IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

Credit Card Fraud Detection using Machine Learning

Karan Raghwani[1], Girish Sawant[2], Prasad Kulkarni[3], Dhananjay Gandal[4], Prof. Sujay Pawar[5]
Students, Department Of Computer Engineering[1,2,3,4]
Assistant Professor, Department Of Computer Engineering[5]
Dr. D. Y. Patil Institute of Technology, Pune, Maharashtra, India

Abstract: *In today's world the credit card fraud is the biggest issue and now there is need to combat against the credit card fraud. "credit card fraud is the process of cleaning dirty money, thereby making the source of funds no longer identifiable." On daily basis, the financial transactions are made on huge amount in global market and hence detecting credit card fraud activity is challenging task. As earlier (Anti- credit card fraud Suite) is introduced to detect the suspicious activities but it is applicable only on individual transaction not for other bank account transaction. To Overcomes issues of we propose Machine learning method using 'Structural Similarity', to identify common attributes and behaviour with other bank account transaction. Detection of credit card fraud transaction from large volume dataset is difficult, so we propose case reduction methods to reduces the input dataset and then find pair of transaction with other bank account with common attributes and behaviour.*

Keywords: Credit Card Fraud

I. INTRODUCTION

Credit card fraud scrub as much as 5 of the world's GDP (Gross Domestic Product.) every year. Combating credit card fraud using AI is to detect the suspicious activities. Combating credit card fraud typically requires most entities that complete financial transactions to keep thorough records of their clients' accounts and activities. If they come across any information that appears to be suspicious, they are required to report it to the government for further investigation. In this Transaction records is check to detect credit card fraud activity if the suspicious data is detected. Here we use Artificial Intelligence and Machine Learning Algorithm to detect the suspicious activities and solve it by training the data of that activity. We are going to use supervised and unsupervised algorithm techniques.

1.1 Motivation

The use of machine learning in fraud detection has been an interesting topic now days. A credit card fraud detection algorithm consists in identifying those transactions with a high probability of being fraud, based on historical fraud patterns. Machine learning, having three types, from that also the supervised and hybrid approach is more suitable for fraud detection.

II. DESCRIPTION OF THE PROBLEM

1.1 Problem Definition

The Credit Card Fraud Detection Problem includes modeling past credit card trans-actions with the knowledge of the ones that turned out to be a fraud. This model is then used to identify whether a new transaction is fraudulent or not.

1.2 Goals and Objectives

- Our main scope is on-line looking, fraud detection system.
- To notice and block from fraud transactions employing a master card.
- To detect and block from fraud transactions using a credit card.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

III. Literature Review

Financial Fraud Detection with Anomaly Feature Detection

In recent years, financial fraud activities such as credit card fraud, credit card fraud, increase gradually. These activities cause the loss of personal and/or enterprises' properties. Even worse, they endanger the security of nation because the profit from fraud may go to terrorism [1][25].Thus, accurately detecting financial fraud and tracing fraud are necessary and urgent. However, financial fraud detection is not an easy task due to the complex trading networks and transactions involved. Taking credit card fraud as an example, credit card fraud is defined as the process of using trades to move money/goods with the intent of obscuring the true origin of funds.

A New Algorithm for credit card fraud Detection Based on Structural Similarity

There are many methods of credit card fraud. Criminals can hide the source of money by using the funds in casinos or real estate purchases, or by over valuing legitimate invoices. In general, a credit card fraud procedure is com-posed of three major steps: placement, layering and integration. Placement is the process of introducing the dirty money into the financial system by some mean. Layering is the processing of carrying out complex transactions to hide the source of the funds. Finally, integration is to withdraw the proceeds from a destination bank account. The purpose of performing complex layering is to confuse anti-credit card fraud instruments.

Credit Card Fraud Detection Using RUS and MRN Algorithms

Currently, enterprise systems have been focusing on expenditure services through credit card broadly because it is convenient and quick to pay for products and services. Thus, this research emphasizes on the fraud detection of credit card payment by using the machine learning technique called RUSMRN. The proposed method adopts three base classifiers which are MLP,B and Naıve Bayes algorithms. In addition, it can analyze the correctness to work with the unbalance data sets. Therefore, this research is focusing on the information of the credit card company of Taiwan for collecting data of customer behaviors in credit card payment. After that, it has brought the information to make prediction for correctness whether it has the risks in payment. The result shows that the proposed method can achieve the best classification performance in terms of accuracy and sensitivity.

Data set shift quantification for credit card fraud detection

Machine learning and data mining techniques have been used extensively in order to detect credit card frauds. However purchase behaviour and fraudster strategies may change overtime. This phenomen on is named data set shift [1] or concept drift in the domain of fraud detection [2]. In this paper, we present a method to quantify day-by-day the dataset shift in our face-to-face credit card transactions dataset (card holder located in the shop) . In practice, we classify the days against each other and measure the efficiency of the classification. The more efficient the classification, the more different the buying behaviour between two days, and vice versa. Therefore, we obtain a distance matrix characterizing the data set shift. After an agglomerative clustering of the distance matrix, we observe that the data set shift pattern matches the calendar events for this time period (holidays, week-ends, etc).We then incorporate this dataset shift knowledge in the credit card fraud detection task as a new feature. This leads to a small improvement of the detection.

Real-time Credit Card Fraud Detection Using Machine Learning

Credit card fraud event stake place frequently and then result in huge financial losses[1]. The number of online transactions has grown in large quantities and online credit card transactions holds a huge share of these transactions. Therefore, banks and financial institutions offer credit card fraud detection applications much value and demand. Fraudulent transactions can occur in various ways and can be put into different categories. This paper focuses on four main fraud occasions in real-world transactions. Each fraud is addressed using a series of machine learning models and the best method is selected via an evaluation. This evaluation provides a comprehensive guide to selecting an optimal algorithm with respect to the type of the frauds and we illustrate the evaluation with an appropriate performance measure. Another major key area that we address in our project is real time credit card fraud detection. For this, we take the use of predictive analytics done by the implemented machine learning models and an API module to decide if a particular transaction is genuine or fraudulent. We also as sessa novel strategy that effectively addresses the skewed distribution of data. The data used in our experiments come from a financial institution according to a confidential disclosure agreement.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

IV. SYSTEM DESIGN AND FLOW

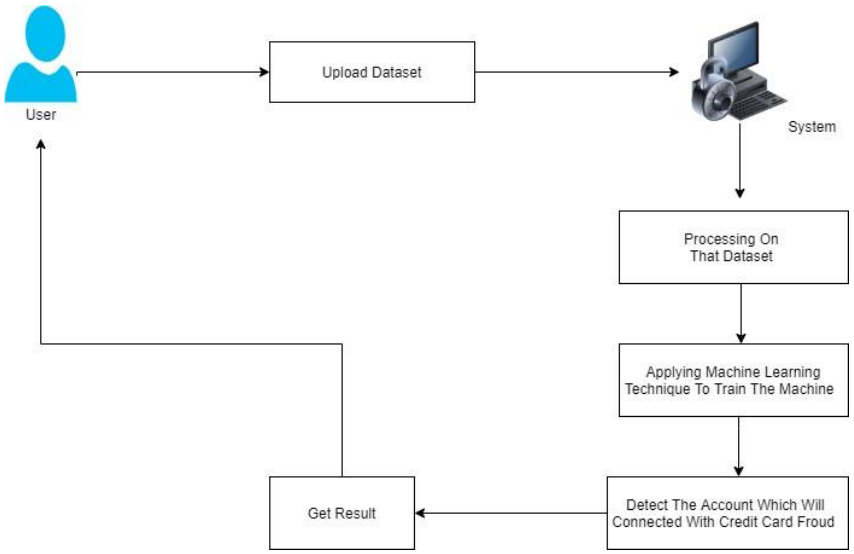

Figure: System Architecture

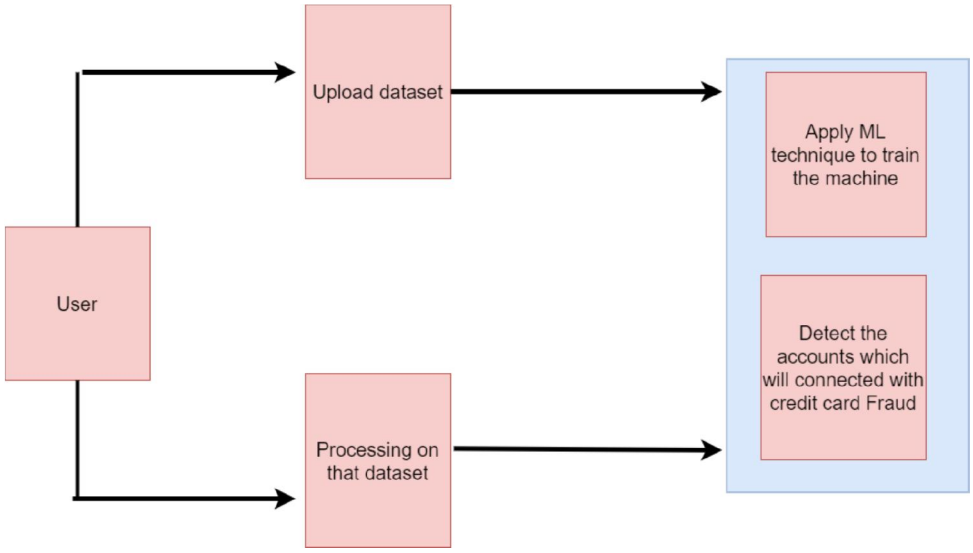

Figure: System Flow

4.1 Convolutional Neural Networks

As in any other neural network, the input of a CNN, in this case an image, is passed through a series of filters in order to obtain a labelled output that can then be classified. The specificity of a CNN lies in its filtering layers, which include at least one convolution layer. These allow it to process more complex pictures than a regular neural network. Whereas the latter is well adapted for simple, well centered images such as hand-written digits, the use of CNNs in image analysis ranges from Facebook's automatic tagging algorithms, to object classification and detection, in particular in the field of radiology.
There are four types of layers in Convolutional Neural Networks:

1. **Convolutional Layer:** In a typical neural network each input neuron is connected to the next hidden layer. In CNN, only a small region of the input layer neurons connect to the neuron hidden layer.
2. **Pooling Layer:** The pooling layer is used to reduce the dimensionality of the feature map. There will be multiple activation pooling layers inside the hidden layer of the CNN.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

3. **Flatten:** Flattening is converting the data into a 1-dimensional array for in- putting it to the next layer. We flatten the output of the convolutional layers to create a single long feature vector.
4. **Fully-Connected layer:** Fully Connected Layers form the last few layers in the network. The input to the fully connected layer is the output from the final Pooling or Convolutional Layer, which is flattened and then fed into the fully connected layer.

V. PROJECT IMPLEMENTATION

- **Pandas:** Pandas is an open-source library that is made mainly for working with relational or labeled data both easily and intuitively. It provides various data structures and operations for manipulating numerical data and time series. This library is built on top of the NumPy library.
- **NumPy:** NumPy is a Python library used for working with arrays. It also has functions for working in domain of linear algebra, fourier transform, and matrices.
- **import cv2**:All packages contain Haar cascade files. cv2.data .haar cascades can be used as a shortcut to the data folder.
- **Pillow:** Pillow is the friendly PIL fork by Alex Clark and Contributors. PIL is the Python Imaging Library by Fredrike Lundh and Contributors.

Figure: Main Window

This is the main window or we can say the first that appears when we run our application. It basically contains three buttons Login, Register and Exit as shown in the above picture. With a background image and our applications title. As per the users requirements the user clicks on the buttons and the next window will appear for the further process.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

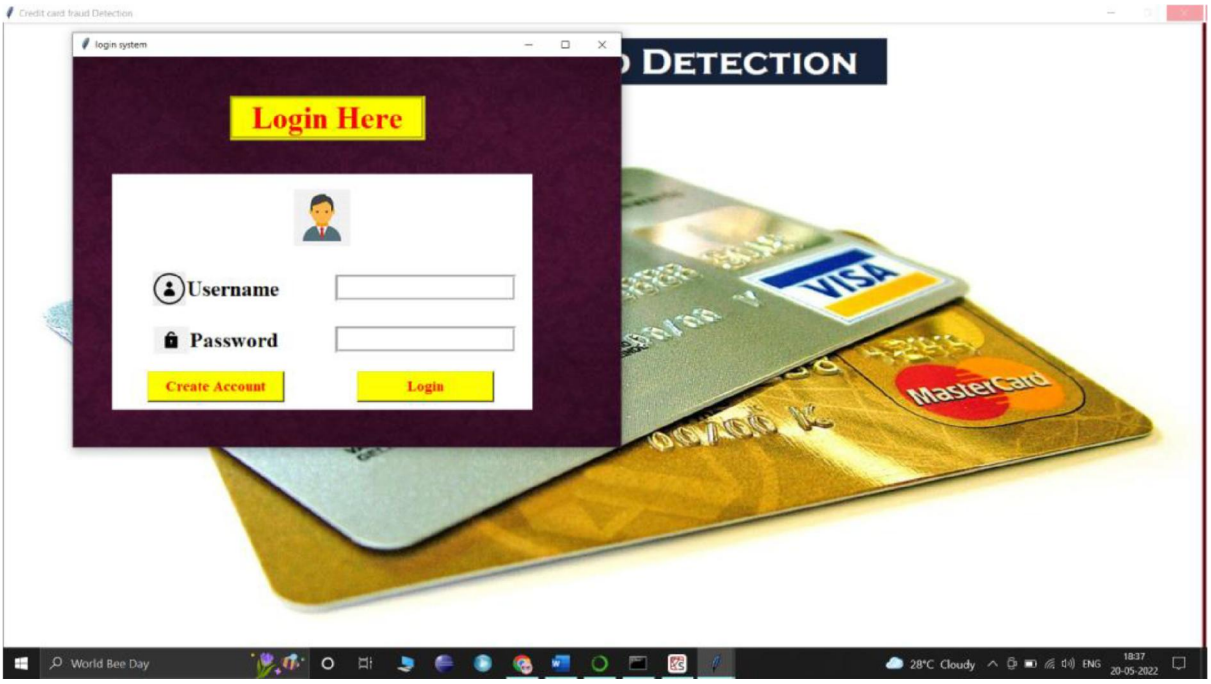

Figure: User Login Window

The next window is our login window where user can login to his/her account where all the users card details and transactions are saved. And head towards the next process and window.

Figure: Registration Window

The next window is registration window. When a new wants to create account this window helps them for the same. Here user has to provide all the required details to successfully create new account. We have provided some format for specific input field such as phone shouldn't more than ten digits or less than ten digits, which helps to organize the data for further use.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

With all the inputs provided correctly user is able to create a new account and a message box appears to let the user know that the account is created successfully and start testing their transactions.

Figure: Card Selection Window

This is the actual window where user can select the card they own and accordingly select them. If they have a banking credit card they can go with card master option or if they have merchant then they can go with it.

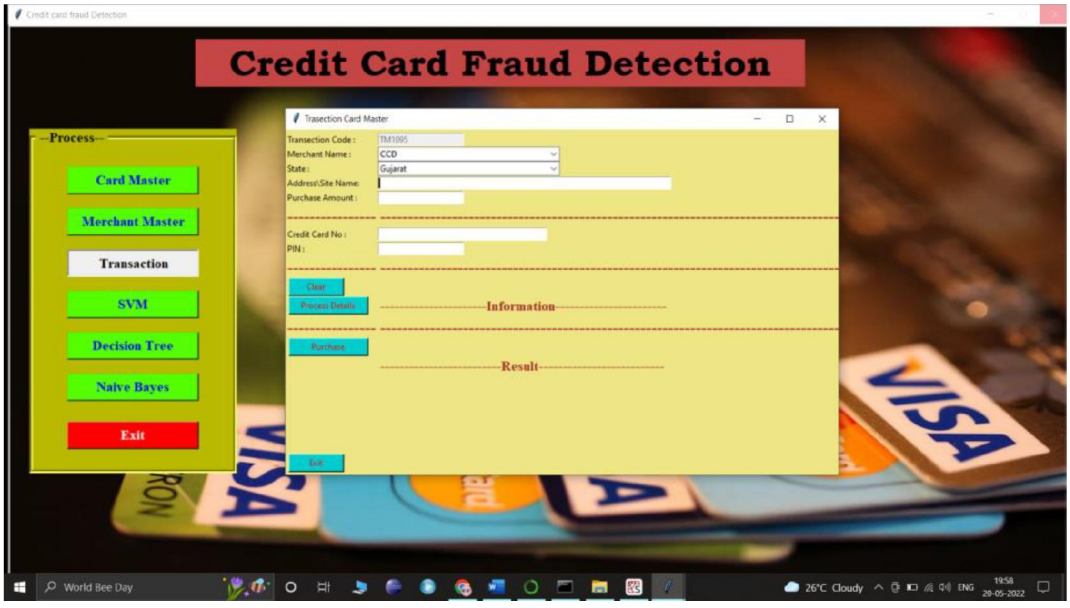

Figure: Transaction Window

This is the final window where transaction details are to be provided and the system shows whether the transaction is a fraud transaction or it is a genuine transaction using the three algorithms that are:

- SVM
- Decision Tree
- Naive Bayes

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

We train the system for the dataset and algorithms provides accuracy of their output, for our system SVM and Decision Tree provides the highest accuracy among three and Naive Bayes provide quite high but not as much as SVM and Decision Tree does.

Figure: SVM Algorithms Accuracy Window

Figure: Decision Tree Accuracy Window

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

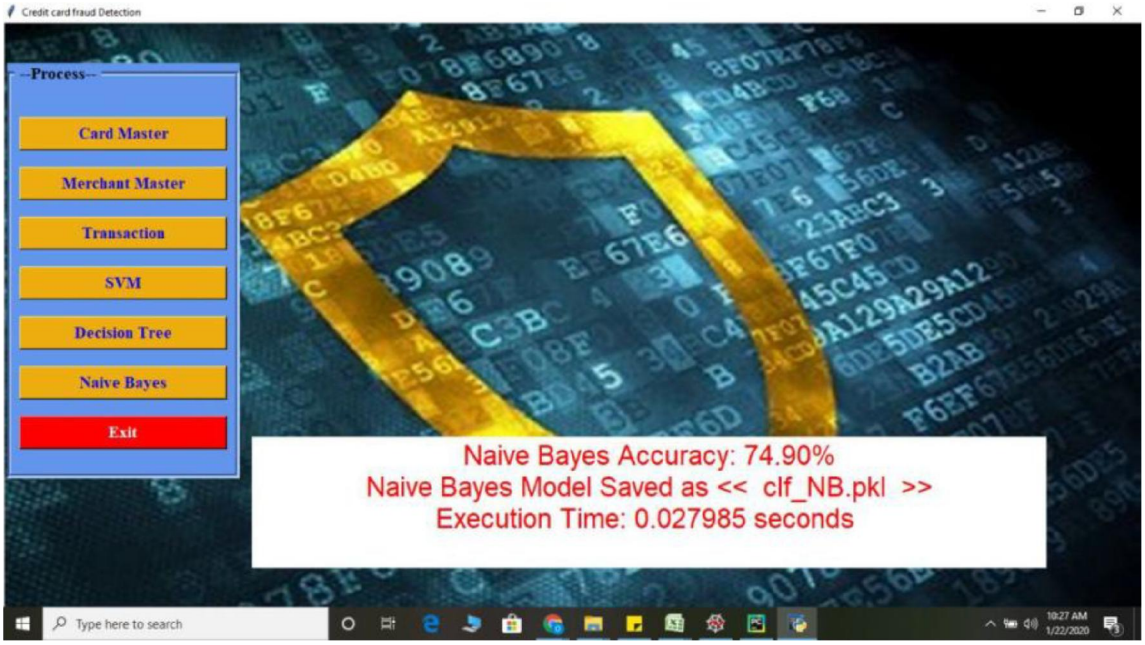

Figure: Naïve Bayes Algorithm Accuracy Window

VI. Advantages And Disadvantages

6.1 Advantages

1. Reduction in number of fraud transactions.
2. User can safely use his credit card for online transaction.
3. Added layer of security.
4. User friendly system.
5. We can identify and group potential credit card fraud accounts.

6.2 Limitations

If details is incorrect format in database then it could be output getting wrong.

6.3 Applications

It is used to predict whether a credit card transaction is fraudulent or not, based on the transaction amount, location and other transaction related data. to track down credit card transaction data, which is done by detecting anomalies in the transaction data. The AI model can also offer cause codes for the transaction being flagged.

VII. Conclusion

The proposed ML framework aims to find potential money-laundering groups among a large number of financial transactions. In order to improve the efficiency of the framework, case reduction methods such as matching transaction detection and balance score filter are used to narrow down the list of potential ML accounts. Next by taking advantage of structural similarity, we can identify and group potential credit card fraud accounts. Our preliminary experimental results show a high degree of accuracy in detection of ML accounts.

Acknowledgements

The completion of our project brings with it a sense of satisfaction, but it is never complete without those people who made it possible and whose constant support has crowned our efforts with success. One cannot even imagine our completion of the project without guidance and neither can we succeed without acknowledging it. It is a great pleasure that we acknowledge the enormous assistance and excellent co-operation to us by the respected personalities.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 4, May 2022
Impact Factor: 6.252

REFERENCES

[1]. "Fatf-gafi.org - Financial Action Task Force (FATF)", Fatf-gafi.org,2016. [Online]. Available: http://www.Fatf-gafi.org. [Accessed: 22-Dec- 2015].

[2]. Fatf-gafi.org, 'credit card fraud - Financial Action Task Force (FATF)', 2014. [On- line]. Available: http://www.fatfgafi.org/faq/moneylaundering/. [Accessed: 22- Dec- 2015].

[3]. Neo4j Graph Database, 'Neo4j, the World's Leading Graph Database', 2014. [On- line]. Available: http://neo4j.com/. [Accessed: 22- Dec- 2015].

[4]. A. C. Bahnsen, A. Stojanovic, D. Aouada, and B. Ottersten. Improving credit card fraud detection with calibrated probabilities. In SDM, 2014.

[5]. M. Gupta, J. Gao, C. C. Aggarwal, and J. Han. Outlier Detection for Tempo- ral Data. Synthesis Lectures on Data Mining and Knowledge Discovery, Morgan Claypool Publishers, 2014.

[6]. Clarke, M. 1994. 'Fraud and the Politics of Morality'. Business Ethics: A European Review, 3: 2, 117-122.

[7]. Encyclopedia Britannica, no date. Credit Card. (Accessed: October 2008).

[8]. Euromonitor International, 2006. Financial cards in Germany Available (Accessed: November 2006).

[9]. European e-Business Market Watch. 2005. ICT Security, e-Invoicing and e-Payment Activities in European Enterprises, Special Report, September.

[10]. Ezawa, K. Norton, S. 1996. 'Constructing Bayesian Networks to Predict Uncollectible Telecommunications Accounts'. IEEE Expert, October; 45-51.