WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Online Financial Fraud Detection

***Aditya Pal*[1], *Yash Singh*[2], *Chitransh Varshney*[3], *Abhishek Kumar*[4], *Santosh Kumar*[5]**

Department of IT, Ghaziabad (UP), India- 201003
aadu24feb@gmail.com[1], yashsinghlkw2015@gmail.com[2], chitranshvarshney4@gmail.com[3], abhi242724@gmail.com[4], skwebfit@rkgit.edu.in[5]

ABSTRACT-

Traditional fraud detection techniques are being tested by the rise in fraudulent activity brought about by the explosion of online financial transactions. In order to increase detection efficiency and accuracy, this work investigates the development of machine learning (ML) algorithms for online financial fraud detection systems. Preprocessing data, feature extraction, and the use of supervised and unsupervised machine learning models, such as Decision Trees, Logistic Regression, and Isolation Forests, are important components. By addressing data imbalance, methods such as Synthetic Minority Oversampling Technique (SMOTE) improve model performance. The work contributes to more reliable and flexible fraud detection systems by reviewing recent research, assessing different machine learning approaches, and discussing potential future paths.

Keywords: — credit card frauds, fraud detection, fraudulent, Support Vector Machine (SVM), One Time Password (OTP), Security questions

I. INTRODUCTION

Online financial transactions have become more common in the digital age, providing companies and customers with a level of ease never seen before. On the other hand, financial fraud has significantly increased as a result of this spike in digital activity. Financial fraud on the internet refers to a broad spectrum of malevolent actions that are intended to steal money or private data from people or businesses. Strong fraud detection systems are more important than ever because rule-based and human verification procedures are sometimes insufficient to combat the advanced tactics used by contemporary fraudsters. [1] [2].

Advanced machine learning (ML) algorithms are used by online financial fraud detection systems to recognize and stop fraudulent transactions in real time. Large volumes of transaction data are analyzed by these systems in order to find trends and abnormalities that could point to fraud. Pre-processing data, feature extraction, model training, and real-time detection are some of the major processes that are usually included in an extensive fraud detection system. [2] [4].

Handling the very unbalanced nature of the data is one of the main obstacles to creating efficient fraud detection systems. Since fraudulent transactions make up a relatively tiny percentage of all transactions, it is challenging for typical machine learning models to detect them correctly without producing a large number of false positives. In order to solve this, the dataset is balanced using methods like the Synthetic Minority Oversampling Technique (SMOTE), and model performance is improved via the use of hybrid and ensemble methodologies.

Furthermore, fraud detection systems must be flexible and updated on a regular basis due to the dynamic and ever-evolving nature of fraudulent activity. Machine learning models need to be able to adapt to new fraud trends and learn from fresh data. This flexibility is essential to preserving the system's long-term efficacy.

The purpose of this study paper is to examine the many elements and approaches that go into creating an online financial fraud detection system. It will go over recent research on the topic, evaluate the efficacy of various machine learning strategies, and talk about the difficulties and potential paths forward in this area. This study aims to provide a thorough review of current solutions and their shortcomings in order to support continued efforts to prevent fraud in online financial transactions.

II. LITERATURE REVIEW

Advanced detection systems are required because of the increased risk of fraud caused by the widespread use of online financial transactions. Conventional approaches, such rule-based systems and human verification, are becoming less and less effective since they can't keep up with the latest fraud strategies. The integration of machine learning (ML) methods in online financial fraud detection is examined in this overview of the literature, with a focus on important methodology, difficulties, and potential areas for future study. [5]

Methods of Machine Learning

Supervised Learning: In the field of fraud detection, supervised learning techniques like Support Vector Machines (SVM) and Artificial Neural Networks (ANN) are widely used. Based on past data (MDPI) (SpringerOpen), these models are trained on labeled datasets to distinguish between authentic and fraudulent transactions. Unsupervised Learning: Without labeled training data, methods such as Local Outlier Factor (LOF) and Isolation Forests are used to find anomalies. These techniques work well for spotting anomalies, which are often signs of fraud and diverge from typical transaction patterns (SpringerOpen).

Hybrid Approaches: Increasing the accuracy of detection by combining supervised and unsupervised techniques. As an example, a system may use supervised learning to verify suspicions of fraud and unsupervised learning to identify possible scams (SpringerOpen).

Data Difficulties

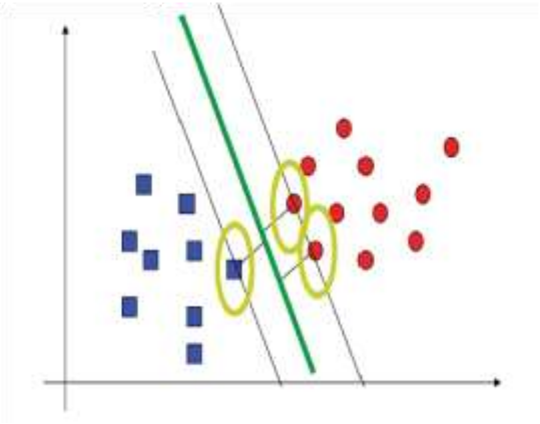
Unbalanced Datasets: Since fraudulent transactions are uncommon in comparison to valid ones, there is a significant imbalance in these datasets. In order to enhance model training (MDPI), strategies such as Synthetic Minority Oversampling Technique (SMOTE) are used to rectify this imbalance by creating synthetic instances of the minority class.

Feature Engineering: A key component of effective fraud detection is feature extraction. It is essential to include features that record transaction patterns, contextual data, and consumer behavior. It takes extensive subject expertise and access to large datasets (SpringerOpen) to develop these features.[6]. Metrics for Evaluation

Area Under the Receiver Operating Characteristic Curve (AUC-ROC), precision, recall, and F1-score are common assessment measures for fraud detection systems. These indicators aid in evaluating the trade-off between preventing false positives (SpringerOpen) and identifying frauds (true positives). [7 Main Results and Omissions

Utilization of ML Models: Both SVM and ANN are popular and quite effective in identifying fraud. However, the accuracy of the data and the suitability of the feature selection (MDPI) determine how well they function. [8] Real-Time Detection: A number of studies highlight the need for systems capable of immediately processing huge numbers of transactions for real-time fraud detection. This calls for scalable and effective massive data management techniques (SpringerOpen). [9] Adaptive Systems: As fraud trends change quickly, systems that are able to continually learn and update their models are required. Future research might focus on feedback loops and reinforcement learning to develop more robust fraud detection systems (SpringerOpen).

Prospective Courses Subsequent investigations need to concentrate on refining data preparation methodologies, creating more intricate approaches for extracting features, and incorporating complex machine learning methods like deep learning and reinforcement learning. Furthermore, investigating the use of cutting-edge anomaly detection techniques with blockchain technology may open up new possibilities for improving fraud detection systems (MDPI) (SpringerOpen).This paper highlights the crucial role that machine learning plays in creating efficient online financial fraud detection systems by combining existing approaches and pointing out shortcomings. It also suggests interesting avenues for further research..



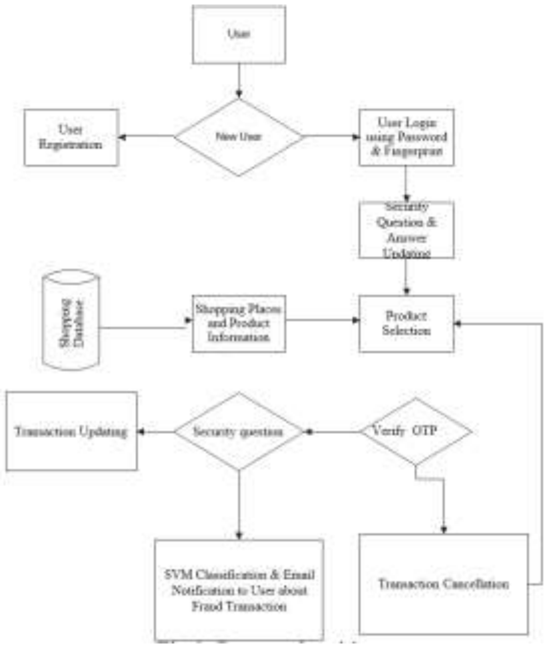

User
New User
User Registration
User Login using Password & Fingerprint
Security Question & Answer Updating
Shopping Database
Shopping Places and Product Information
Product Selection
Transaction Updating
Security question
Verify OTP
SVM Classification & Email Notification to User about Fraud Transaction
Transaction Cancellation

Fig.1: Block diagram of proposed system

III. ALGORITHM

Support Vector Machine (SVM):

VM may also be an algorithm for supervised machine learning that can be used for both classification and regression problems. It's often used in classification issues, however. In this algorithm, we plot each data item in a "N" dimensional space to some degree, with the price of each function being the price of a chosen coordinate showing the hyper-plane differentiating the 2 groups from the hyper-plane. A supervised machine learning algorithm may also be a support vector machine (SVM), which can be used for both classifications. During this algorithm, we take the shopping data item of each consumer to some degree in a 'n' dimensional space (where n is the number of features you have) with the price of each function being the price of a chosen coordinate. Within the linearly separable example, the hyper-plane differentiating, there is one or more hyperplane that can distinguish the two groups defined by training data with 100%.

IV. EXPERIMENTAL METHODOLOGY

A. New User Registration with fingerprint image

Here, we develop an integrated online shopping module for user. The new user registration form consists of username, mobile number, address, city, E-Mail ID, etc. The secondary security clearance level which deals with the fingerprint scanning. It is a standalone fingerprint identification device with many excellent features such as high identification performance. user's fingerprints collected and unique hash code value generated using one-way hash function and it's stored into the bank server's database.

B. User Login authentication using fingerprint Image

The user is requested to enter their login details such as username, password and fingerprint image. A fingerprint consists of ridges and valleys (lines between fingerprints) (spaces between ridges). For every person, the pattern of the ridges and valleys is special. There are two main methods of

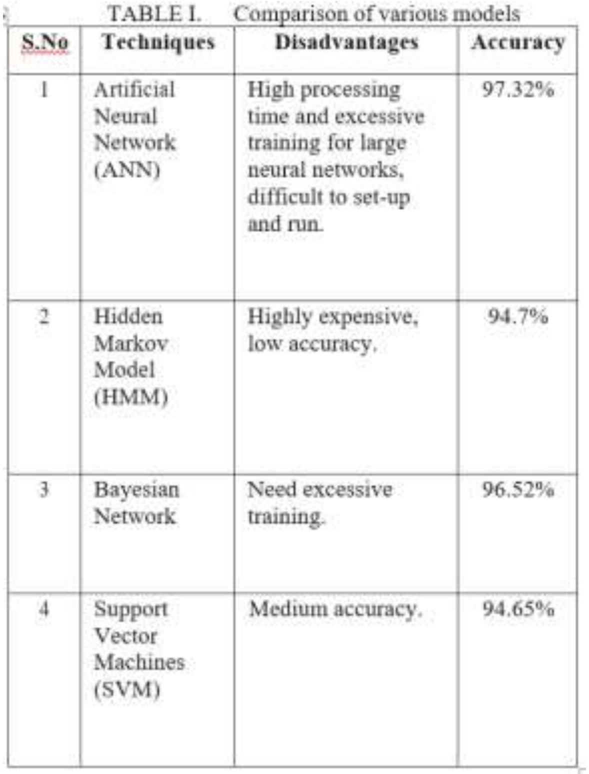


TABLE I. Comparison of various models

S.No	Techniques	Disadvantages	Accuracy
1	Artificial Neural Network (ANN)	High processing time and excessive training for large neural networks, difficult to set-up and run.	97.32%
2	Hidden Markov Model (HMM)	Highly expensive, low accuracy.	94.7%
3	Bayesian Network	Need excessive training.	96.52%
4	Support Vector Machines (SVM)	Medium accuracy.	94.65%

Fig.2: Flow chart of the system

C. Security question Customization and Verification module

To enroll for the 'Security Questions' verification, the Login user selects several questions and supplies confidential answers that only the user knows. The banking server system provides a set of default questions to users. Security questions are stored into database by securely on a user and can only be answered by you during online shopping's checkout verification. There is a provision available for individual user to read or modify questions and/or answers.

D. One-Time Password (OTP) Verification

The aim of authentication is to prove that the shopping user is that the authentic user or fraud user for suspicious shopping transaction identification. One-Time Passwords are utilized as a supplemental think about multi-factor sanction/authentication applications. they're only valid for precisely one sanction or authentication request. To evade password lists, a convenient thanks to provide the user with an OTP is to send it through email. the e-mail id of the user must be registered for the accommodation that gives email OTPs for authentication or sanction. OTPs are quite popular as a supplemental sanction or authentication think about web- predicated accommodations. These passwords are often utilized to authenticate a user, i.e., the user needs a legitimate OTP to prove his identity to authenticate in to a application or to access the network. Email OTPs are withal utilized for account verification, e. g., Google Mail. After successful OTP verification process, Online shopping's banking server will through Security questions which may be used as an additional level of security when your clients contact you so as to verify their identify. User can then prompt them for the solution to the question they chose during the registration phase.

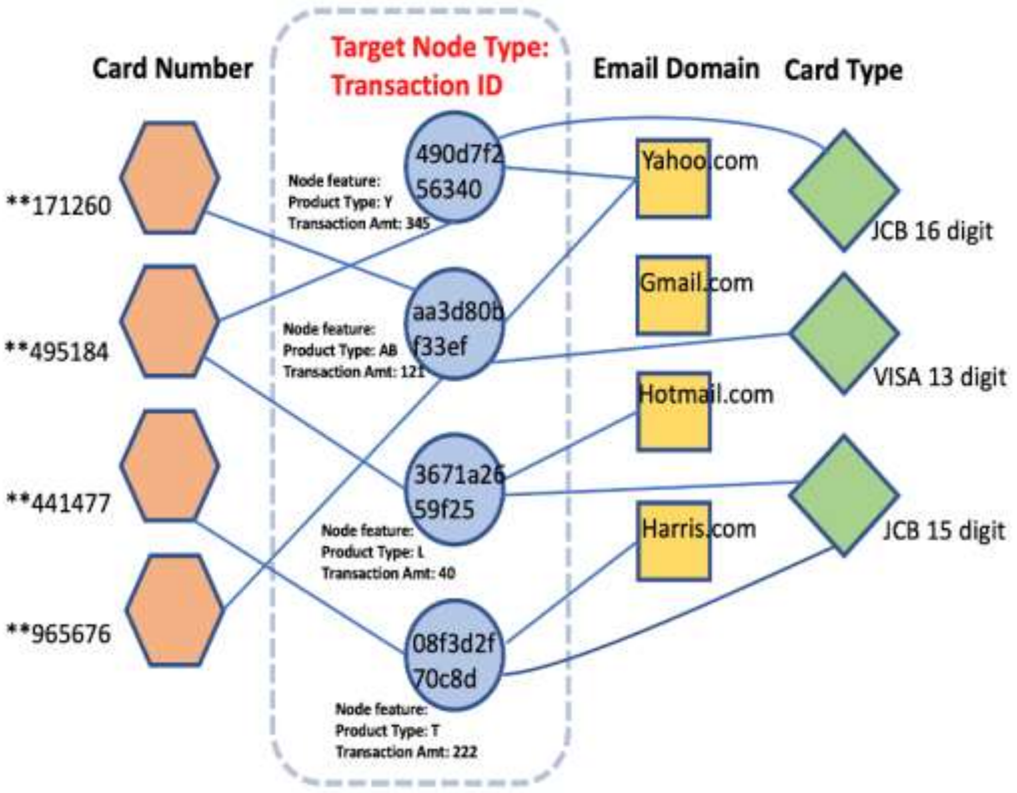
E.Online Fraud Transaction Classification using SVM

In order to predict or classify patterns into two categories, SVM may be a classifier: fraudulent or non-fraudulent. During this algorithm, we take the shopping data item of each user to some extent in a 'n' dimensional space with the value of each function being the value of a particular coordinate. Within the linearly separable case, the hyper-plane differentiating, there is one or more hyperplane that separates the two classes represented by training data with 100 percent. If a shopping transaction is classified as fraud, payment deduction and checkout process will be cancelled by server and fraudulent user's transaction details are sent to that specific user thru email else shopping is completed and transaction stored into database.

V PERFORMANCE EVALUATION

The system presents classification of online credit / open- end credit the challenges faced by cardholder also because the card issuer, verity of fraud implemented by the persons who commit that fraud, some latest news regarding master card fraudster and provide some prevention techniques that need to

Hereby, it concludes that each system faces its own problems while dealing with dataset description. Even in early system of SVM model which deals with preprocessing datasets faces low accuracy rate. So, the proposed system come with the SVM model of real databases which helps in acquiring the maximum of 99.9% accuracy.



VI Scope

1. Overview and Context

Definition and Types of Financial Fraud: Clearly define financial fraud and group its many forms, including transaction, phishing, identity theft, and credit card fraud.

Importance and Effects: Talk about how financial fraud affects individuals, companies, and the financial system on an economic and social level.

2. Modern Methods and Technologies

AI and Machine Learning: Examine how supervised and unsupervised learning techniques are used in machine learning to find irregularities and fraudulent activity.

Data Mining Techniques: Talk about how data mining may be used to find trends in big datasets that might indicate fraud.

Blockchain and Cryptography: Investigate how to improve the security and traceability of financial transactions using blockchain technology and cryptographic techniques.

Rule-based Systems: Research the efficacy of conventional rule-based detection systems.

3. Sources and Gathering of Data

Financial Data Types: List the several kinds of financial data that are used to identify fraud, including transaction logs, user activity data, and data from outside sources like social media.

Data Collection Methods: Talk about the many ways that data is gathered and the significance of data integrity and quality.

4. Difficulties with Fraud Detection

Changing Fraud Tactics: Talk about how fraud tactics are always changing and how adaptive detection systems are necessary.

Data Security and Privacy: When putting fraud detection systems into place, consider the moral and legal issues around data security and privacy.

Scalability: Talk about the challenges of expanding fraud detection systems to manage high transaction volumes in real time.

5. Performance Metrics and Evaluation

Accuracy and Precision: Describe and go over the meaning of measures like F1-score, accuracy, precision, recall, and their significance in assessing how well fraud detection systems work.

Examine the effects that false positives and false negatives have on the financial institutions and user experience.

6. Applications and Case Studies

Provide case studies of financial institutions or businesses that have effectively adopted fraud detection systems as examples from the real world.

Success Stories and Failures: Examine instances of fraud prevention that have been successful as well as noteworthy failures, together with the lessons that may be drawn from them.

7. Upcoming Patterns and Prospects for Research

Future developments in machine learning and artificial intelligence (AI): Project how these fields will develop and how they can improve fraud detection.

Integration with Other Technologies: Look into possible connections to other cutting-edge technologies, such as quantum computing and the Internet of Things (IoT).

Impact of Policies and laws: Talk about how the creation and use of fraud detection systems may be impacted by changing policies and laws.

8. Final Thoughts

Summary of Findings: Provide an overview of the main ideas covered in the study.

ramifications for academia and industry Emphasize how your results will affect academic research as well as the financial sector.

Suggestions: Make suggestions for upcoming studies and useful applications in the area of identifying online financial fraud.

Methods of Research

Literature Review: Perform a thorough analysis of the body of knowledge about the identification of online financial fraud.

Analyze data using statistical and computational tools to confirm the efficacy of different fraud detection strategies.

Interviews and Surveys: Gather qualitative data by interviewing practitioners and industry experts and conducting surveys.

VII Future Motives

1. Adjusting to Changing Dangers

Dynamic Fraud strategies: Adaptive detection systems are necessary due to the ongoing development of fraud strategies. Subsequent investigations have to concentrate on creating algorithms with the ability to instantly recognize and adjust to novel fraudulent conduct patterns.

Real-time Detection: It's critical to improve real-time detection skills to stop fraud before it starts. This entails enhancing the fraud detection systems' accuracy and speed.

2. Making Use of Modern Technologies

Machine learning and artificial intelligence (AI): These fields will see more application as more complex models that are better able to identify and stop fraud are developed.

Deep Learning: Examining how deep learning methods might enhance detection rates and provide more profound insights into intricate fraud patterns.

Quantum Computing: Investigating how quantum computing may be used to process complicated calculations and huge datasets more quickly, leading to more precise and quick fraud detection.

3. Improving the Use of Data

Big Data Analytics: Applying big data analytics to handle and examine enormous volumes of transaction data in order to spot odd trends and stop fraud.

Behavioral Biometrics: Adding an extra degree of security by using behavioral biometrics to comprehend and validate user behavior.

Cross-platform Data Integration: Combining information from many platforms (such as social networking, banking, and e-commerce) to provide a thorough understanding of user behavior and identify fraudulent activity across various media.

4. Strengthening Security Protocols

Blockchain Technology: By using blockchain technology to create immutable, transparent, and secure transaction records, fraud risk may be decreased.

Enhanced Encryption: Creating more robust encryption methods to stave against fraud and illegal access to private financial information.

Using more reliable multi-factor authentication techniques can help to guarantee that only authorized users are able to carry out financial transactions.

5. Adherence to Policies and Regulations

Regulatory adaptation is the study of how to create fraud detection systems that adhere to changing norms and laws while maintaining moral and legal propriety.

Fostering International Cooperation: Promoting international cooperation between regulatory agencies and financial institutions to exchange data and create standardized methods for detecting fraud.

6. Awareness and Education of Users

Customer Education Programs: Creating initiatives to educate clients about typical fraud strategies and self-defense techniques.

Interactive Security Features: Adding interactive elements to financial systems to assist users in real-time fraud detection and prevention.

7. Moral Aspects

Privacy Concerns: Striking a balance between people's right to privacy and the efficiency of fraud detection. ensuring that user privacy is not violated by detecting systems.

Bias in AI Models: To guarantee impartial and equitable fraud detection for all users, biases in AI models are examined and mitigated.

8. Social and Economic Repercussions

Effectiveness in terms of cost: Creating affordable fraud detection systems that banks of all sizes may use.

Social Responsibility: Examining fraud detection's wider societal ramifications, such as how it affects confidence in financial institutions and how it helps to lower financial crime.

In summary

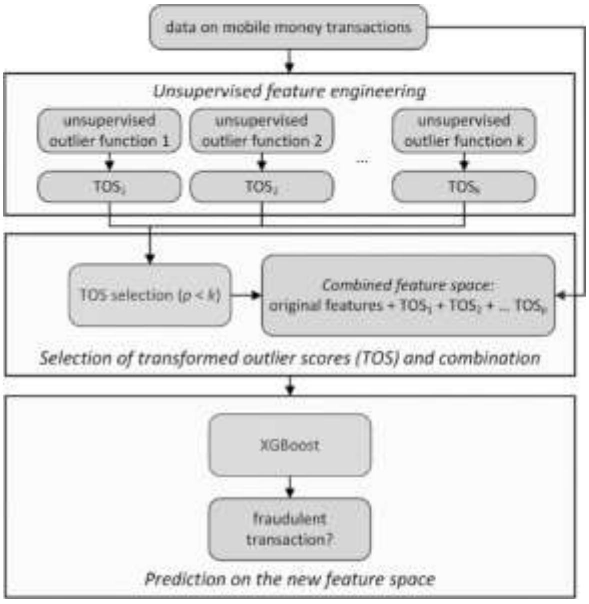
Technology, regulations, and the growing demand for security are all expected to drive major breakthroughs in online financial fraud detection in the future. Future research may help preserve the financial ecosystem by concentrating on these motivations in order to design fraud detection systems that are more resilient, adaptable, and user-friendly.

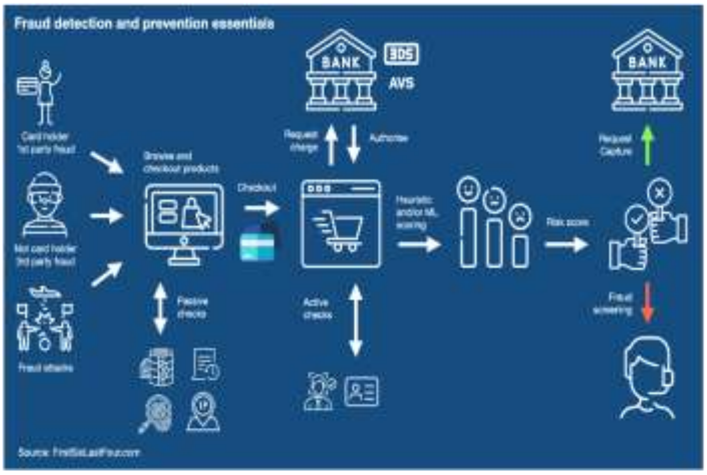
Suggestions for Upcoming Multidisciplinary Research Approaches: integrating knowledge from criminology, psychology, finance, and computer science to create comprehensive fraud detection techniques.

Encouraging constant innovation in detecting technology and procedures in order to keep ahead of new fraud strategies.

Conducting longitudinal research is one way to evaluate the long-term efficacy of fraud detection systems and how they affect the decrease of fraud.

Through examining these motivations, scholars may enhance the identification of online financial crime and guarantee that the financial industry is robust to constantly changing risks.


data on mobile money transactions
Unsupervised feature engineering
unsupervised outlier function 1
unsupervised outlier function 2
...
unsupervised outlier function k
TOS1
TOS2
TOSk
TOS selection (p < k)
Combined feature space: original features + TOS1 + TOS2 + ... TOSp
Selection of transformed outlier scores (TOS) and combination
XGBoost
fraudulent transaction?
Prediction on the new feature space


Fraud detection and prevention essentials
Card holder 1st party fraud
Not card holder 3rd party fraud
Fraud attacks
Browse and checkout products
Passive checks
Checkout
BANK
3DS
AVS
Request charge
Authorise
Active checks
Heuristic and/or ML scoring
Risk score
BANK
Request Capture
Fraud screening

Source: FirstSixLastFour.com

VIII. CONCLUSION

This project provides a classification of online credit and open-end credit, details the difficulties that cardholders face due to the card issuer, verifies the fraud committed by the individuals committing the fraud, provides some recent news regarding master card fraudsters, and offers some preventive techniques that the cardholder should use to stop fraudulent activity. In recent years, master cards have emerged as the most widely used form of payment, and with a rise in master card transactions comes an increase in fraud. The good news is that advances in technology have made it easier to avoid online fraud, and lower computing costs have made it possible to put in place sophisticated systems that can quickly identify fraudulent activity.

In this research, the SVM behavior-based classification algorithm is used. SVM is commonly used to provide excellent fraud detection performance. Usually, SVM offers an original answer. The SVM becomes more flexible in terms of the kind of threshold used to separate the data. These characteristics enable the SVM to solve the classification issue in this complicated area and provide a reliable result.The proposed method is scalable for handling a high volume of transactions and provides increased identification accuracy.

IX. REFERENCE

[1] KaithekuzhicalLeenaKurien& Dr. AjeetChikkamannur (2019), DETECTION AND PREDICTION OF CREDIT CARD FRAUD TRANSACTIONS USING MACHINE LEARNING, Issue & volume : 8 (3) , p.p - 199 - 207.

[2] SamanehSorournejad , Zahra Zojaji , Reza EbrahimiAtani , Amir Hassan Monadjemi , A Survey of Credit Card Fraud Detection Techniques: Data and Technique Oriented Perspective 19 Nov 2016

[3] Nana Kwame Gyamfi , Jamal-DeenAbdulai , (2018) , Bank Fraud Detection Using Support Vector Machines

[4] Yashvi Jain, NamrataTiwari, ShripriyaDubey, Sarika Jain in A Comparative Analysis of Various Credit Card Fraud Detection Techniques ISSN: 2277-3878, Volume-7 Issue-5S2, January 2019

[5] Linda Delamaire, Hussein Abdou , John Pointon in Credit card fraud and detection techniques Volume 4, Issue 2, 2009

[6] International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056, p-ISSN: 2395-0072, Volume: 05 Issue: 03 | Mar2018 "Fraud Detection in Online Credit Card Payment ". Aishwarya Kaneri1, Anugrah S2, Isha Bharti3, Samruddhi Jadhav4, Mitali Kadu5