IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CREDIT CARD FRAUD DETECTION SYSTEM USING MACHINE LEARNING

[1] Arati Kale, [2]Sakshi Akangire, [3]Sakshi Akangire, [4]Apurva Vajarinkar, [5]Snehal Nalawade

[1]Prof., [2]Student, [3]Student, [4]Student, [5]Student

[1]Department of Information Technology,

[1]Department of Information Technology, JSCOE,

Pune.

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 butit 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 andthen find pair of transaction with other bank account with common attributes and behaviour.

***Key Words*: Fraud Detection, Credit Card, SVM, Decision Tree, Machine Learning.**

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.

2.PROBLEM STATEMENT

The aim of this project is to detect the suspicious activities of bank transaction to identifying the Credit card Fraud . Also to reduces the amount of criminal activities. The people involved in Credit card obviously try to conceal the real purpose of money transfer used in this process.

3.LITERATURE SURVEY

Fraud act as the unlawful or criminal deception intended to result in financial or personal benefit. It is a deliberate act that is against the law, rule or policy with an aim to attain unauthorized financial benefit.

Numerous literatures pertaining to anomaly or fraud detection in this domain have been published already and are available for public usage. A comprehensive survey conducted by Clifton Phua and his associates have revealed that techniques employed in this domain include data mining applications, automated fraud detection, adversarial detection. Unconventional techniques such as hybrid data mining/complex network classification algorithm is able to perceive illegal instances in an actual card transaction data set, based on network reconstruction algorithm that allows creating representations of the deviation of one instance from a reference group have proved efficient typically on medium sized online transaction. The fraud detection is a complex task and there is no system that correctly predicts any transaction as fraudulent.

4. METHODOLOGY

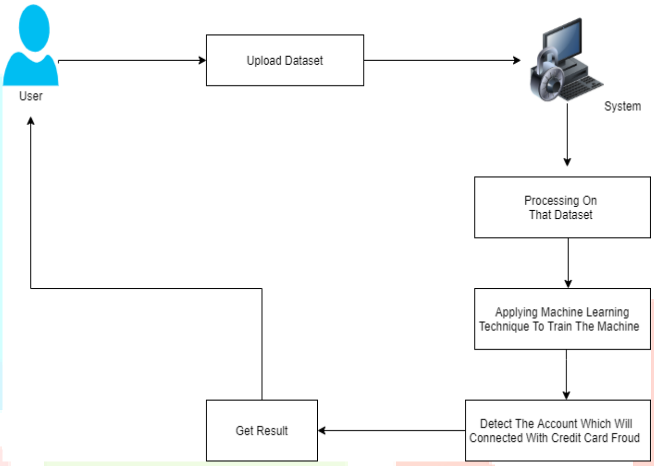

User
Upload Dataset
System
Processing On That Dataset
Applying Machine Learning Technique To Train The Machine
Detect The Account Which Will Connected With Credit Card Froud
Get Result

Figure 1: System Architecture

At The Balance, we are dedicated to giving you unbiased, comprehensive credit card reviews. To do this, we collect data on hundreds of cards and score more than 55 features that affect your finances, such as interest rates, fees, and rewards. We score each attribute on a scale of 0 to 5. We then weight these scores to determine the star ratings you see on our review pages. The following elements are generally ordered from highest to lowest in how heavily they factor into our overall evaluation of credit cards. Clearly, credit card fraud is an act of criminal dishonesty.

This article has reviewed recent findings in the credit card field. This paper has identified the fraud, and discussed measures to detect them. Such measures have included clustering techniques algorithms. From an ethical perspective, it can be argued that banks and credit card companies should attempt to detect all fraudulent cases. Yet, the unprofessional fraudster is unlikely to operate on the scale of the professional fraudster and so the costs to the bank of their detection may be uneconomic.

The bank would then be faced with an ethical dilemma. Should they try to detect such fraudulent cases or should they act in shareholder interests and avoid uneconomic costs? As the next step in this research program, the focus will be upon the implementation of a 'suspicious' scorecard on a real data-set and its evaluation. The main tasks will be to build scoring models to predict fraudulent behavior , taking into account the fields of behavior that relate to the different types of credit card fraud identified in this paper, and to evaluate the associated ethical implications.

The plan is to take one of the European countries, probably Germany, and then to extend the research to other EU countries. clustering techniques for behavioural fraud. The peer group analysis is a system that allows identifying accounts that are behaving differently from others at one moment in time whereas they were behaving the same previously. Those accounts are then flagged as suspicious. Fraud analysts have then to investigate those cases.

The hypothesis of the peer group analysis is that if accounts behave the same for a certain period of time and then one account is behaving significantly differently, this account has to be notified. Breakpoint analysis uses a different approach. The hypothesis is that if a change of card usage is notified on an individual basis, the account has to be investigated. In other words, based on the transactions of a single card, the break-point analysis can identify suspicious behavior . Signals of suspicious behavior are a sudden transaction for a high amount, and a high frequency of usage.

5.ALGORITHM

1.Decision Tree:

A decision tree is a non-parametric supervised learning algorithm, which is utilized for both classification and regression tasks. It has a hierarchical, tree structure, which consists of a root node, branches, internal nodes and leaf nodes.

2.SVM Algorithm:

Support Vector Machine(SVM) is a supervised machine learning algorithm used for both classification and regression. Though we say regression problems as well its best suited for classification. The objective of SVM algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points We use SVM for identifying the classification of genes, patients on the basis of genes and other biological problems. Protein fold and remote homology detection – Apply SVM algorithms for protein remote homology detection. Handwriting recognition – We use SVMs to recognize handwritten characters used widely. The objective of applying SVMs is to find the best line in two dimensions or the best hyperplane in more than two dimensions in order to help us separate our space into classes. The hyperplane (line) is found through the maximum margin, i.e., the maximum distance between data points of both classes

Input Layers:
It's the layer in which we give input to our model. The number of neurons in this layer is equal to the total number of features in our data (number of pixels in the case of an image).

Hidden Layer: The input from the Input layer is then feed into the hidden layer. There can be many hidden layers depending upon our model and data size. Each hidden layer can have different numbers of neurons which are generally greater than the number of features. The output from each layer is computed by matrix multiplication of output of the previous layer with learnable weights of that layer and then by the addition of learnable biases followed by activation function which makes the network nonlinear. Output Layer: The output from the hidden layer is then fed into a logistic function like sigmoid or softmax which converts the output of each class into the probability score of each class.

6. RESULT AND DISCUSSION

Project Task Set Major
Tasks in the Project stages are:

• Task 1: correctness

• Task 2: availability

• Task 3: integrity

1.Task Network

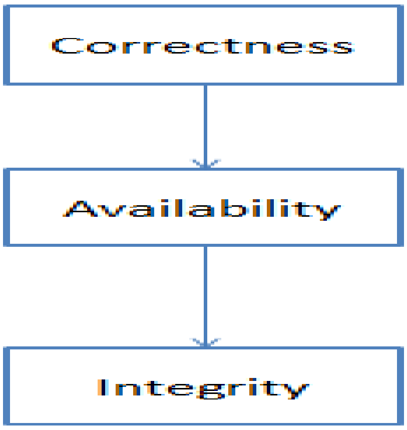
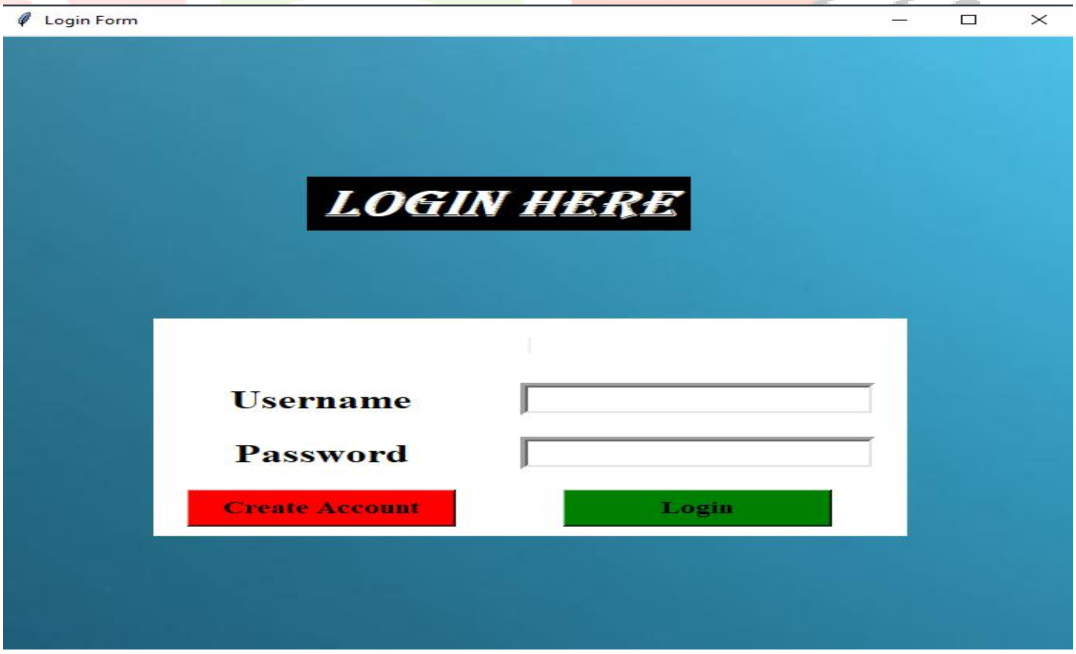


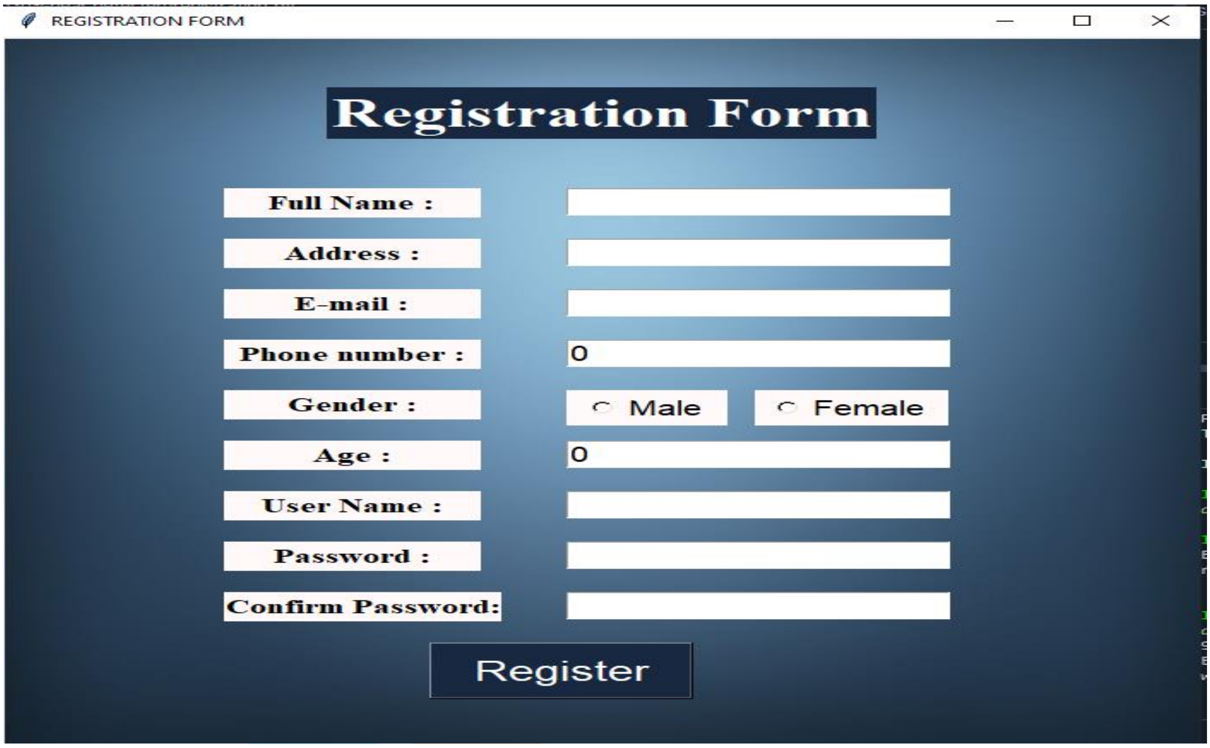
Fig: Task Network

Login Test Cases:

Test Case ID	Test Case	Test Case I/P	Actual Result	Expected Result	Test case criteria(P/F)
001	Enter The Wrong username or password click on submit button	Username or password	Error comes	Error Should come	P
002	Enter the correct username and password click on submit button	Username and password	Accept	Accept	P



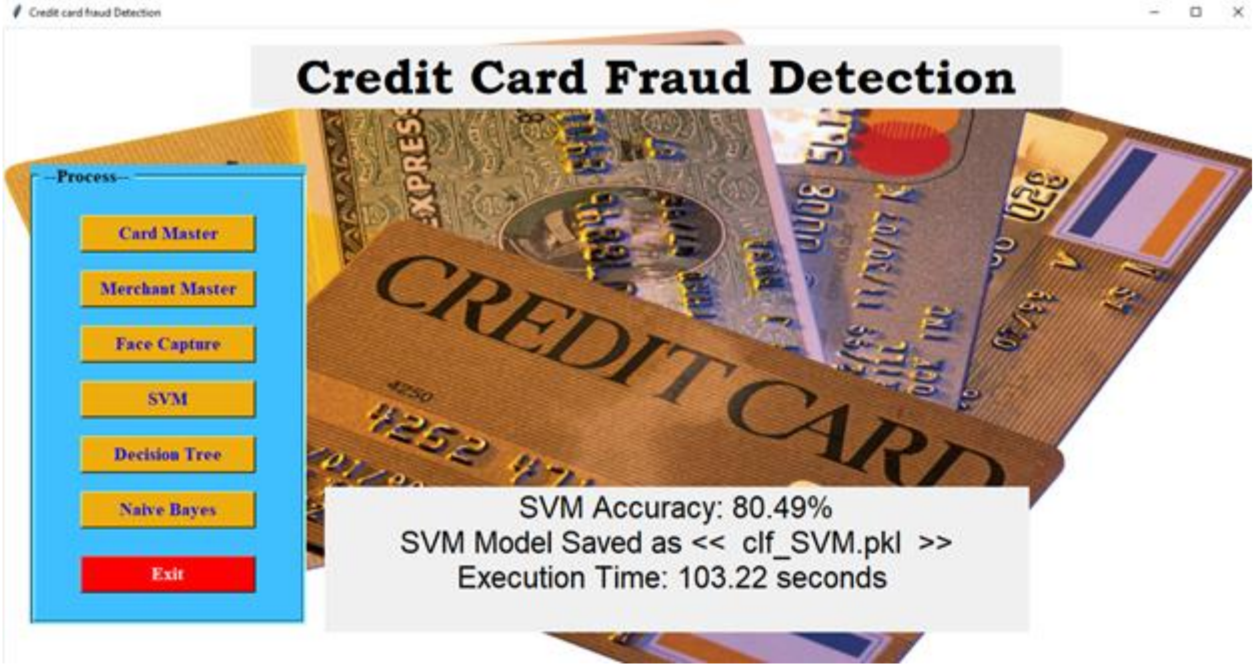
Resistration test Cases:



Phone number : 0

Age : 0

Test Case ID	Test Case	Test Case I/P	Actual Result	Expected Result	Test case criteria(P/F)
001	Enter the number in username, middle name, last name field	Number	Error Comes	Error Should Comes	P
001	Enter the character in username, middle name, last name field	Character	Accept	Accept	p
002	Enter the invalid email id format in email id field	Kkgmail.com	Error comes	Error Should Comes	P
002	Enter the valid email id format in email id field	kk@gmail.com	Accept	Accept	P
003	Enter the invalid digit no in phone no field	99999	Error comes	Error Should Comes	P
003	Enter the 10 digit no in phone no field	9999999999	Accept	Accept	P



SVM Accuracy: 80.49%
SVM Model Saved as << clf_SVM.pkl >>
Execution Time: 103.22 seconds

7.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.

REFERENCES

1 [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 [2] Fatf-gafi.org, ’credit card fraud - Financial Action Task Force (FATF)’, 2014. [Online]. Available: http://www.fatfgafi.org/faq/moneylaundering/. [Accessed: 22- Dec- 2015].

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

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

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