Credit Card Fraud Detection Using Clustering Based Approach

Yogesh Bharat Sonawane [1], Akshay Suresh Gadgil [2], Aniket Eknath More4 [3], Niranjan Kamalakar Jathar [4]

1,2,3,4 Student, Department of Computer Engineering, Matoshree College of Engineering and Research Center Nashik, Savitribai Phule, Pune University, Maharashtra, India

ABSTRACT

Now days the use of credit card has dramatically increased. As credit card become the most popular manner for both online as well as orderly purchase cases of fraud joined with it are also increasing. In this paper, k-means clustering is used for credit card fraud detection. Data is developing haphazardly for credit card and the k-means algorithm is used for discovering transaction weather it is fraud transaction or legitimate transaction. Cluster is developed to discover fraud in credit card fraud transaction which is divided into low risky cluster, high risky cluster and medium risky cluster. K-means clustering algorithm is easy and powerful algorithm for credit card fraud detection.

Keyword: *Fraud detection, Electronic commerce, Credit card fraud, Clustering, Outlier detection, data mining.*

---

1. INTRODUCTION

In modern world most of the people are using credit cards. It is the most popular purchase manner-means clustering assists in gathering the transaction dataset into number of clusters that assists in unrelated retrieval of dataset. K-means cluster analysis is a technique for smashing dataset down into narrated constituents in such a way that patterns and order becomes perceivable. Searching outliers is a main duty in k-means clustering. The outlier detection is a bough of data mining technique has lot of consequence applications and worthy of large attention from data mining ownership. In modern years, assembly database querying technique are not enough to pull out important information, and therefore investigation nowadays are tried to implement new techniques to meet the increased needs. It is to be remarked that the addition in dimensionality of transaction dataset gives rise to a number of new computational invitation because of increase in number of dataset and terms also.

Outlier detection is a main investigation problem that aims to search dataset that are considerably alike and inconsistent in the database. Therefore, there is a need to implement a model to discover these frauds in the business. Architecture of credit card fraud detection system is shown in fig 2.4. User first request for transaction. Then it will enter username and password. If it is valid, then it will allow for transaction. Otherwise user requires login again. It also ask security question for cardholder. If it is not valid, then it will detect fraud transaction. If it valid, then it will detect legitimate transaction.

2. SYSTEM DESCRIPTION

The following figure shows the system architecture of Credit Card Fraud Detection Using Clustering Based Approach

A. Transaction variable declaration:

First, the variables used in this program are declared such as transaction amount, credit card number, city, transaction, transaction date, merchant category id, transaction id and country [3].

B. Transaction Validation:
It shows the validity of the particulars needed for the transaction such as transaction amount, credit card number, city, transaction, transaction date, merchant category id, and transaction id and transaction country. Think the transaction amount, name of transaction country is entered and if you forget to enter the credit card number then it will not able to process further as well as the dialog box will open with an instruction to enter the credit card number. All the details are filled it will work further [3].

C. Constructing an entry of the input transaction into database:
The table that is produced before is now entered in the database. It contains transaction country, transaction amount, transaction id, transaction city, credit card number, merchant category id, transaction date [3].

D. Getting training transaction data set:
The data which is removing from the table is now entered to get transaction data [3].

E. Transforming list of transaction data object to multi-dimensional array:
Here, the array is used so that the transaction detail will produce row wise. They are transaction amount, credit card number, transaction country and transaction id etc [3].

F. Assigning cluster name:
The three clusters which are formed are named/labeled as low cluster, high cluster, and middle risky cluster [3].

G.Commit transaction to database either as fraud transaction or legitimate transaction:
The current transaction details were taken and by using k- means clustering algorithm the fraud is discovered. If it is fraud then the message will display 'fraud transaction' or else it will display 'legitimate transaction'.

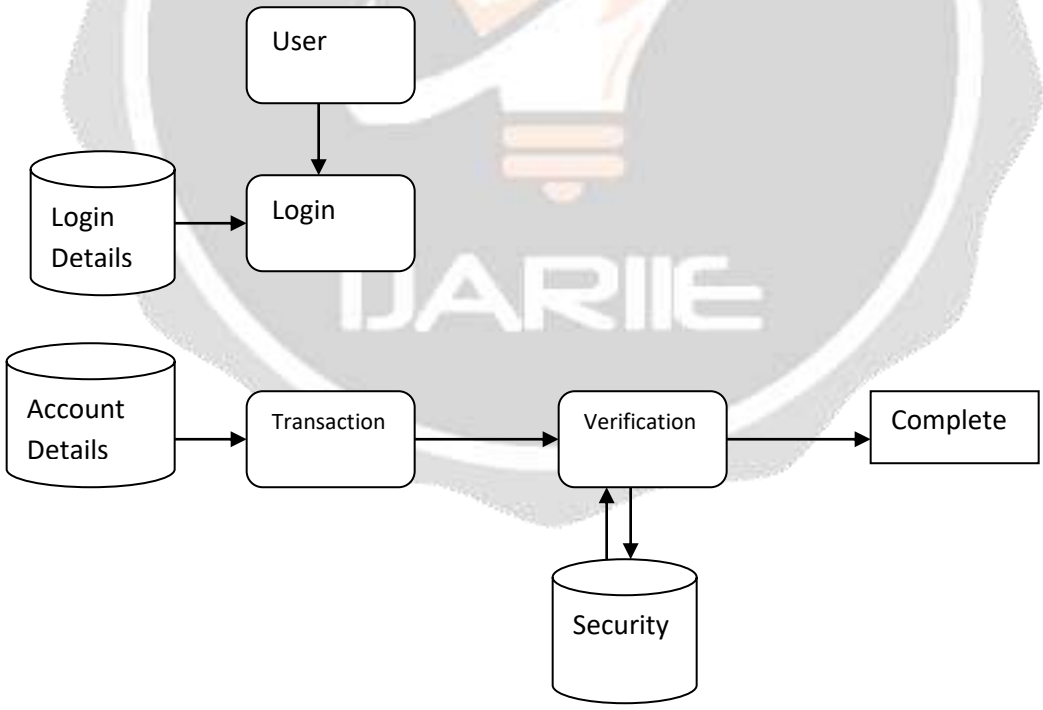

Fig-1 System Architecture

3. OVERVIEW OF SYSTEM

Following use-case diagram shows the overview of the Credit Card Fraud Detection Using Clustering Based Approach. The user actor can have the New card, login credentials, User can store information and Verification etc.

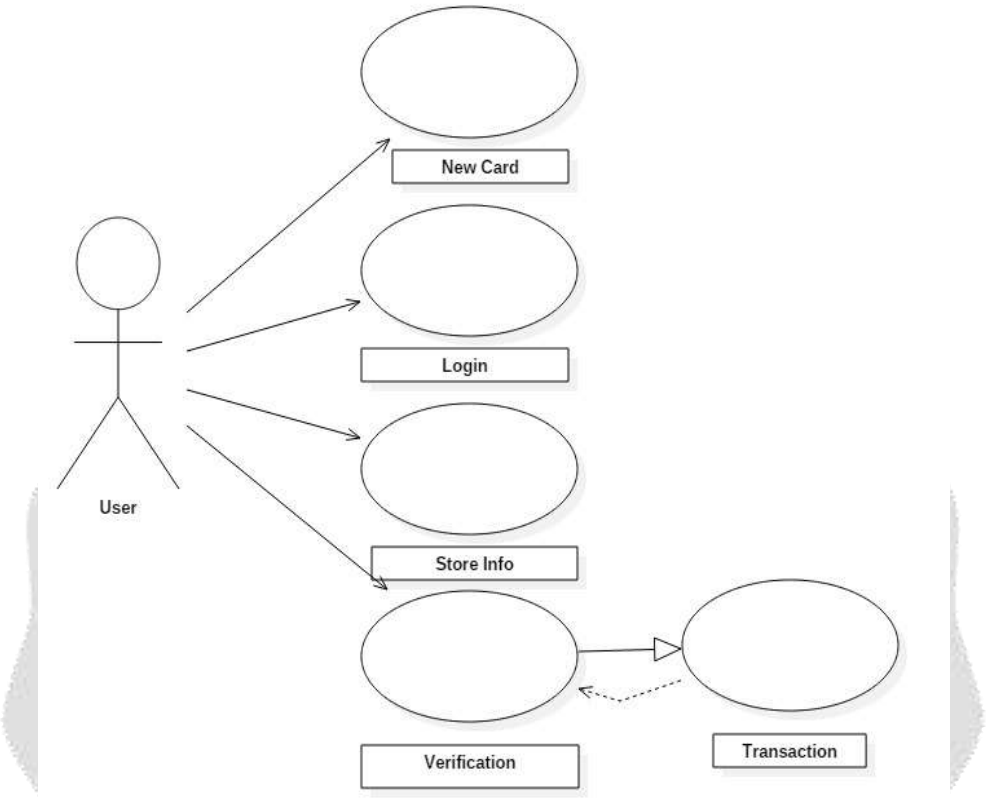

Fig-2 Use case Diagram

4. CONCLUSIONS

The paper aims to discover outliers which are vulgarly outliers different form or unsuitable with the remaining dataset. The usage of credit card is for both online and regular purchase and fraud join with it also increase. Criminal build it as business. The main thing is that technique for hindering credit card fraud is also improving many folds with passage of time. Minimizing price of computing assist in minimizing introducing composite system, which can analyze outlier transaction in a manner of fraction of second .It is similarly consequence to identify right place of transaction, which should be subject to reconsideration and as every transaction has not same amount of risk join with it .Finally optimally balanced price of fraud associate with it and issuing bank in combating fraud more competently.

5. ACKNOWLEDGEMENT

We take this opportunity to thank our guide Mr. R. M. Gawande and Project Coordinator Prof. S. Bhawasar and Head of the Department Dr. V. H. Patil for their valuable guidance and providing all necessary facilities, which were indispensable in the completion of the survey. We are also thankful to all the staff member of the Computer department of Matoshree Collage of Engineering and Research Center, Nashik for their valuable time, guideline, suggestions and comments.

6. REFERENCES

[1] Surbhi Agarwal,Santosh Upadhayay"A Fast Fraud Detection Approach Using Clustering Based Method", Journal of Basic and Applied Engineering Research, Volume 1, Number 10; October, 2014.

[2] A. Shrivastava, A. Kundu, S. Sural and A. Mujamdar, "Credit Card Fraud Detection using Hidden Markov Model",IEEE Transaction On Dependable And Secure Computing, vol. 5, No. 1, January-March 2008.

[3] Vaishali, "Fraud Detection in Credit Card by Clustering Approach", Iternational Journal of Computer Applications, vol. 98, No. 3, July 2014.

[4] A.Dharmarajan, T. Velmurugan, "Applications of partition based clustering algorithms: A survey", IEEE International Conference on Computational Intelligence and Computing Research 2013.

[5] Mohammed Al-Maolegi, B. A.: An improved apriori algorithm for association rules of mining. International Journal on Natural Language Computing, 3(1), pp. 942–946. doi:10.1109/ITIME.2009.5236211, (2014)

[6] François, D.: Binary classification performances measure cheat sheet. Journal of Machine Learning Research, 7, pp. 1–30, (2006).