Utilitas Mathematica

ISSN 0315-3681 Volume 122, 2025

Credit Card Fraud Detection

[1] Pabba Harish Goud

PG Scholar, Department of Computer Science & Engineering, Siddhartha Institute of Technology and Sciences, Hyderabad, India.

[2] B. Ravinder Reddy

Assistant professor, Department of Computer Science & Engineering, Siddhartha Institute of Technology and Sciences,

ABSTRACT

In our project, mainly focussed on credit card fraud detection for in real world. Initially I will collect the credit card datasets for trained dataset. Then will provide the user credit card queries for testing data set. After classification process of random forest algorithm using to the already analysing data set and user provide current dataset. Finally optimizing the accuracy of the result data. Then will apply the processing of some of the attributes provided can find affected fraud detection in viewing the graphical model visualization. The performance of the techniques is evaluated based on accuracy, sensitivity, and specificity, precision. The results indicate about the optimal accuracy for Decision tree are 98.6% respectively.

1. INTRODUCTION

Now a day the usage of credit cards has dramatically increased. As credit card becomes the most popular mode of payment for both online as well as regular purchase, cases of fraud associated with it are also rising. In this paper, we model the sequence of operations in credit card transaction processing using a Decision tree and Deep Neural Network show how it can be used for the detection of frauds. An both algorithms is initially trained with the normal behaviour of a cardholder. If an incoming credit card transaction is not accepted by the trained with sufficiently high probability, it is considered to be fraudulent. At the same time, we try to ensure that genuine transactions. We present detailed experimental results to show the effectiveness of our approach and compare it with other techniques available in the literature.

2. LITERATURE SURVEY

1) **The Use of Predictive Analytics Technology to Detect Credit Card Fraud in Canada (Kosemani Temitayo Hafiz, Dr. Shaun Aghili, Dr. Pavol Zavarsky)**

This research paper focuses on the creation of a scorecard from relevant evaluation

criteria, features, and capabilities of predictive analytics vendor solutions currently being used to detect credit card fraud. The scorecard provides a side-by-side comparison of five credit card predictive analytics vendor solutions adopted in Canada. From the ensuing research findings, a list of credit card fraud PAT vendor solution challenges, risks, and limitations was outlined.

2) BLAST-SSAHA Hybridization for Credit Card Fraud Detection

(Amlan Kundu, Suvasini Panigrahi, Shamik Sural, Senior Member, IEEE, and Arun K. Majumdar)

In this paper, we propose 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 credit card with the genuine cardholder's past spending sequences. The unusual transactions traced by the profile analyser are next passed on to a deviation analyser (DA) for possible alignment with past fraudulent behaviour. The final decision about the nature of a transaction is taken on the basis of the observations by these two analysers. In order to achieve online response time for both PA and DA, we suggest a new approach for combining two sequence alignment algorithms BLAST and SSAHA.

3. SYSTEM ANALYSIS

3.1 EXISTING SYSTEM:

In existing System, a research about a case study involving credit card fraud detection, where data normalization is applied before Cluster Analysis and with results obtained from the use of Cluster Analysis and Artificial Neural Networks on fraud detection has shown that by clustering attributes neuronal inputs can be minimized. And promising results can be obtained by using normalized data and data should be MLP trained. This research was based on unsupervised learning. Significance of this paper was to find new methods for fraud detection and to increase the accuracy of results.

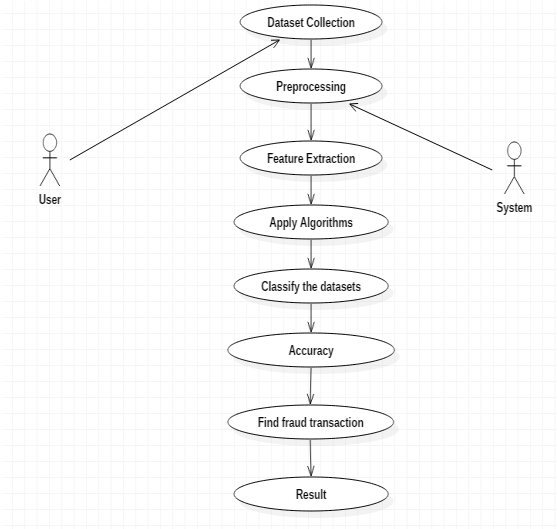
3.2 PROPOSED SYSTEM:

In proposed System, we are applying random forest algorithm for classify the credit card dataset. Decision tree is an algorithm for classification and regression. Summarily, it is a collection of decision tree classifiers. Decision tree has advantage over decision tree as it corrects the habit of over fitting to their training set. A subset of the training set is sampled randomly so that to train each individual tree and then a decision tree is built, each node then splits

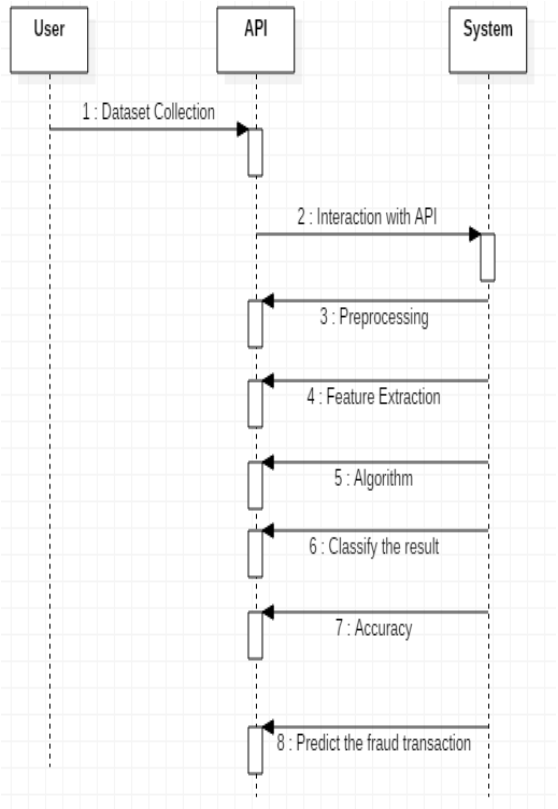
on a feature selected from a random subset of the full feature set.Use different methods to collect the customer data instead of physical forms

4. SYSTEM DESIGN

4.1 USE CASE DIAGRAM



4.2 SEQUENCE DIAGRAM



5. IMPLEMENTATION

5.1 DATA COLLECTION

Data used in this paper is a set of dataset. This step is concerned with selecting the subset of all available data that you will be working with. ML problems start with data preferably, lots of data (examples or observations) for which you already know the target answer. Data for which you already know the target answer is called labelled data.

5.2 DATA PRE-PROCESSING

Organize your selected data by formatting, cleaning and sampling from it.

Three common data pre-processing steps are:

Formatting: The data you have selected may not be in a format that is suitable for you to work with.

Cleaning: Cleaning data is the removal or fixing of missing data. There may be data instances that are incomplete and do not carry the data you believe you need to address the problem.

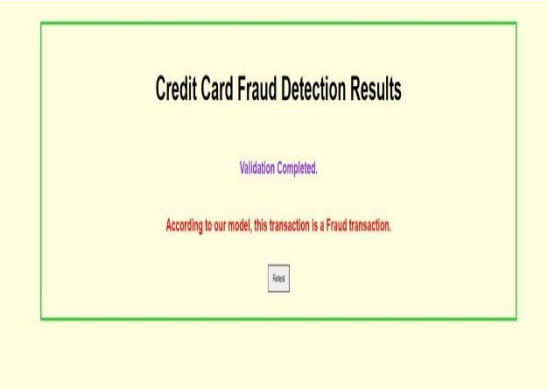
5.3 FEATURE EXTRATION

Next thing is to do Feature extraction is an attribute reduction process. Unlike feature selection, which ranks the existing attributes according to their predictive significance, feature extraction actually transforms the attributes. The transformed attributes, or features, are linear combinations of the original attributes.

5.4 EVALUATION MODEL

Model Evaluation is an integral part of the model development process. It helps to find the best model that represents our data and how well the chosen model will work in the future. Accuracy is defined as the percentage of correct predictions for the test data. It can be calculated easily by dividing the number of correct predictions by the number of total predictions.

6. OUTPUT SCREENS



7. CONCLUSION

The proposed paper evaluate that the Decision tree and support vector machine algorithm will perform better with a larger number of training data comparing to Adaboost classifier , but speed during testing and application will suffer. Application of more pre-processing techniques would also help. The SVM

algorithm still suffers from the imbalanced dataset problem and requires more pre-processing to give better results at the results shown by SVM is great but it could have been better if more pre-processing have been done on the data.so, in proposed work we balanced the imbalanced data with up-sampling technique during pre-processing .

8. REFERENCES

I. P. Richhariya and P. K. Singh, "Evaluating and emerging payment card fraud challenges and resolution," International Journal of Computer Applications, vol. 107, no. 14, pp. 5 – 10, 2014.

II. S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data mining for credit card fraud: A comparative study," Decision Support Systems, vol. 50, no. 3, pp. 602–613, 2011.

III. A.DalPozzolo,O.Caelen,Y.-A.LeBorgne,S.Waterschoot,andG. Bontempi, "Learned lessons in credit card fraud detection from a practitioner perspective," Expert systems with applications, vol. 41, no. 10, pp. 4915– 4928, 2014.

IV. C. Phua, D. Alahakoon, and V. Lee, "Minority report in fraud detection: classification of skewed data," ACM SIGKDD explorations newsletter, vol. 6, no. 1, pp. 50–59, 2004.

V. Z.-H. Zhou and X.-Y. Liu, "Training cost-sensitive neural networks with methods addressing the class imbalance problem," IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 1, pp. 63–77, 2006.