operations and raising the possibility of fraudulent activity in the digital sphere. It is therefore more important than ever to create reliable fraud detection systems in the face of the epidemic highlight how important it is for both customers and financial institutions to strengthen their defences against fraud. The increasing trend of financial transactions occurring on digital platforms underscores the need for advanced security measures and flexible approaches to protect the integrity of online banking systems.

Exiting System

Regarding the field of online banking fraud detection, the systems that are now in place primarily depend on conventional techniques and rule-based methods. In order to identify potentially fraudulent transactions, these systems frequently use static rules and thresholds, usually based on established patterns or anomalies. Although these systems have shown some degree of effectiveness, they have shortcomings in terms of accuracy, scalability, and adaptability, especially when it comes to sophisticated and ever-evolving fraud techniques.

A prevalent obstacle within the current framework is its dependence on static rules, which may not be able to identify subtle or evolving patterns that point to fraudulent activity. Furthermore, the rule-based approach frequently finds it difficult to manage the subtleties and intrinsic complexity of transaction data, particularly given the ever-changing environment of online banking transactions. The current system's vulnerability to false positives and false negatives is another drawback. Static rules have the potential to unintentionally identify genuine transactions as fraudulent (false positives) or fail to identify actual fraudulent activity (false negatives), which can result in inefficiencies, unhappy customers, and financial losses for financial institutions as well as consumers. Furthermore, particularly in the context of the Unified Payments Interface (UPI), the current systems would find it difficult to handle the enormous volume and variety of transaction data generated in real-time. The scalability and computing efficiency needed to efficiently process and analyse large-scale transaction datasets may be lacking in traditional methodologies.

II. PROPOSED SYSTEM:



We present a unique strategy that uses convolutional neural networks (CNNs) for increased fraud detection in online banking transactions in order to overcome the shortcomings of current fraud detection systems. Compared to conventional rule-based approaches, our suggested solution has various advantages, including better accuracy, flexibility, and scalability. We support the use of convolutional neural networks (CNNs) as the primary technology for online banking transaction fraud detection. Especially in image analysis tasks, CNNs have shown impressive skills in feature extraction and pattern detection. We intend to leverage CNNs' capacity to automatically learn hierarchical characteristics and identify complex patterns suggestive of fraudulent activity by applying them to transactional data. Important elements of the system we've suggested include: Adaptive Learning: Unlike static rule-based systems, our suggested approach makes use of CNN-enabled adaptive learning techniques. With the ability to dynamically modify their internal representations and adjust their parameters in response to incoming data, these neural networks are able to react in real-time to evolving fraud trends and new threats. Its ability to adjust strengthens the system's defences against changing fraud strategies and guarantees steady advancement over time.

Feature extraction and transformation: To extract pertinent characteristics from transactional data and turn them into interpretable representations for fraud detection, our suggested approach makes use of CNNs. CNNs can identify subtle patterns and abnormalities that may escape conventional rule-based methods by automatically learning discriminative features from raw transactional attributes. The method of feature extraction improves the accuracy with which the system can distinguish between authentic and fraudulent transactions.

Real-time Processing: Our suggested solution allows for real-time fraud detection and response in high-volume transaction environments like the Unified Payments Interface (UPI) by utilising the parallel processing powers of CNNs. CNNs streamline the simultaneous processing of massive amounts of transaction data, making it easier to identify fraudulent activity quickly and take prompt action to prevent losses. Despite the inherent complexity of CNNs, we have given priority to interpretability and openness in our suggested solution. We include methods for illustrating and interpreting CNN decision-making processes so that interested parties can comprehend the reasoning behind reported transactions and develop faith in the dependability of the system. Our technology facilitates cooperation between users, regulators, and internal auditors by improving interpretability, which in turn promotes responsibility and adherence to regulatory standards.

Continuous Monitoring and Evaluation: To evaluate the efficacy and performance of our suggested system over time, it is equipped with mechanisms for ongoing monitoring and assessment. We make sure that the fraud detection system is continuously optimised and improved by examining user feedback, comparing model predictions to ground truth labels, and monitoring important performance metrics like precision, recall, and F1-score. The overall effectiveness and dependability of the system are increased by this iterative process, which permits ongoing enhancement and adaptability to changing fraud environments.

Advantages

- Enhanced accuracy
- Adaptability to changing patterns
- Continuous improvement
- Comprehensive fraud detection
- Scalability and efficiency
- Real time detection

III. Results and Discussion

The evaluation of our implemented models has provided valuable insights into their performance. Notably, the Feedforward Neural Network (FNN) and Convolutional Neural Network (CNN) have exhibited commendable accuracy in test sets, showcasing their ability to accurately classify transactions. However, the relatively low average precision scores for both models indicate potential challenges with false positive rates. This emphasizes the significance of considering precision-recall trade-offs, especially in the context of imbalanced datasets like ours.

In contrast, the machine learning models, including Decision Tree, Naive Bayes, Logistic Regression with L1 and L2 regularization, and K-Nearest Neighbors (KNN), have demonstrated robust overall performance

with varying precision scores. Particularly noteworthy is Logistic Regression with L1 regularization (LR1), which stands out by achieving the best overall accuracy among all algorithms. This remarkable accuracy underscores LR1's effectiveness in correctly identifying fraud cases with minimal false positives. Such precision is of paramount importance in fraud detection, where minimizing false positives is crucial to avoid inconveniencing legitimate users. These results provide actionable insights for refining fraud detection strategies in real-world applications. For instance, the outstanding precision achieved by LR1, along with the perfect precision scores attained by Decision Tree, Naive Bayes, and KNN models, suggests their suitability for deployment in scenarios where false positives must be minimized. On the other hand, while FNN and CNN models demonstrate high accuracy, further optimization or consideration of additional factors may be necessary to address potential challenges with false positives.

Moreover, in the context of user interface design, presenting these insights in a comprehensible manner becomes essential. A user-friendly interface should provide clear visualizations and summaries of the model performance, emphasizing the trade-offs between accuracy and precision. This aids decision-makers in selecting models aligned with the specific requirements of the application. Additionally, incorporating user feedback mechanisms into the interface can enhance the adaptability of the system, allowing for iterative improvements based on real-world usage and evolving fraud patterns. Overall, the integration of insightful model evaluations and a user-friendly interface forms a cohesive strategy for refining and deploying effective fraud detection systems in practical applications.

IV. Conclusion

We have gone through several phases of data collection, pre-processing, algorithm selection, and system implementation in this extensive effort to create a strong fraud detection system. The end result is a solution that has the potential to greatly improve the security and dependability of financial transactions. We started our trip by obtaining a large and comprehensive dataset that included complex transactional information, which served as the foundation for our later investigations.

Setting up a strong basis for later model training required the first stage of data pre-processing. We carefully addressed issues, including the dataset's unequal class distribution, using calculated pre-processing techniques to reduce biases and guarantee the stability of our models. Methods like standardization based feature scaling and careful treatment of missing data played a crucial role in getting the dataset ready for efficient model training.

The selection of algorithms was an important factor in guaranteeing the effectiveness of the system in detecting fraudulent actions. We made sure our system could detect a wide range of fraudulent patterns by implementing a diverse ensemble of machine learning algorithms, from more sophisticated techniques like convolutional neural networks to more conventional methods like logistic regression and decision trees. With the contributions of each algorithm, a comprehensive fraud detection system that could handle a variety of fraud scenarios was created.

The system architecture was crucial in guaranteeing scalability, efficiency, and usability during the development and deployment stages. By using technologies like Flask, HTML, CSS, Python, and other programming languages, we were able to smoothly integrate the fraud detection system into the current financial infrastructure, increasing its usability and accessibility for both stakeholders and end users. The seamless deployment and functioning of the system in real-world settings were made possible by this harmonious technology integration.

The fraud detection system is a major improvement in protecting financial transactions from fraudulent activity, as we can see when we consider the results of our work. It is crucial to preserving the integrity and reliability of financial systems because of its capacity to precisely detect fraudulent transactions while reducing false positives. But our adventure doesn't end here. Maintaining protection and security in the ever-changing world of financial transactions will require constant system improvement and refinement to meet changing fraud trends and new security risks. We are prepared to meet upcoming challenges and protect the integrity of financial systems around the globe with a dedication to innovation and vigilance.

REFERENCES

1. ALESKEROV E, FREISLEBEN, B., and, RAO B (1997) CARDWATCH: A neural network-based database mining system for credit card fraud detection. In Conference (pp. 220–226). IEEE, Piscataway, NJ

2. Sahin M (2017) Understanding Telephony Fraud as an Essential Step to Better Fight it [Thesis]. École Doctorale Informatique, Télécommunication et Électronique, Paris

3. Abdallah A, Maarof MA, Zainal A (2016) Fraud detection system: A survey. J Netw Comput Appl 68:90–113

4. ANDREWS PP, PETERSON MB (eds) (1990) Criminal Intelligence Analysis. Palmer Enterprises, Loomis, CA

5. ARTÍS M, AyUSO M, GUILLÉN M (1999) Modeling different types of automobile insurance fraud behavior in the Spanish market. Insurance Math Econ 24:67–81

6. BARAO MI, TAWN JA (1999) Extremal analysis of short series with outliers: Sea-levels and athletics records. Appl Stat 48:469–487

7. BLUNT G, HAND DJ (2000) The UK credit card market. Technical report, Department of Mathematics, Imperial College, London.

8. BOLTON RJ, HAND DJ (2001) Unsupervised profiling methods for fraud detection. In Conference on Credit Scoring and Credit Control 7, Edinburgh, UK, 5–7 Sept

9. Phua C, Lee V, Smith K, Gayler R (2010) A comprehensive survey of data mining-based fraud detection research. https://doi.org/10.48550/ARXIV.1009.6119

10. Summers SL, Sweeney JT (1998) Fraudulently misstated financial statements and insider trading: An empirical analysis. 73(1):131– 146https://www.jstor.org/stable/248345.