



A SURVEY PAPER ON – FRAUD DETECTION IN BANKING SYSTEM USING MACHINE LEARNING

PROF.SNEHA TIRTH, Trinity College Of Engineering And Research, Pune-48, Maharashtra, India

PROF.SAI TAKAWALE, Trinity College Of Engineering And Research, Pune-48, Maharashtra, India

VIRAJ BHALEKAR, Trinity College Of Engineering And Research, Pune-48, Maharashtra, India

PRANALI SAKPAL, Trinity College Of Engineering And Research, Pune-48, Maharashtra, India

SAURABH JAGTAP, Trinity College Of Engineering And Research, Pune-48, Maharashtra, India

SHRAMIK RATHOD Trinity College Of Engineering And Research, Pune-48, Maharashtra, India

❖ ABSTRACT

In today's world, online payment are mostly used making it easy for the fraudsters to cause fraud in banking. Various security techniques are involved in this. Machine learning can be de-facto in order to detect the fraud efficiently. Machine Learning/Deep Learning techniques can be used for fraud detection in better way for those attackers who surpasses basic cryptographic methods. The results suggest that in terms of standard classification measures, the proposed semi-supervised ensemble model integrate multiple unsupervised outlier detection algorithms such as XGBoost classifier which achieves the best results, while the highest cost savings. It can be achieved by combining random under-sampling and XGBoost methods. This study has therefore an the financial implications and organizations to make appropriate decisions to regarding the implementation of effective fraud detection systems.

Keywords:

Credit Card Mobile payment Fraud detection Machine learning Imbalanced data Outlier detection

❖ I - INTRODUCTION

Now-a-days mobile payment has become one of the mainstream payment methods. Thousands of transactions are carried out on the online trading platform all the time. Personal property in the complex network environment has the risk of theft, which not only damages the interests of consumers, but Nowadays, the number of smartphones are increasing, due to this online payment systems are becoming more popular, which in turn attracts more fraudsters. For this extensive research is required for detection of fraud using machine learning techniques. In [3] The author has proposed the extra boost based framework for detection of fraud in financial transactions. For efficiency of this framework a comparative study of machine learning techniques of modelling and outlier detection was done.

In [3] The author has used various resources and methods like risk models, algorithms, human action, tools, web technology tools and business system in risk management. Now-a-days, online fraud detection is a difficult problem which requires a great understanding to deal with the large data set.

Therefore, the transaction fraud detection is one of the key and tools to solve the problem of network transaction fraud. That are the Since they are verification techniques, it is difficult to obtain the laws hidden behind the fraud transaction data. The big data technology and machine learning algorithm provide efficient detection from methods for transaction fraud detection [2]. The data received that here are unbalanced data so to detect fraud in efficient way [1] has mentioned the ML/DL methods. Here Bayesian optimization methods are used to optimize the hyperparameters. The pre-processing is done here through weight-tuning method. To uplift the pre-processing, they added CATboost and XGBoost method. In order to fine tune the hyperparameters they further use deep learning methodologies. While using the machine methods the unbalanced data-sets need to be preprocessed where feature selection is wisely improve the efficiency of fraud detection. In [2] the author used the marine predator algorithm for feature selection.

In this machine learning project, we solve the problem of transactions fraud detecting using machine numpy, scikit learn, & few other python libraries that are learn . We overcome the problem can be creating a binary classifier and experimenting with machine learning techniques to see which fits better.

❖ II - OBJECTIVE

- We are not only detecting the fraud in transactions but we are blocking the whole transaction.
- We can also use Human Intelligence for better context and insights.
- Interpretability is needed so that all the people can use this system.
- We can prevent financial loss of users.
- Risk of non-compliance is reduced

Risk Assessment: Analyzing the data to identify potential risks, predict default probabilities, and enhance risk management strategies.

Fraud Detection: Using data analytics to detect unusual patterns and behaviors that could indicate fraudulent activities, helping banks prevent financial losses.

Customer Insights: Gaining a deeper understanding of customer behavior, preferences, and needs to improve customer experience and tailor services.

Operational Efficiency: Optimizing banking processes by analyzing data to identify bottlenecks, streamline operations, and reduce costs.

Credit Scoring: Developing predictive models to assess creditworthiness and determine appropriate lending terms for customers.

Detection of Fraudulent Transactions: The primary goal is to accurately detect fraudulent transactions, such as credit card fraud, insurance fraud, or online payment fraud. This involves identifying unusual patterns or anomalies in the data that may indicate fraudulent behavior.

Minimize False Positives: While detecting fraud is crucial, it's equally important to minimize false positives, where legitimate transactions are incorrectly flagged as fraudulent. Balancing accuracy with the reduction of false positives is a key objective.

Real-time Detection: In some cases, the system may need to operate in real-time, detecting fraud as transactions occur. This requires efficient algorithms and data processing.

Market Trends: Analyzing market data to understand current needs and technologies.

High Security: This application provides high security to users with no loss of data.

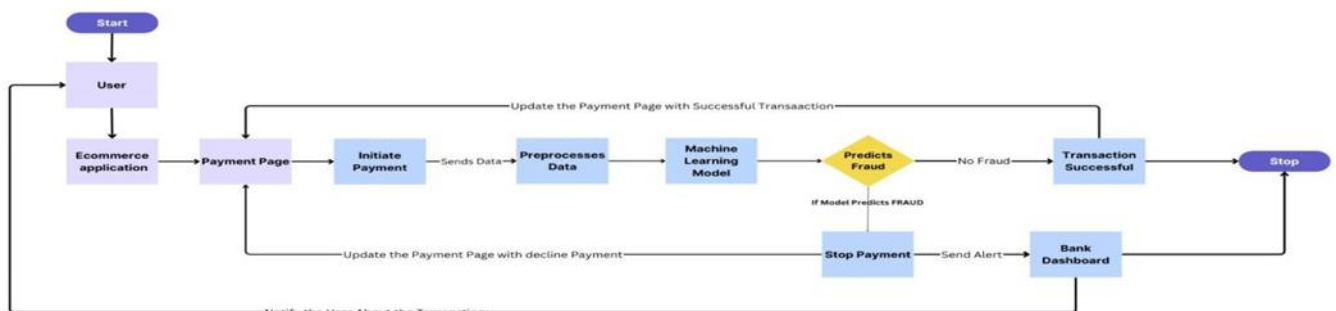


Figure: Fraud Detection Model

❖ III - LITERATURE SURVEY

1. **Fraud Detection using Machine Learning and Deep Learning (2019)** Many scammers steal the credit card' information for unauthorized purchases which causes fraudulent activities everyday. Various website and bank figure out fraudulent activities and stop them by using Machine learning & Deep learning.
2. **Predicting Mobile Money Transaction Fraud using Machine Learning Algorithms(2022)** [4]



This study mainly focuses on the machine learning classifiers in online money transfer to predict the transactions whether they are accurate or not. The classifiers such as Random forest gives the best performance in terms of fraud detection also the techniques such as logistic regression are also used.

3. **Big Data Application the Banking Sector A Bibliometric Analysis Approach(2021) [6],[7]** This study focuses on the literature study on the applications of big data in the banking sector. They use a bibliometric analysis method. This approach describes all the research outputs, implementations streams and citation rates and research agenda.

4. **Online transaction fraud detection system based on machine learning(2023)** detection algorithm based on XGBoost. This algorithm constructs the XGBoost classifier with best parameters by using Hyperopt [7][8] transaction fraud detection system finally provides services in the form of Web. In order to ensure the detection accuracy, we built an online detection platform based on XGBoost model with Django framework.

5. **Fraud Detection in Banking Data by Machine Learning Techniques Banking** sector are best use now-a-days in online mode only making it easy for the fraudsters in order to avoid such fraud in banking various security techniques are involved. Machine learning can be de-facto in order to detect the fraud efficiently. ML/DL can be utilize for fraud detection in better way for those attackers who surpasses basic cryptographic methods. The data received here are unbalanced data so to detect fraud in efficient way [1] has mentioned the ML/DL methods. Here Bayesian optimization methods are used to optimize the hyperparameters.

6. **An Amalgamated Novel IDS Model for Misbehaviour Detection using VeReMiNet** While using the machine methods the unbalanced data-sets need to be preprocessed where feature selection is wisely improve the efficiency of fraud detection. In [2] the author used the marine predator algorithm for feature selection algorithm which nature inspired meta-heuristic algorithm and follow the concept the marine creature prey and predator methods. This is one of the latest algorithm and now it is utilizing for machine learning feature selection.

7. **FRAUD DETECTION USING MACHINE LEARNING** Various machine learning algorithm are used to detect the fraud in financial transaction. To increase the precision in financial transaction of detecting fraud various strategies and algorithms are used. In [3] the author reviews research in addressing advantages and disadvantages

8. **Fraud Detection in Mobile Payment Systems using an XGBoost-based Framework** Nowadays, the number of smartphones are increasing, due to this online payment systems are becoming more popular, which in turn attracts more fraudsters. For this extensive research is required for detection of fraud using machine learning techniques. In [3] The author has proposed the extra boost based framework for detection of fraud in financial transactions. For efficiency of this framework a comparative study of machine learning techniques of modelling and outlier detection was done.

9. **Online payment fraud: from anomaly detection to risk management** In [3] The author has used various resources and methods like risk models, algorithms, human action, tools, web technology tools and business system in risk management. Now-a-days, online fraud detection is a difficult problem which requires a great understanding to deal with the large data set.

10. **The Impact of Big Data Analytics on the Banking Industry(2022)[1] ,[2]** in this paper, we propose an efficient approach for detecting on publicly available datasets and has used optimised algorithms LightGBM, XGBoost, CatBoost, and logistic regression individually.

11. **Online Banking Fraud Detection Based on Local and Global Behavior** [3] Payment fraud risk management, Anomaly detection, Ensemble models, Integration of machine learning and statistical risk modelling, Economic optimization machine learning outputs.

12. **Big data analysis in banking sector(2021)** The reminder of this paper is organized as follows: In Section [2] we review the related state-of-the-art. The proposed approach for credit card fraud detection including the dataset, pre-processing, feature extraction and feature selection, algorithms, framework, and evaluation metrics

❖ **IV. COMPARATIVE STUDY**

Sr /No	Paper Name	Author Name	Methodology	Advantage	Disadvantage
1	Fraud Detection using Machine Learning and Deep Learning (2019)	Pradheepan Raghavan	Fraud Detection Technique 1. Neural Network	A.very powerful and accurate with large transactions volume	Based on a large-scale sample detection algorithm. B. Needs high computing capabilities. C. Accuracy rate depends on the sample size (moderate accuracy).
2	Fraud Detection in Banking Data by Machine Learning Techniques	Seyedeh Leili Mirtaheri	available datasets and has used optimised algorithms LightGBM, XGBoost, CatBoost, and logistic regression individually	It Uses Evaluation metrics like Precision, Recall and F1-Score for Accuracy	More Memory & Time Needed -Hardware Limitations
3	Online payment fraud:from anomaly detection to risk management(2023)	Sebastiano Rossi Ermin Zvizdic Thomas Domenig	Payment fraud risk management, Anomaly detection, Ensemble models, Integration of machine learning and statistical risk modelling, Economic optimization machine learning outputs	Anomaly detection reduces risk	Received no funding rights from any agency-statistical uncertainty and variability in the driving forces of the defined models
4	Predicting Mobile Money Transaction Fraud using Machine Learning Algorithms(2022)	Mark Lokanan	Predicting Mobile Money Services (MMS) or Mobile Money Transfer Services (MMTS)	Random Forest is used to achieve high precision	We can also use human intelligence for better context & insights which is needed for an automated system
5	Fraud Detection in Mobile Payment Systems using an XGBo	Mohammad Zoynul Abedin Uthayasankar Sivarajah	Mobile payment transactions are carried out using mobile phone technologies that allow users to deposit, withdraw,	Speeds up the transaction	Interpretability is needed

	ost-based Framework(2022)		spend, transfer and send money		
6	Big Data Application in the Banking Sector A Bibliometric Analysis Approach(2021)	HAitham Nobanee Mehroz Nida Dilshad Mona Ai Neyadi Sultan Ai Qubaisi Saeed AI Shamsi	different metrics have been evaluated with the help of bibliometric tools to demonstrate and comprehend emerging research trends on this research topic	Used by bibliographic review to analyze risk assessment, credit card risk, etc	Common topics are explored by researchers like database banking, bankcards, bankcard transaction, etc
7	Big Data Analysis in Banking Sector(2021)	Rahul More Yash Moily	transaction fraud detection system finally provides services in the form of Web. In order to ensure the detection accuracy, we built an online detection platform based on XGBoost model with Django framework	Study of Indian Economy	Risk of non-compliance
8	Online Transaction Fraud Detection System Based on Machine Learning(2023)	Bocheng Liu, Xiang Chen, and Kaizhi Yu	detection algorithm based on XGBoost. This algorithm constructs the XGBoost classifier with best parameters by using Hyperopt	-XGBoost classifier is used -XGBoost performs better and uses shorter time for prediction	Data Dependency Frequent cyber attacks may take place

❖ V. CONCLUSION

One of the leading frauds in the past few decades is Online Payment Fraud. In this research paper, we discussed and studied the concept of online payment fraud detection. It was seen that feature selection techniques are very important and can be implemented to attain lower false positive rate. We also implemented various machine learning algorithms like Logistic Regression, Random Forest for prediction. It is used for detecting if a particular transaction is fraudulent or not. A good fraud detection system should be accurate to predict if a given transaction is fraudulent or not. To improve the performance of the models, various techniques such as handling class imbalance, feature selection was used. Confusion matrix was used to evaluate the performance of our models, however we did not attain 0 False Positive and false negative score. It is important for a financial organization to attain 0 false



positive and negative score as we discussed it impacts on the customer retention and costs lot of money for the refunds. More future works can be done on this research in order to attain the 0 false positive and negative score. Combination of models can be used to attain high accuracy in identifying the transactions as fraudulent and non-fraudulent.

VI. REFERENCES

- I. S. K. Hashemi, S. L. Mirtaheri and S. Greco, "Fraud Detection in Banking Data by Machine Learning Techniques," in *IEEE Access*, vol. 11, pp. 3034-3043, 2023, doi: 10.1109/ACCESS.2022.3232287.
- II. Saleha Saudagar, Rekha Ranawat, An Amalgamated Novel IDS Model for Misbehaviour Detection using VeReMiNet, *Computer Standards & Interfaces*, Volume 88, 2024, 103783, ISSN 0920-5489, <https://doi.org/10.1016/j.csi.2023.103783>
- III. Raghavan, Pradheepan & Gayar, Neamat. (2019). Fraud Detection using Machine Learning and Deep Learning. 334-339. 10.1109/ICCIKE47802.2019.9004231.
- IV. Hajek, PetrAbedin, Mohammad ZoynulSivarajah, Uthayasankar2023 2023/10/01Fraud Detection in Mobile Payment Systems using an XGBoost-based Framework *Information Systems Frontiers* 1572-9419 <https://doi.org/10.1007/s10796-022-10346-6>
- V. JOUR Vanini, Paolo Rossi, Sebastiano Zvizdic, Ermin Domenig, Thomas 2023 2023/03/13 Online payment fraud: from anomaly detection to risk management *Financial Innovation* <https://doi.org/10.1186/s40854-023-00470-w>
- VI. Amakobe, Moody. (2015). The Impact of Big Data Analytics on the Banking Industry. 10.13140/RG.2.1.1138.4163.
- VII. @inproceedings{Kovach2011OnlineBF, title={Online Banking Fraud Detection Based on Local and Global Behavior}, author={Stephan Kovach and Wilson Vicente Ruggiero}, booktitle={International Conference on the Digital Society}, year={2011}, url={https://api.semanticscholar.org/CorpusID:64246376} }
- VIII. Rahul More, et. al. *International Journal of Engineering Research and Applications* www.ijera.com ISSN: 2248-9622, Vol. 11, Issue 4, (Series-II) April 2021, pp. 01-05
- IX. P. Raghavan and N. E. Gayar, "Fraud Detection using Machine Learning and Deep Learning," 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, 2019, pp. 334-339, doi: 10.1109/ICCIKE47802.2019.9004231.
- X. Emmanuel Gbenga Dada, Joseph Stephen Bassi, Haruna Chiroma, Shafi'i Muhammad Abdulhamid, Adebayo Olusola Adetunmbi, Opeyemi Emmanuel Ajibuwa, Machine learning for email spam filtering: review, approaches and open research problems, *Heliyon*, Volume 5, Issue 6 2019, e01802, ISSN 2405-8440, <https://doi.org/10.1016/j.heliyon.2019.e01802>.
- XI. Ali, Abdulalem, Shukor Abd Razak, Siti Hajar Othman, Taiseer Abdalla Elfadil Eisa, Arafat Al-Dhaqm, Maged Nasser, Tusneem Elhassan, Hashim Elshafie, and Abdu Saif. 2022. "Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review" *Applied Sciences* 12, no. 19: 9637. <https://doi.org/10.3390/app12199637>
- XII. Ayesha S. Dina, D. Manivannan, Intrusion detection based on Machine Learning techniques in computer networks, *Internet of Things*, Volume 16, 2021, 100462, ISSN 2542-6605, <https://doi.org/10.1016/j.iot.2021.100462>.
- XIII. S. K. Hashemi, S. L. Mirtaheri and S. Greco, "Fraud Detection in Banking Data by Machine Learning Techniques," in *IEEE Access*, vol. 11, pp. 3034-3043, 2023, doi: 10.1109/ACCESS.2022.3232287.
- XIV. Ali, A.; Abd Razak, S.; Othman, S.H.; Eisa, T.A.E.; Al-Dhaqm, A.; Nasser, M.; Elhassan, T.; Elshafie, H.; Saif, A. Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review. *Appl. Sci.* **2022**, *12*, 9637. <https://doi.org/10.3390/app12199637>