ISSN2394-3777 (Print)
ISSN2394-3785 (Online)
Available online atwww.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology
(IJARTET) Vol. 5, Special Issue 14, April 2018

CREDIT CARD FRAUD DETECTION MODEL USING BIG DATA TECHNOLOGY

Dr. SOHAN KUMAR GUPTA[1], RAMESH SINGH YADAV[2], RAJU KHADKA[3], KRISHNA NARAYAN BANIYA[4], SATISH KUMAR[5]

[1]*Head of Department, Department of computer science and engineering,BTI,Bengaluru,India*
[2]*UG Student, Department of computer science and engineering, BTI, Bengaluru, India*
[3]*UG Student, Department of computer science and engineering, BTI, Bengaluru, India*
[4]*UG Student, Department of computer science and engineering, BTI, Bengaluru, India*
[5]*UG Student, Department of computer science and engineering, BTI, Bengaluru, India*

ABSTRACT

In this paper, we focus on designing an online credit card fraud detection model using big data technology. Along with the advancement in the technology nowadays credit card has become one of the prominent and the most used payment method in every area irrespective of it is an online or regular purchase. So in this case all of the users or the card holders may not be smart enough to secure their card by themselves. Hence we are developing a model that can detect if any fraud activities are going on with the card without the knowledge of the card holder or not by using the machine learning having algorithm like Naïve Bayes, logistics regression. If our model finds out something unusual or fraud cases are going on with the particular creditcard, then transaction will be stopped and the user will be taken to give more verification.

Keywords*- credit card: big data technology; fraud detection; machine learning; Naïve Bayes; Logistics regression*

1.INTRODUCTION

Abhinav Srivastava et al [1] the popularity in the use of online credit card has been growing day by day. By looking to the current scenario more than 80% of the people make use of the credit card for their payment. Either of the way that it's easy to make payment in online or offline and it may be so also just because of the profitable usage policy of credit card company and the merchants they are buying from. Whatever the reason may be but the study shows that the credit card's usage is increased rapidly. So as its usage is increasing in a huge ratio day by day it makes a sense to have the chances of the card being misused by other so-called bad intension people. Hence in this project we trying to do something on this issue to detect such kind of credit card fraud activities. So here in this project we are going to train a model using machine learning algorithms with big data technology here we will have a two component one is genuine training model and the other is fraud training model. [1] Then based on the previous and the current user behavior we collect the data and then we fetch the data into our model. Then our model slowly learns to detect both the fraud and the genuine cases hence in the future days if user uses the card it(model) checks the current behavior with its previous behavior and if this behavior matches with the fraud case then transaction will not be proceeded and user will be redirected to make some verification. If the behavior matches with the genuine scenario. Then transaction will not be disrupted.

2. LITERATURE REVIEW

Tej Paul et al [2] Plenty of the literature survey have been done in finding out the credit card fraudulent activities detection and prevention. And some of the author have made a very good case studies and other authors like NMalini,Dr. M. Pushpa and lot more people have done the research and published their paper to detect the fraudulent activities by using some outlier detection technique. The author uses a very popular algorithm so-called KNN algorithm with the outlier detection technique.The outlier detection technique is easily performed in two ways:

1. By using supervised machine learning algorithm.
2. By using unsupervised machine learning algorithm.

Supervised way for outlier detection: Here the data are given as a pre-labelled set of information that are organized in a way so that the model can treat them as a data to be analyzed.

Unsupervised way of the outlier detection: The inputs here are the raw data and these data are not so organized data which model can use directly to recognize as an input. The data are associated and clustered and k-means algorithms help is taken here.

Some other author like AmalnKundu,Samik Sural have also used a model so-called HMM (Hidden Markov Model).In this model the fraudulent activities are decided based on the transaction behavior of the cash holder. If there is a huge deviation in the transactionbehavior, then the approach says some fraud cases are going on.

A. EXISTING SYSTEM

The existing way of predicting the credit card fraud detection and the prevention of this issue was a login verification in the back-end and the two and fro transfer of the server generated secrete key between the user and the server. And the other way of detecting the fraud activities was testing the user spending behavior. It was not so convenient to work with the huge amount of dataset or the large transaction were not captured easily to test and validate.

B.PROPOSED SYSTEM:

We are going to develop a training model using big data technology like Spark on Hadoop platform, our training model is not limited to spending behavior alone. But, it also analyzes a lot more user related data to detect and distinguish between fraudulent and genuine transaction. Such as usual password and card details entering keypad speed, Location variation, frequent change of the card accessing devices etc. And the transaction data is passed to our training model via MVC. And it contains: (See fig.1)

1.Controller
2.Services
3.DAO (Data Access Object)
4.Apache Kafka
5.Spark Engine

Benefits of proposed system:

- Fast fraud detection in Real-time.
- Feasibility of standalone working.
- Independent Platform compatibility
- Enables massive data storage mechanism.

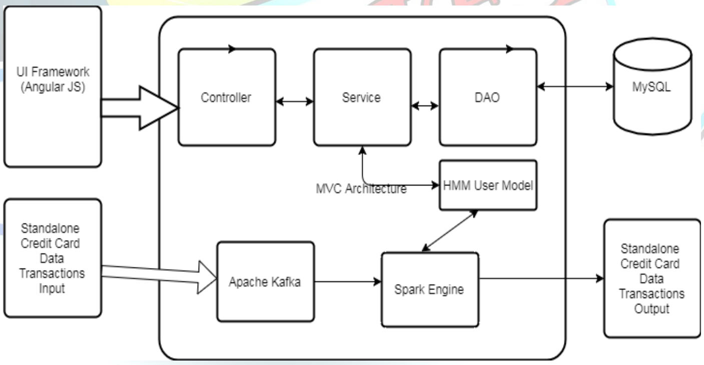
UI Framework (Angular JS)
Controller
Service
DAO
MySQL
MVC Architecture
HMM User Model
Standalone Credit Card Data Transactions Input
Apache Kafka
Spark Engine
Standalone Credit Card Data Transactions Output

fig.1. Architectural diagram

3. METHODOLOGY

A. Data Preparation

Because of the bank's non-disclosure agreement (NDA) we are initially taking the dataset from the public domain. [3] And the data can be further made more labelled or organized by capturing the transaction and other user behavior. Each and every time the uses the credit card more data related to the card will be generated and these huge collected data are fetched to our training model.

B. Training Model

Collected dataset area fed into our training model. The training model consist of two parts:

1.Genuine transaction data
2.Fradulent transaction data

Genuine transaction data:

Certain parameters are set to resemble that the data are genuine data and its labelled as a genuine kind of data so this process is continued each and every time the user makes further transaction. Hence in future if the transaction made matches this category then the transaction is not disrupted.

Fraudulent transaction data:

Again in this section of the training model, these kind of data are the data which do not fall into the category of the genuine data. So in the future days, once again the same scenario of capturing data will be repeated at each and every time the user makes transaction. So the data that fall into fraud category are used to build the model that can detect the fraud type of transaction. If the entered data is not matching with the dataset which we have then it will cluster the different behaviors and different activities based on which we are going to predict is the transaction genuine or not. Once our model predicts the chance of transaction being more fraudulent than genuine it will not allow for the further transaction and makes user to go through other verification process.

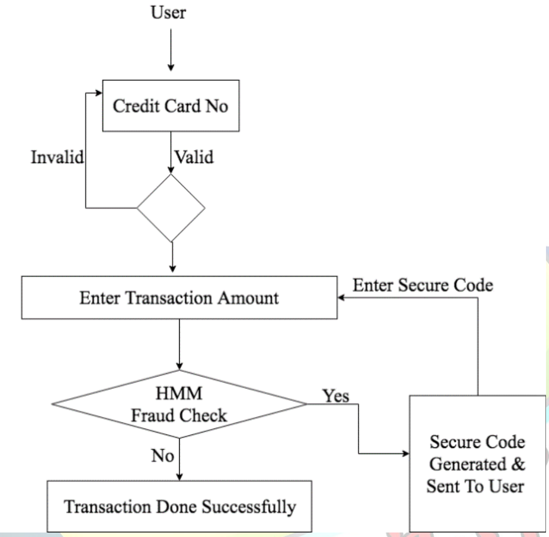

Fig.2. Workflow diagram

4. CONCLUSION

In this paper, we have proposed a model for credit card fraud detection.[4] The different steps in credit card transaction processing are represented as the underlying stochastic process of training model. We have used the ranges of transaction amount as the observation symbols, whereas the types of item have been considered to be states of the model building. [5] We have suggested a method for finding the spending profile of cardholders, as well as application of this knowledge in deciding the value of other kind of observations like very frequent change of the accessing devices. It has also been explained how our model can detect whether an incoming transaction is fraudulent or not. The system is also scalable for handling large volumes of transactions. Credit card fraud can cost businesses billions of dollars each year, making fraud detection and prevention crucial to maintaining a successful company. [6] Educating yourself and your employees about credit card fraud is the first step in detecting unauthorized transactions. Although not all fraudulent charges can be prevented, implementing some of the practices discussed can give your business an advantage over criminals.

REFERENCES

[1] Abhinav Srivastava, AmlanKunda, Shamik Sural, and Arun K. Majumdar" Credit Card Fraud Detection Using Hidden Markov Model" VOL. 5, NO. 1, JANAUARY –MARCH 2008

[2] Tej Paul Bhatla, VikramPrabhu, & Amit Dua, "Understanding Credit Card Frauds".Tata Consultancy Services. June 2003

[3] Statistic Brain Research Institute (2014, July 12). Credit Card Fraud Statistics (2014). Available: http://www.statisticbrain.com/credit-cardfraud-statistics/

[4] Haiying Ma & Xin Li, "Application of Data Mining in Preventing Credit Card Fraud" in International Conference on Management and Service Science, 2009 ©, DOI: 10.1109/ICMSS.2009.5304330, pp 1 – 6

[5] Wen-Fang Yu & Na Wang, "Research on Credit Card Fraud Detection Model Based on Distance Sum" in International Joint Conference on Artificial Intelligence, 2009 ©, DOI: 10.1109/JCAI.2009.146, pp 353 – 356

[6] Republic Act No. 1405 – An act prohibiting disclosure of or inquiry into, deposits with any banking institution and providing penalty therefore (01 November 2015). Available: http://www.pdic.gov.ph/index.php?nid1=10&nid2=3