ATM fraud detection system using machine learning algorithms

Gurjeet Ujjainwar[1], Ashish Gacche[1] ,Aniket Bawankule[1], Aditya Sahare[1], Kajal Labhane[1], Priyanka Gonnade[2]

[1]G H Raisoni Academy of Engineering & Technology (GHRAET), Nagpur, Maharashtra, India - 440028
[2]Assistant Professor, CSE Department of G H Raisoni Academy of Engineering & Technology (GHRAET), Nagpur, Maharashtra, India - 440028

---***---

Abstract - *The banking industry has long been an essential institution that contributes significantly to the sustainability and maintenance of a country's economy. When banking transactions are tampered with by intruders or fraudsters, the results can be dire. This article aims to examine the current ATM electronic transfer (EFT) system for cash withdrawals, money transfers, password cracking, PIN misplacing and biotechnology. The article will look at many types of fraud and try to find a solution to solve and detect ATM fraud, as well as a more advanced machine that can accept security technology. EDP) in the banking sector; this will be achieved by mining biometric data by biometric combo operations at the first opening of the account and will therefore comply with the proposed algorithm; The study looked at a lot of current literature and systems to see how biometric approaches may be combined with existing methodology to come up with a comprehensive proposal for a dynamic design. The ATM engine will have an embedded fingerprint capture area and eye scanning capability and also ensure that this does not slow down the process at an unacceptable rate.*

***Key Words*: ATM, Fraud Detection , Machine Learning, Fraud Detection Systems, Security, Electronic fund transfer(EFT), Electronic data processing.**

1. INTRODUCTION

In recent years, the most notable data mining concern has been people using data mining-based ATM card fraud detection algorithms. Data mining is essential when it comes to detecting ATM card fraud in online transactions. Because our topic is addressed as a classification problem, traditional data mining methodologies are ineffective. As a result, a new method based on general-purpose meta heuristic methods, such as machine learning techniques, has been devised. The goal of this study is to create an ATM card fraud detection system based on genetic algorithms. Iterative algorithms that strive to improve solutions over time are known as machine learning algorithms.In order to reduce false alarms, it employs machine learning techniques to optimize a collection of interval valued parameters. Create an ATM card fraud detection system using a genetic algorithm. During an ATM card transaction, the fraud is detected, and a genetic algorithm is utilized to reduce the number of false alarms. Instead of maximizing the number of transactions that have been successfully classified, we created an objective function with variable misclassification costs, such that correctly categorizing certain transactions is more essential than correctly categorizing others. This information is important for the proposed system's study. Here are all of the cost and performance factors that will influence the project's viability as well as its goal.

2. METHODOLOGY

There are three types of machine learning algorithms: supervised, unsupervised, and semi supervised. Supervised Machine Learning Algorithms: You may use tagged examples to predict future occurrences by applying prelearning to new data. Based on the evaluation of a given training dataset, the learning algorithm develops a derived function to predict the output value. The system may build a goal for each new input after sufficient training. To detect and correct model flaws, the training process might compare its output to the right intended output. When there is no method to categorize or tag training data, unsupervised machine learning algorithms are used. The study of how computers extract functions from unlabeled input and explain hidden structures is known as unsupervised learning.The system evaluates the input and utilizes the dataset to infer hidden structures from unlabeled data, rather than picking the best output. Semisupervised machine learning methods, which train with both labeled and unlabeled data, fall between supervised and unsupervised learning, with a small amount of labeled data and a large amount of unlabeled data. both unlabeled and identified In systems that adopt it, this strategy can considerably improve learning accuracy. Semi Supervised learning is typically used when the gathering of labeled data involves the use of appropriate and skilled training/learning resources. Collecting unlabeled data, on the other hand, seldom necessitates the use of additional resources.

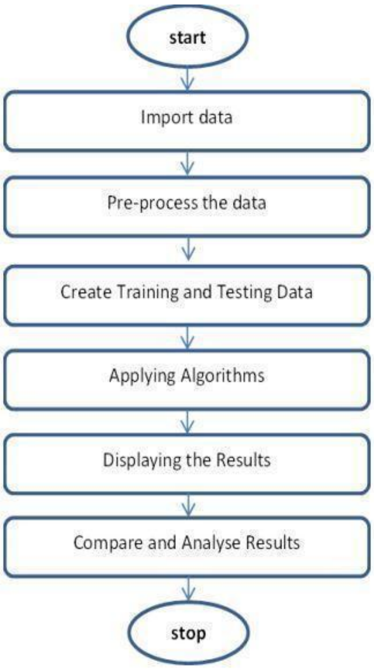


Chart -1: Flowchart of Algorithm

The algorithm's flowchart is depicted in chart-1.

2.1 HMM Based Fraud Detection:

Hidden Markov Chains Model A Markov model is a statistical model based on the assumption that the system being studied is a Markov process with an unobserved state. It detects fraud by examining user spending profiles, which are classified into three groups: The three possibilities are low profile, moderate profile, and high profile. Training, detection, and prevention are all phases of the procedure.In this setup, start the bank server and the HMM server first. When the client initiates a transaction, HMM starts observing and comparing it. The transaction is delayed and halted if fraud is discovered. Users type their password on their phone and send it over Bluetooth or SMS to the same ATM.

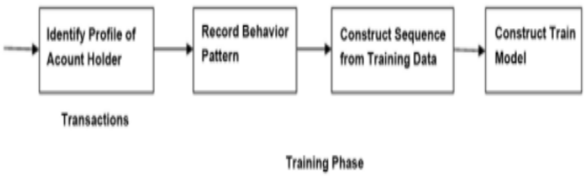

Identify Profile of Acount Holder
Record Behavior Pattern
Construct Sequence from Training Data
Construct Train Model
Transactions
Training Phase

Fig -1: Training Phase of model

Fig.1 describes training phase of model

2.2 Support Vector Machine (SVM) Based Fraud Detection:

Support Vector Machines, or SVMs, are statistical learning systems that have been applied to a wide range of issues. In the SVM classification strategy, the hyperplane is employed as the decision plane, and the distance between the positive and negative modes is maximized. SVM is a typical machine learning algorithm for classification, regression, and other problems. LIBSVM is an SVM library. LIBSVM is typically applied in two stages: first, to train a model, and then, to predict data from a test dataset. The following are the main characteristics of SVM:

1) First, create the model's training data.

2) Next, set SVM parameters for the newly created dataset and send it to SVM Training.

3) SVM Trainer: Every data point in the large dataset is trained by this software..

4) The SVM Predictor predicts the training data when the dataset has been fully trained.

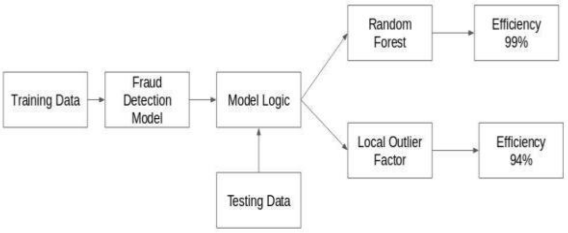


Fig -2: Flow of Implementation

The fIG. 2 describes the flow of implementation

Algorithms:

The Decision Tree Algorithm and How to Use It

Step 1:The data must first be imported.

Step 2: Determine the training-to-testing data ratio.

Step 3: Using the most significant attribute as the root, divide the dataset into subgroups.

Step 4: Determine the number of buckets associated with the construction parameters.

Step 5: The Confusion Matrix and the Underlying Data Accuracy

3. RESULT

3.1 Dataset Description:

A customer of a fake bank made an ATM card transaction in the dataset. In the previous 48 hours, this dataset contains 492 fraud cases from 2,84,807 transactions. The sample is strongly biased because the positive category (rogue) accounts for 0.172 percent of all transactions. In this study, only PCA-transformed digital input variables are used. We are unable to reveal the data's original attributes or underlying information due to security concerns. PCA has no effect on the attributes of "time" and "quantity." The Time method keeps track of how many seconds have passed since the collection's first transaction. The Amount parameter represents the transaction's total value. This capability is useful for both cost-conscious and example-based learning. The Grade response variable is set to 1 if cheating occurs. It is set to 0 otherwise.

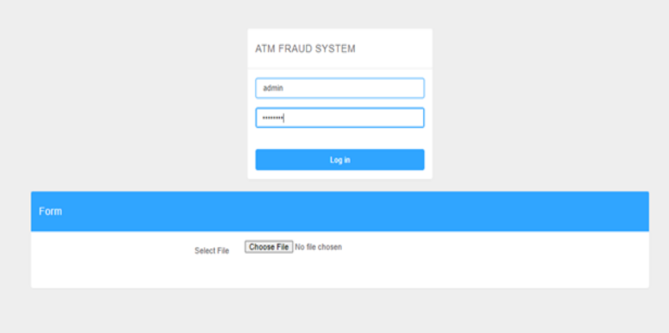


Fig.4 Login Page

The fig.4 describes the login page information

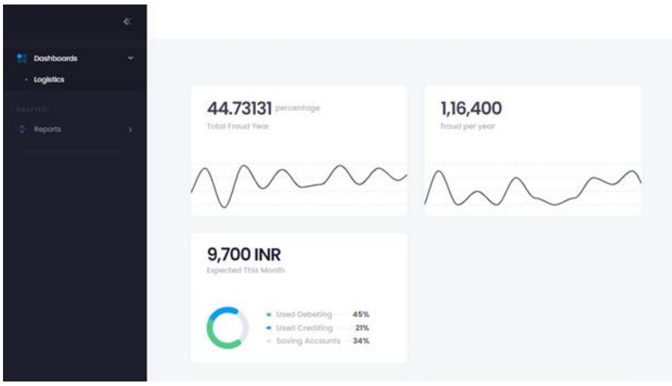


Fig. 5 Dashboard page

The fig.5 describes the Annual fraud detected in system

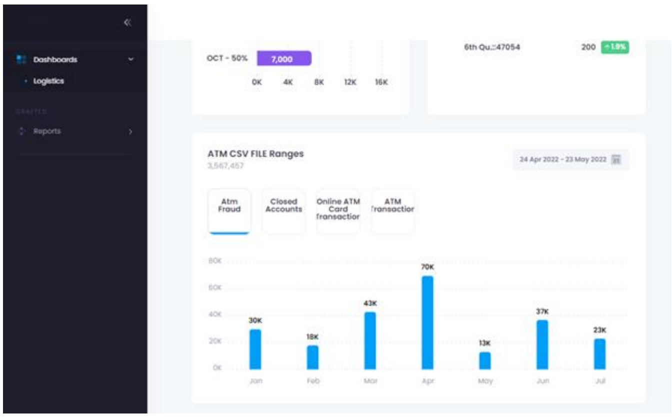


Fig. 6 Bar graph representing the Fraud

The fig.6 describes the detected fraud in bar graph format

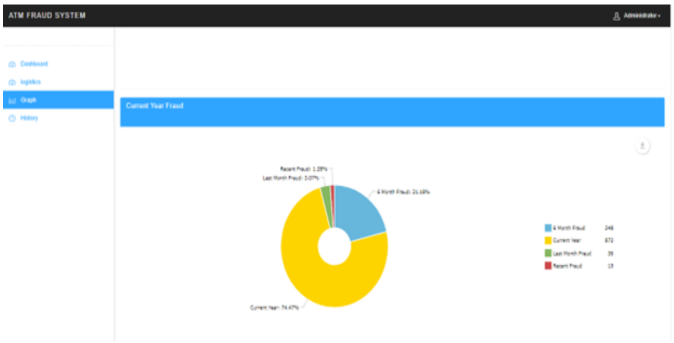
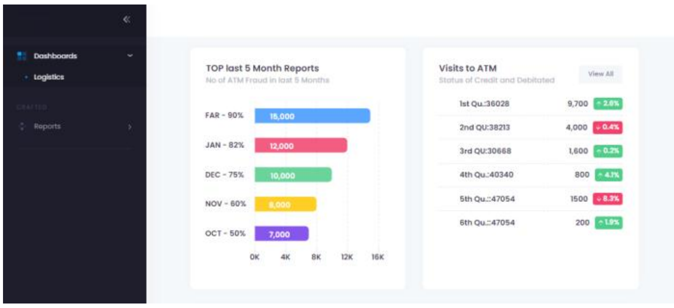
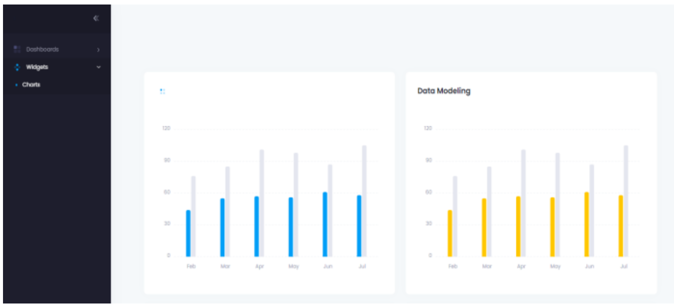


Fig. 7 Pie Chart for Annual detected fraud

The fig.6 describes the detected fraud in bar graph format



1st Qu.:36028	9,700	2.6%
2nd QU:38213	4,000	0.4%
3rd QU:30668	1,600	0.2%
4th Qu.:40340	800	4.1%
5th Qu.::47054	1500	8.3%
6th Qu.::47054	200	1.9%



The graphics above depict the output screens. The output of the algorithm on the original dataset is displayed on these screens. The following columns make up the summary result list:

- Total Fraud in Hours
- Dealings that are fraudulent
- Fraudulent transactions as a graph
- Number of the client

4. CONCLUSIONS

This approach is effective in detecting fraudulent transactions while also minimising false positives. Machine learning approaches are described in this article as being novel in terms of breadth. This method can be expected to be used in a bank's ATM card fraud detection system immediately after a fraudulent transaction happens. Financial organisations can also utilise fraud prevention techniques to protect themselves from large losses and reduce risk. The goal of the inquiry was treated differently from the standard classification work due to altering consequences for misclassification. We opted to apply machine learning techniques across several populations to obtain the optimal parameters because the usual data mining methodology did not match this case well.

REFERENCES

[1] Intelligent Anomaly Detection Model For Atm Booth Surveillance Using Machine Learning Algorithm: Intelligent ATM Surveillance Model:- S. Viji; R. Kannan; N. Yogambal Jayalashmi IEEE 19 Feb,2021

[2] Financial Fraud Detection Using Machine Learning:- C.Maheswara Reddy, Marri Saiteja, B.Shashank, Mrs.Dhikhi ACADEMIA Nov,2020.

[3] Detection of Frauds for Debit Card Transactions at Automated Teller Machine in Indonesia Using Neural Network:-Ermatita, Indrajani Sutedja IOP Publishing 10 Oct,2019

[4] Bethapudi, Dr Prakash & Murthy, G. & Ashok, P. & Prithvi, B. & Kira, S.. (2018). ATM

[5] Card Fraud Detection System using Machine Learning Techniques. International Journal

[6] for Research in Applied Science and Engineering Technology. 5.10.22214/ijraset.2018.4836

[7] Abdullahi, Ibrahim & Mishra, Amit & Ahmad, Barroon. (2010). Fraud Detection and Control on ATM Machines, an Algorithm for Combating Cash and Fund Transfer International journal of Physical Science. 5

[8] R. Laimek and N. Kaothanthong, "ATM Fraud Detection using Behavior Model," 2018 5th Asian Conference on Defense Technology (ACDT), 2018, pp. 21-25, doi:10.1109/ACDT.2018.8593092.

[9] Diebold Inc.(2006,"ATM Fraud and Security", White paper , www.diebold.com, pp.2

[10] Case, P. and S. N. Sisat, "Secured Automatic Teller Machine (ATM) and Cash Deposit Machine (CDM)," vol. 7782, pp. 118–121, (2014)

[11] Bank Indonesia, "Mengenal Kartu Debit & ATM," pp. 1–2, (2009).

[12] Departemen Kebijakan Makroprudensial, "Kajian Stabilitas Keuangan," Igarss 2014, no. 1, pp. 1–5, (2014).

[13] Otoritas Jasa Keuangan, "Bijak Ber-eBanking," Bank Indonesia, (2015).

[14] Auditor General Western Australian, "Fraud Prevention and Detection in the Public Sector," pp. 1–24, (2013).

[15] Jog, Vivek V., and Nilesh R. Pardeshi. "Advanced Security Model for Detecting Frauds in ATM Transaction," International Journal of Computer Applications 95, no. 15, pp. 47-50, (2014). ICONISCSE IOP Conf. Series: Journal of Physics: Conf. Series 1196 (2019) 012076 IOP Publishing doi:10.1088/1742-6596/1196/1/012076 9

[16] Anderka, M., T. Klerx, S. Priesterjahn, and H. K. Büning, "Automatic ATM Fraud Detection as a Sequence Based Anomaly Detection Problem," Proc. 3rd Int. Conf. Pattern Recognit. Appl. Methods (ICPRAM 2014), (2014).

[17] Lepoivre, M. R., C. O. Avanzini, and G. Bignon, "Credit Card Fraud Detection with Unsupervised Algorithms," vol. 7, no. 1, (2016).

[18] Kass, G. V., "An Exploratory Technique for Investigating Large Quantities of Categorical Data," Applied Statistics, Vol. 29, No. 2, pp. 119–127, (1980).

[19] Svigals, J.;"The Long Life and Imminent Death of the Mag-Stripe Card"; IEEE Spectrum; (June 2012); 73-76

[20] M. Paliwal and U. A. Kumar: Neural networks and statistical techniques : A review of applications. Expert Syst. Appl., vol. 36, no. 1. (2009) 2–17

[21] Indrajani, Prabowo, Harjanto, Meyliana, "Learning Fraud Detection from Big Data in Online Banking Transactions: A Systematic Literature Review." Journal of Telecommunication, Electronic and Computer Engineering (JTEC) 8.3 (2016): 127-131.ISSN 2180-1843 eISSN 2289-8131