CREDIT CARD FRAUD DETECTION USING DECISION TREE AND RANDOM FOREST ALGORITHMS

[1]Mr. Manas Kumar ,[2]K. Pravallika , [3]Rishik raj ,[4]J. Manikanta

[1]Assistant Professor, Department of Information Technology, CMR College of Engineering & Technology

[2, 3, 4] B-Tech, Department of Information Technology, CMR College of Engineering & Technology

Abstract:

In the world of finance, as the technology grown, new systems of business making came into picture. Credit card system is one among them. But because of lot of loop holes in this system, lot of problems are aroused in this system in the method of credit card scams. Due to this the industry and customers who are using credit cards are facing a huge loss. There is a deficiency of investigation lessons on examining practical credit card figures in arrears to privacy issues. In the manuscript an attempt has been made for finding the frauds in the credit card business by using the algorithms which adopted machine learning techniques. In this regard, two algorithms are used viz Fraud Detection in credit card using Decision Tree and Fraud Detection using Random Forest. The efficiency of the model can be decided by using some public data as sample. Then, an actual world credit card facts group from a financial institution is examined. Along with this, some clatter is supplemented to the data samples to auxiliary check the sturdiness of the systems. The significance of the methods used in the paper is the first method constructs a tree against the activities performed by the user and using this tree scams will be suspected. In the second method a user activity-based forest will have constructed and using this forest an attempt will be made in identifying the suspect. The investigational outcomes absolutely show that the mainstream elective technique attains decent precision degrees in sensing scam circumstances in credit cards.

INTRODUCTION:

Nowadays Credit card usage has been drastically increased across the world, now people believe in going cashless and are completely dependent on online transactions. The credit card has made the digital transaction easier and more accessible. A huge number of dollars of loss are caused every year by the criminal credit card transactions. Fraud is as old as mankind itself and can take an unlimited variety of different forms. The PwC global economic crime survey of 2017 suggests that approximately 48% of organizations experienced economic crime. Therefore, there is positively a necessity to unravel the matter of credit card fraud detection. Moreover, the growth of new technologies provides supplementary ways in which criminals may commit a scam. The use of credit cards is predominant in modern day society and credit card fraud has been kept on increasing in recent years. Huge Financial losses have been fraudulent effects on not only merchants and banks but also the individual person who are using the credits. Fraud may also affect the reputation and image of a merchant causing non-financial losses. Credit card scam finding is while a trade receipts steps to preclude whipped cash, merchandises, or amenities attained via an illegal credit card business. Credit card scam can occur together by the customer or by somebody else. To avoid happening such frauds, there are many techniques invented. If such frauds happen, then how to track the misused transactions are also improvised. There are number of novel and unique algorithms are proposed to provide the security to the digital data transactions from unauthorized access. But still, there are some drawbacks in one or the other way. This paper deals with methodologies in detection of credit card frauds. In this system, users would register on the portal and can take up the role of a buyeror seller accordingly. The seller needs to upload all requisite details whereas the buyer can then buy the lands on the portal that are verified by the smart contract. Further users can get deeds digitally which will be uploaded as a new block in the chain. In this way this proposed system does not involve any middleman and all transactions are directly dealt between the buyer and the seller. Transactions will be backed up in all legal servers of all the parties involved in a cryptographic format and the audit ability of transactions will be stronger now that they are associated with timestamps. Business runs on information. The faster it is received and the more accurate it is, the better.

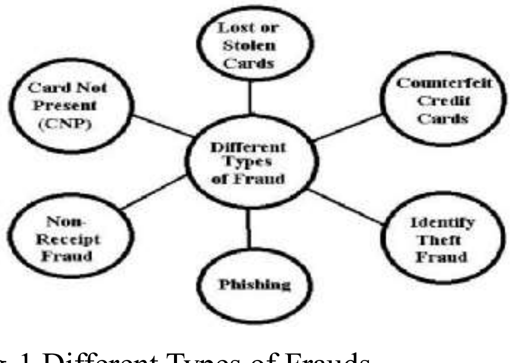


Fig-1 Different Types of Frauds

OBJECTIVE:

We propose a Machine learning model to detect fraudulent credit card activities in online financial transactions. Analyzing fake transactions manually is impracticable due to vast amounts of data and its complexity. However, adequately given informative features, could make it is possible using Machine Learning. This hypothesis will be explored in the project. To classify fraudulent and legitimate credit card transaction by supervised learning Algorithms such as Random Forest and Decision Tree.

PROPOSED MODEL

In proposed System, we are applying random forest algorithm for classification of the credit card dataset. Random Forest is an algorithm for classification and regression. Summarily, it is a collection of decision tree classifiers. Random forest has an 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. Even for large data sets with many features and data instances training is extremely fast in random forest and because each tree is trained independently of the others. The Random Forest algorithm has been found to provide a good estimate of the generalization error and to be resistant to over fitting. Random Forest Algorithm is used to detect the accuracy of the fraud in the transaction. Random choice forests are another name for random forests. These are a categorization, regression, and other tasks that use an ensemble learning strategy that involves teaching a greater number of decision trees and then determining the norm of the classifications (categorization) or the overall prediction (regression) of each tree. Random Forest is a supervised classification technique that uses ensemble learning. Ensemble model is a type of machine learning in which multiple versions of a same algorithm are combined to create a far more effective predictive model. Decision tree classifiers are used successfully in many diverse areas. Their most important feature is the capability of capturing descriptive decision-making knowledge from the supplied data. Decision tree can be generated from training sets. One of the major advantages of this algorithm is that it enforces the consideration of all the probable outcomes of decision and it keeps track of each path to a conclusion and creates a comprehensive analysis of the consequences

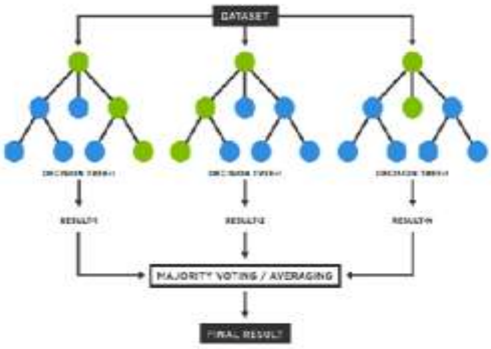


Fig-2 Random Forest

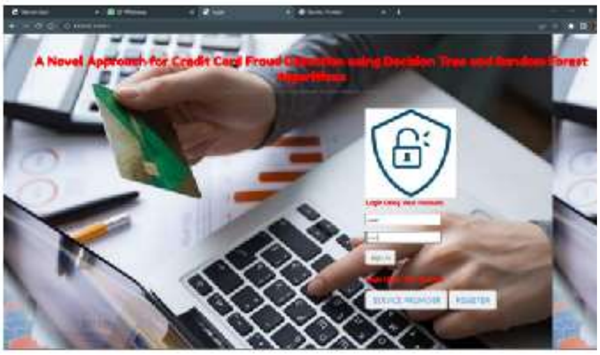
RESULTS AND DISCUSSIONS

In the above-mentioned existing systems as compared to our proposed system in which we calculate the true positive, true negative, false positive and false negative generated by a system or an algorithm and use these in quantitative measurements to evaluate and compare performance of different systems. True Positive (TP) is number of transactions that were fraudulent and were also classified as fraudulent by the system. True Negative (TN) is number of transactions that were legitimate and were also classified as legitimate. False Positive (FP) is number of transactions that were legitimate but were wrongly classified as fraudulent transactions. False Negative (FN) is number of transactions that were fraudulent but were wrongly classified as legitimate transactions by the system. 4.2 Data Collection and Performance Metrics Decision trees and random forest classifiers have been used in credit card fraud detection, with promising results. In supervised learning algorithms, decision trees are used for classification of datasets. Random forest classifiers have been found to have a high accuracy of 99.7% in detecting fraudulent transactions, as well as high precision, recall, and F1 score for class 1. These performance metrics demonstrate the effectiveness of decision tree and random forest classifiers in credit card fraud detection.

Screenshots

To run project, install python 3.7 and then install XAMPP and then create Pythonfolder inside system 'C' directory and then put 'aNovel_approach_forCredit_cardfraudDetection' folder inside that C:/Python folder and then start XAMPP server by clicking on 'START' option. Now type commands which gives you an URL http://127.0.0.1:8000/index.html and press enter key to get below home page

Fig-3 Command Prompt Commands



A Novel Approach for Credit Card Fraud Detection using Decision Tree and Random Forest Algorithms

Fig-4 Login Page

CONCLUSION

In conclusion, both random forest algorithm and decision tree have been successfully applied to credit card fraud detection with promising results. Random forest algorithm can provide a more accurate and robust detection performance by combining multiple decision trees, while decision tree can provide an interpretable and easy-tounderstand model for fraud analysts. Ultimately, the choice of algorithm depends on the specific needs and constraints of the application, such as the trade-off between accuracy and interpretability, computational resources, and the complexity of the dataset. Beside all these metric evaluation methods for the performance evaluation of the proposed algorithms, the algorithms can predict the fraud in credit card business up to some level, whereas the possibility of fraud occurrences in credit card business is through many intermediate channels. Regardless of the chosen algorithm, credit card fraud detection is an essential task for financial institutions and e-commerce businesses to protect their customers and assets from fraudulent activities

REFERENCES

[1] O. Adewumi and A. A. Akinyelu, “A survey of machine learning and natureinspired based credit card fraud detection techniques,” International Journal of System Assurance Engineering and Management, vol. 8, pp. 937–953, 2017J.

[2] Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73

[3] A. Srivastava, A. Kundu, S. Sural, A. Majumdar, “Credit card fraud detection using hidden Markov model,” IEEE Transactions on Dependable and Secure Computing, vol. 5, no. 1, pp. 37–48, 2008S.

[4] Bansal, J. C., Singh, P. K., Saraswat, M., Verma, A., Jadon, S. S., and Abraham, A. (2011). Inertia weight strategies in particle swarm optimization. In Nature and Biologically Inspired Computing (NaBIC), (Salamanca, Spain, October 19 - 21, 2011). IEEENaBIC'11,633--640.

[5] Bello-Orgaz, G., Jung, J. J., & Camacho, D. (2016). Social big data: Recent achievements and new challenges. Information Fusion. 28 (Mar. 2016), 45—59

[6].https://www.kaggle.com/code/hassana

min/credit-card-fraud-detection-usingrandom- forest/notebook

[7] https://www.tutorialspoint.com/what-is-a-neural-network-in-machine-learning

[8]https://www.simplilearn.com/tutorials/ machine-learning-tutorial/random-forestalgorithm [9] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C., “Data mining for credit card fraud: A comparative study,” Decision Support Systems, vol.50, no. 3, pp. 602–613, 2011.

[10] J. T. Quah, and M. Sriganesh, “Real-t ime credit card fraud detection using computational intelligence,” Expert Systems with Applications, vol. 35, no. 4, pp. 1721– 1732, 2008.

[11] Reddy, b. V. R., dasari, n., & venkateswararao, k. (2021). A steganography system with gausian markov random fields and error detection codes.

[12] Dr. S.Balamurugan, & Aurchana, Aurchana & Gurumoorthi Elangovan, Dr & Govindharaj, I. (2022). Augmentation of Decision Tree Characteristics for Agri-Food Supply Chain using Internet of Things.

[13]Muthubalaji, S., Divya Devi, B., Sangeetha, S., 2022, Performance Analysis of Rung Ladder-Structured Multilevel Inverter with PV Application, Cognitive Science and Technology, 10.1007/978-981-19-2350-0_11

[14] Sabitha, R., Shukla, A.P., Mehbodniya, A., Shakkeera, L., Reddy, P.C.S., 2022, A Fuzzy Trust Evaluation of Cloud Collaboration Outlier Detection in Wireless Sensor Networks, Ad-Hoc and Sensor Wireless Networks, 10.32908/ahswn.v53.8447

[15] Vemula, P., Dhar, R.S., 2022, Design of 8T SRAM using 14nm FINFET Technology [Konstrukcja 8T SRAM przy uÅ¼yciu technologii 14nm FINFET], Przeglad Elektrotechniczny, 10.15199/48.2022.10.07