Credit Card Fraud Detection Using Machine Learning

Sk.Anjaneyulu Babu[1], Devarampati Sisindri[2]

[1]Associate Professor, Department of MCA at QISCET (Autonomous), Vengamukkapalem, Prakasam (DT)
[2] PG Scholar in the Department of MCA at QISCET (Autonomous), Vengamukkapalem, Prakasam (DT)

ABSTRACT_

In this paper the main focal point on savings credit card fraud identification in real world. Here the savings credit card fraud detection is primarily based on transactions. Generally savings card fraud things to do can take place in each on line and offline. But in present day world on-line fraud transaction things to do are growing day by way of day. So in order to locate the on line fraud transactions quite a number techniques have been used in current gadget.In this paper we have proposed different machine learning algorithms or finding the fraudulent transactions and the accuracy of those transactions.

1. INTRODUCTION

There are quite a number fraudulent things to do detection strategies has applied in credit score card transactions have been stored in researcher minds to strategies to strengthen fashions primarily based on synthetic talent , records mining, fuzzy good judgment and computer learning. Credit card fraud detection is drastically difficult, however additionally famous trouble to solve. In our proposed device we constructed the savings card fraud detection the use of Machine learning. With the development of computing device gaining knowledge of techniques. Machine getting to know has been recognized as a profitable measure for fraud detection. A giant quantity of information is transferred for the duration of on-line transaction processes, ensuing in a binary result: true or fraudulent. Within the pattern fraudulent datasets, facets are constructed. These are information factors particularly the age and fee of the client account, as properly as the starting place of the savings card. There are heaps of elements and every contributes, to various extents, toward the fraud probability. Note, the degree in which every characteristic contributes to the fraud rating is generated via the synthetic talent of the computing device which is pushed by means of the education set, however is no longer decided by way of a fraud analyst. So, in regards to the card fraud, if the use of playing cards to commit fraud is tested to be high, the fraud

weighting of a transaction that makes use of a savings card will be equally so. However, if this had been to shrink, the contribution degree would parallel. Simply make, these fashions self-learn except express programming such as with guide review. Credit card fraud detection the usage of Machine studying is finished by means of deploying the classification and regression algorithms.

2.LITERATURE SURVEY

[1] The Use of Predictive Analytics Technology to Detect Credit Card Fraud in Canada. "KosemaniTemitayo Hafiz, Dr. Shaun Aghili, Dr. PavolZavarsky."

This look up paper focuses on the introduction of a scorecard from relevant contrast criteria, features, and talents of predictive analytics supplier choices in modern times being used to word credit score score card fraud. The scorecard gives a side-byside distinction of 5 financial savings card predictive analytics supplier choices adopted in Canada. From the ensuing look up findings, a checklist of credit card fraud PAT vendor reply challenges, risks, and obstacles was once as soon as outlined.

[2] BLAST-SSAHA Hybridization for Credit Card Fraud Detection. "AmlanKundu, SuvasiniPanigrahi, ShamikSural, Senior Member, IEEE, and Arun K. Majumdar"

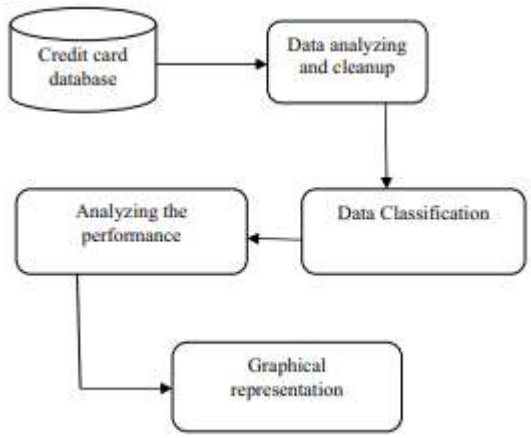
This paper recommend to use two-stage sequence alignment in which a profile Analyser (PA) first determines the similarity of an incoming sequence of transactions on a given financial savings card with the proper cardholder's preceding spending sequences. The exclusive transactions traced by means of skill of the profile analyzer are subsequent surpassed on to a deviation analyzer (DA) for potential alignment with preceding fraudulent behavior. The ultimate resolution about the nature of a transaction is taken on the basis of the observations by using the usage of these two analysers. In order to gain on line response time for every PA and DA, we advocate a new method for combining two sequence alignment algorithms BLAST and SSAHA.

[3] Fraudulent Detection in Credit Card System Using SVM & Decision Tree. "Vijayshree B. Nipane, Poonam S. Kalinge, DipaliVidhate, Kunal War, Bhagyashree P. Deshpande". As the field of electronic commerce gets better, fraud is

spreading all over the world and costing a lot of money. In the world we live in now, credit card fraud is the main cause of financial losses. It affects both businesses and individual clients. Methods like the decision tree, genetic algorithm, meta-learning strategy, neural network, and HMM are shown to find credit card fraud. Support Vector Machine (SVM) and decision tree, which are both parts of artificial intelligence, are used to solve the problem in the contemplate system for detecting fraud. By using this hybrid method, financial losses can be cut down to a greater extent.

3.PROPOSED WORK

We think that a Machine Learning model should be used to find credit card fraud in online financial transactions that use credit cards. Because there is so much data and the process is hard to understand, it is not possible to look for fake transactions by hand. Machine Learning could make it possible, though, if enough useful information is given. This idea will be looked into as part of the project. supervised learning algorithms like Random forest can be used to figure out which credit card transactions are real and which are fake. money.



3.1 MODULES

1: Exploratory Data Analysis: In this module we will first collect all the credit card dataset and store it in a database. Then we will perform some descriptive analysis about the dataset.

2: Data Cleaning In the next step, after analyzing the dataset then we have to clean the data. In this cleaning process all the duplicate values and null values that are present in the dataset will be removed and a new dataset will be obtained.

3: Preprocessing of dataset In this module, the cleaned dataset will be preprocessed by dividing it into groups based on the amount and time of each transaction.

4 : Dataset Partition In this module first the dataset will be divided into two partitions as trained dataset and testing dataset. After the data partitions the Random Forest Algorithm is applied. After applying Different machine learning algorithmsfinally a confusion matrix is obtained.

5:Evaluation Now the resultant data obtained in the form of confusion matrix can be evaluated by using graphical representation which gives better accuracy

4.ALGORTITHMS

4.1 Random Forest Algorithm

Random forests is an algorithm for learning with help. Both classification and regression can be done with it. It is also the most flexible and easiest algorithm to use. There are trees in a wooded area. Some people say that a wooded area is stronger the more trees it has. Random forests make decision trees based on random samples of data, get predictions from each tree, and choose the best answer based on how well the trees vote. It also gives an extremely useful sign of how important the characteristic is. Python's built-in SKLEARN has a guide for CART with all types of wood and a random wooded area classifier.

Random forests can be used for many things, such as making suggestions, sorting photos, and choosing functions. It can be used to sort mortgage applicants by how likely they are to pay back their loans, find out about fake activities, and predict diseases. It is the basis for the Boruta algorithm, which picks out the most important parts of a dataset..

5. RESULTS AND DISCUSSIONS

Fig 1: In above screen we can see Random Forest generate 99.78% percent accuracy while building model on train and test data. Now click on 'Detect Fraud From Test Data' button to upload test data and to predict whether test data contains normal or fraud transaction

Fig 2:In above screen beside each test data application will display output as whether transaction contains cleaned or fraud signatures. Now click on 'Clean & Fraud Transaction Detection Graph' button to see total test transaction with clean and fraud signature in graphical format. See below screen

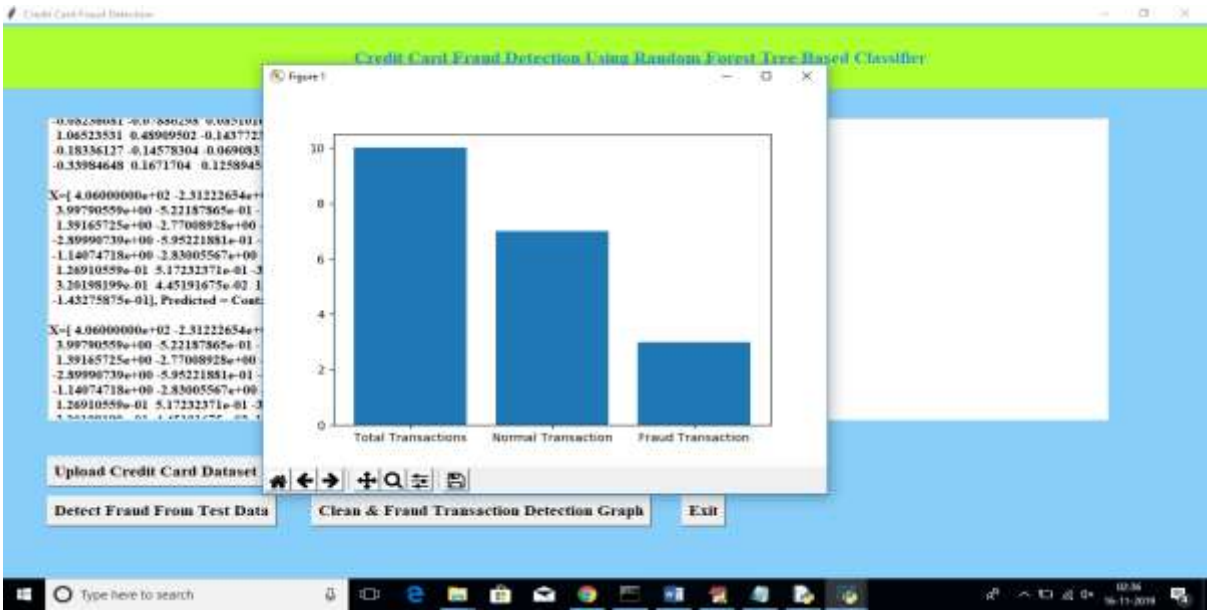


Fig 3:In above graph we can see total test data and number of normal and fraud transaction detected. In above graph x-axis represents type and y-axis represents count of clean and fraud transaction

6.CONCLUSION

In this project we have applied Diff machine learning algorithms for detecting credit cards frauds.as per our analysis each algorithms working fine but accuracy wise there is slight variation.The decision tree will provide better performance with many training data, but speed during testing and application will still suffer. Usage of more pre-processing techniques would also assist. Our future work will try to represent this into a software application and provide a solution for credit card fraud using the new technologies like Machine Learning, Artificial Intelligence and Deep Learning.

REFERENCES

[1] Sudhamathy G: Credit Risk Analysis and Prediction Modelling of Bank Loans Using R, vol. 8, no-5, pp. 1954-1966.

[2] LI Changjian, HU Peng: Credit Risk Assessment for ural Credit Cooperatives based on Improved Neural Network, International Conference on Smart Grid and Electrical Automation vol. 60, no. - 3, pp 227-230, 2017.

[3] Wei Sun, Chen-Guang Yang, Jian-Xun Qi: Credit Risk Assessment in Commercial Banks Based On Support Vector Machines, vol.6, pp 2430-2433, 2006.

[4] AmlanKundu, SuvasiniPanigrahi, ShamikSural, Senior Member, IEEE,"BLAST-SSAHA Hybridization for Credit Card Fraud Detection", vol. 6, no. 4 pp. 309-315, 2009.

[5] Y. Sahin and E. Duman, "Detecting Credit Card Fraud by Decision Trees and Support Vector Machines, Proceedings of International Multi Conference of Engineers and Computer Scientists, vol. I, 2011.

[6] Sitarampatel, SunitaGond, "Supervised Machine (SVM) Learning for Credit Card Fraud Detection, International of engineering trends and technology, vol. 8, no. -3, pp. 137- 140, 2014.

[7] SnehalPatil, HarshadaSomavanshi, JyotiGaikwad, AmrutaDeshmane, RinkuBadgujar," Credit Card Fraud Detection Using Decision Tree Induction Algorithm, International Journal of Computer Science and Mobile Computing, Vol.4 Issue.4, April- 2015, pg. 92-95

[8] Dahee Choi and Kyungho Lee, "Machine Learning based Approach to Financial Fraud Detection Process in Mobile Payment System", vol. 5, no. - 4, December 2017, p. 12-24.

[9] Golden.,& Walker, (1993). Corporate Governance in the financial Institutions. Journal of Business Finance and Accounting,16(3), 385-398 Inkumbi, M. (2009).

[10] Nagarajan, G. (2001). Can Microfinance Meet the Poor's Needs in Times of Natural Disaster? Microenterprise Best Practices (Bethesda: Development Alternatives, Inc).

[11] Craig, C., & Dan, C. (2001). Care Microfinance Risk Management Handbook. Available online: http://www.adb.org/doc./books/interest. Crosbie , (2003) Journal of Planktonic Research, plankt.oxfordjournals.org/content

[12] Ekhator, V.E. (2002). Role of Small and Medium Scale Enterprises in the Nigerian Economy. Paper Presented to the Department of Public Administration, University of Abuja.

[13] Lang. W. &Jagtiani, J. A. (2010). The Mortgage and Financial Crises: The Role of Credit Risk Management and Corporate Governance, Atlantic Economic Journal, 38, 295- 316

[14] Mohammed, A, (2002). The Dynamism of the banking industry: Research Report Mordi, F. (2002). "Why Government Must Protect Small and Medium Scale Enterprises". Financial Standard, August 19.

[15] Longstaff, F., & Schwartz, E. (1995). A simple approach to valuing risky fixed and floating rate debt. Journal of Finance, 50, 789–819

AUTHOR PROFILE

SK. ANJANEYULUBABU, Associate Professor in the department of MCA at QIS College Engineeringand Technology. He is having over 20 Research publications. His area of interest is Machine learning &Data mining.

D. SISINDRI, PG Scholar in the department of MCA, QIS College of engineering and Technology(Autonomous), Vengamukkapalem, Prakasam (DT) His areas of Interests are Networking & Cloud Computing.