CREDIT CARD FRAUD DETECTION SYSTEM

DR.P.SRIVANI

Professor & HOD CSE(Cybersecurity)

MALLA REDDY ENGINEERING COLLEGE FOR WOMEN

HYDERABAD

Pulla.srivani@gmail.com

1. B.PALLAVI B.TECH III YEAR MRECW HYDERABAD

2. A.RACHANA B.TECH III YEAR MRECW HYDERABAD

3. B.UDAYSRI B.TECH III YEAR MRECW HYDERABAD

ABSTRACT

One of the primary purposes of banking is to manage credit risk. Risk is categorised by banks based on their characteristics. Although various techniques have been developed, the problem has not yet been resolved. Data normalisation is currently used prior to cluster analysis, and the findings on fraud detection from these two techniques have indicated that clustering properties and neural inputs can be decreased. Finding an algorithm to lower the cost measure is the paper's main contribution. The algorithm employed was Minimum Bayesian-Risk, and the outcome was 23%. (MBR). Random Forest Algorithm is utilised for both classification and regression in the suggested system. Decision trees tend to overfit their training data sets, but random forest has an advantage over them in that it breaks this propensity..

KEYWORDS

Random Forest

Logistic Regression

Databases

Servers

Supervised Learning

INTRODUCTION

Every year, fraudulent credit card transactions cause billions of dollars in losses. Fraud is as old as humanity and can take on an infinite number of different forms. According to the 2017 PWC global economic crime survey, approximately 48% of organisations experienced economic crime [3]. As a result, there is unquestionably a need to address the issue of credit card fraud detection.

Credit card use is widespread in modern society, and credit card fraud has increased in recent years. Huge financial losses have resulted from fraud, affecting not only merchants and banks, but also individual credit users. Fraud can also harm a merchant's reputation and image, resulting in non-monetary losses that are difficult to quantify.

LITERATURE SURVEY

Title: Credit Card Fraud Detection Model Based on Distance

Summation Research W. Yu and N. Wang are the authors. Year: 2009 Description: Credit card fraud is on the rise in China, alongside the growth of credit cards and trade volume. How to improve the detection and prevention of credit card fraud has become the focus of bank risk management. This paper proposes a credit card fraud detection model based on distance sum based on the infrequency and irregularity of fraud in credit card transaction data, incorporating outlier mining into credit card fraud detection. Experiments show that this model can detect credit card fraud and is accurate.

EXISTING SYSTEM

Supervised Learning

In the current system, research on a case study involving credit card fraud detection has shown that by clustering attributes neuronal inputs can be minimised and using normalised data, promising results can be obtained. Data normalisation is applied before Naive Bayer's and Cluster Analysis, and with results obtained from the use of these methods on fraud detection. This study used unsupervised learning as its foundation. Detecting significance and improving the accuracy of outcomes. Personal information in the data is kept private and is based on actual transactional data from a major European corporation. An algorithm's accuracy hovers around 50%.

PROPOSED SYSTEM

The suggested system classifies the credit card data set using the Random Forest algorithm. An algorithmic software for classification and regression is called Random Forest. It is, in essence, a collection of decision tree classifiers. Because Random Forest corrects the behaviour of overfitting to their training set, it has an advantage over Decision Tree. To train each individual tree, a subset of the training set is randomly sampled, and after a decision tree has been built, each node is divided on a feature chosen at random from the entire feature set. Because each tree is trained independently of the others in Random Forest, training is incredibly quick even for big data sets with

numerous characteristics and data instances.

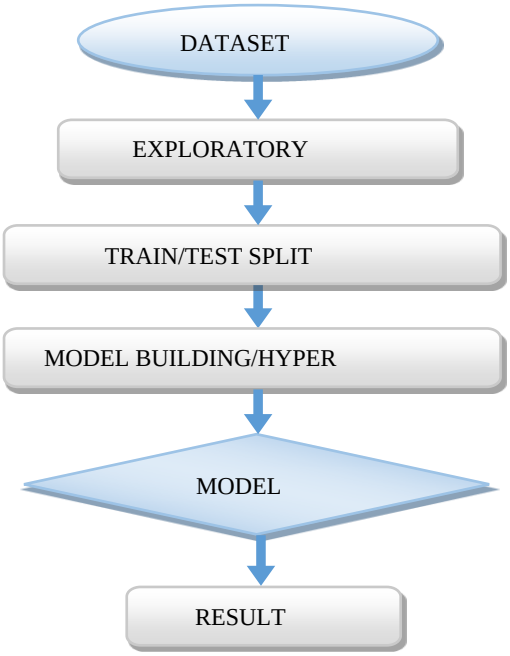
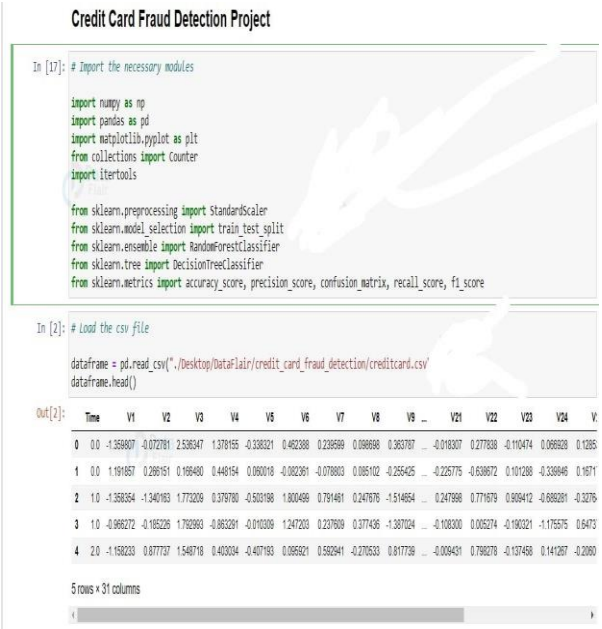


Fig: System Architecture

EXPERIMENTAL RESULTS



CONCLUSION

In this study, an unbalanced dataset was utilised to test the applicability of various supervised machine learning models for predicting the likelihood of a fraudulent transaction occurring. To arrive at a certain result, we used the criteria of sensitivity, precision, and time. Since accuracy does not provide a clear conclusion and is not sensitive to unbalanced data, it was not employed as a criteria. In this study, we examined the models of kNN, Naive Bayes, Decision Tree, Logistic Regression, and Random Forest. The result attained is 99.9%, and the Random Forest Algorithm will perform better

with more training data. The SVM method can replace Random Forest, however it still has the issue of an unbalanced data set.

REFERENCES

[1] W. Yu and N. Wang, "Research on Credit Card Fraud Detection Model Based on Distance Sum," *2009 International Joint Conference on Artificial Intelligence*, Hainan Island, 2009, pp. 353-356. [2] Vijayshree B. Nipane, Poonam S. Kalinge, Dipali Vidhate, Kunal War, Bhagyashree P. Deshpande, Fraudulent Detection in Credit Card System Using SVM & Decision Tree.

[3] Sitaram patel, Sunita Gond, Supervised Machine (SVM) Learning for Credit Card Fraud Detection.

[4] Y. Sahin and E. Duman, Detecting Credit Card Fraud by Decision Trees and Support Vector Machines.

[5] Snehal Patil, Harshada Somavanshi, Jyoti Gaikwad, Amruta Deshmane, Rinku Badgujar, Credit Card Fraud Detection Using Decision Tree Induction Algorithm.

[6] E. Aleskerov, B. Freisleben, and B. Rao, “CARDWATCH: A neural network based database mining system for credit card fraud detection,” in Proc. IEEE/IAFE Computat. Intell. Financial Eng., Mar. 1997, pp. 220–226.