Credit Card Fraud Detection using Machine Learning Algorithms

Mr. Aniket K. Shewade, Nilesh S. Bissa
Department Of Computer Science & Engineering
VidarbhInstitude of Technology of Engineering, Nagpur, India.
RashtrasantTukdojiMaharaj Nagpur University,
Nagpur 2019-2020

Abstract—Credit card payment has become very popular today. Most of people used online transaction (credit card or debit card). Credit card is an easiest way to pay directly through your bank account. But we know now a day's credit card fraud transaction. It is very important to prevent the account transaction from unauthorized user. This paper intends to illustrate the modelling of a data set using machine learning with Credit Card Fraud Detection In this paper we used three discrete algorithm ((BNN, Isolation forest, Random forest algorithm)) to detect fraud.The Credit Card Fraud Detection Problem includes modelling past credit card transactions with the dataof the ones that turned out to be fraud. Our objective to detect 100% of fraudulent transaction. In this process, we have focused on analyzing and preprocessing data sets as well as the deployment of multiple anomaly detection algorithms such as Local Outlier Factor and Isolation Forest algorithm on the PCA transformed Credit Card Transaction data.

Keywords— ***Machine Learning; Isolation forest, Building a random Forest Model ,BNN.***

1. Introduction

Credit card fraud is an inclusive term for fraud committed using a payment card, such as a credit card or debit card.The purpose may be to obtain goods or services, or to make payment to another account which is controlled by a criminal.The Payment Card industry security standard (PCI DSS) is the data security standard created to help businesses process card payments securely and reduce card fraud. Credit card fraud can be authorized, where the genuine customer themselves processes a payment to another account which is controlled by a criminal, or unauthorized, where the account holder does not provide authorization for the payment to proceed and the transaction is carried out by a third party.In 2018, unauthorized financial fraud losses across payment cards and remote banking totaled £844.8 million in the United Kingdom. Whereas banks and card companies prevented £1.66 billion in unauthorized fraud in 2018. That is the equivalent to £2 in every £3 of attempted fraud being stopped. Credit cards are more secure than ever, with regulators, card providers and banks taking considerable time and effort to collaborate with investigators worldwide to ensure fraudsters aren't successful. Cardholders' money is usually protected from scammers with regulations that make the card provider and bank accountable. The technology and security measures behind credit cards are becoming increasingly sophisticated making it harder for fraudsters to steal money.

Currently, credit-card companies attempt to predict the legitimacy of a purchase through the analyzing anomalies in various fields such as purchase location, transaction amount, and user purchase history. However, with the recent increases in cases of credit card fraud it is crucial for credit card companies to optimize their algorithmic solutions. This paper compares various machine learning and regression algorithmic models to explore which algorithm and combination of factors provides the most accurate method of classifying a credit-card transaction as fraudulent or non-fraudulent (normal).

2. LITERATURE SURVEY

1. In this paper, S.P. Manirajan describe Random forest algorithm applicable on Find fraud detection. Random forest has two types. They describe in detail and their accuracy 91.96% and 96.77% respectively. This paper summaries second type is better than the first type.

2. Suman Arora, in this paper, many supervised machine learning algorithms apply on 70% training and 30% testing dataset. Random forest, stacking classifier, XGB classifier, SVM, Decision tree and KNN algorithms compare each other i.e. 94.59%, 95.27%, 94.59%, 93.24%, 90.87%, 90.54% and 94.25% respectively. Summaries of this paper, SVM has the highest ranking with 0.5360 FPR, and stacking classifier has the lowest ranking with 0.0335.

3. KosemaniTemitayoHafiz, in this paper, they describe flow chart of fraud detection process. I.e. data Acquisition, data pre-processing, exploratory data analysis and methods or algorithms are in detail. Algorithms are K- nearest neighbor (KNN), random tree and Logistic regression accuracy are 96.91%, 94.32%, 57.73% and 98.24% respectively.

3. PROPOSED SYSTEM

1. The proposed model is introduced to overcome all the disadvantages that arises in the existing system.

2. This system will increase the accuracy of the classification results by classifying the data based on the attacks and others using naive-bayes classification algorithm.

3.It enhances the performance of the overall classification results.

4. METHODOLOGY

Datasets are an integral part in the field of machine learning. Gathering data is one of the hardest tasks, especially when it is related to the financial domain like credit card fraud.

1. This work focusing on an application which is used to detect the fraudulent credit card activities on internet transaction. In this peculiar type, the pattern of current fraudulent usage of the credit card has been analyzed with the previous transaction. By using the BNN in algorithm of machine learning algorithm.

2. In credit card fraud detection train an auto encoder neural network (BNN) (implemented in keras) in unsupervised or semi-supervised fashion for anomaly detection.

3. The train model will be evaluated on pre label an anomalymized data set.

i. Will be using:

ii. Tenser flow

iii. Keras

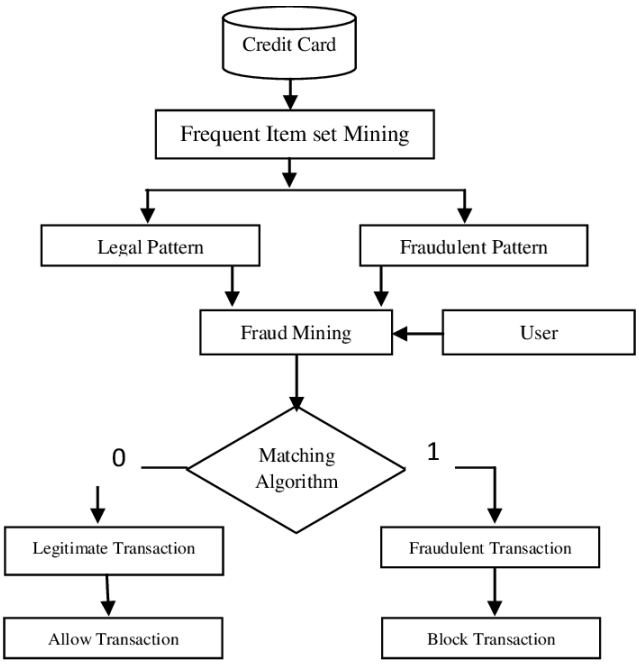

Fig.1: Credit Card Fraud Detection

5. CONCLUSION

Fraud detection is a complex issue that requires a substantial amount of planning before throwing machine learning algorithms at it. Nonetheless, it is also an application of data science and machine learning for the good, which makes sure that the customer's money is safe and not easily tampered with.

5. REFERENCES

1. Credit Card Fraud Detection Based on Transaction Behavior -by John Richard D. Kho, Larry A. Vea published by Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, and November 5-8, 2017

2. L.J.P. van der Maaten and G.E. Hinton, Visualizing High-Dimensional Data Using t-SNE (2014), Journal of Machine Learning Research

3. Machine Learning Group — ULB, Credit Card Fraud Detection (2018), Kaggle

4. Nathalie Japkowicz, Learning from Imbalanced Data Sets: A Comparison of Various Strategies (2000), AAAI Technical Report WS-00–05

5. R. M. jamailesmaily, "Intrusion detection system based on multilayer perceptron neural networks and decision tree," in International conference on Information and Knowledge Technology, 2015.

6. J. K. T. J. C. W. Siddhatha Bhattacharya, "Data Mining for credit card fraud: A comparative study," Elsevire, vol. 50, no. 3, pp. 602613, 2011.

7. RaghavendraPatidar and LokeshSharma International Journal of soft computing and engineering, vol. 1, no. NCAI2011, 2011.

8. S.P. Tanmaykumarbehera, "credit card fraud detection: a hybrid approach using fuzzy clustering and neural network," in international conference on advances in computing and communication Engineering, 2015.

9. N. W. Wen -Fang Yu, "Research on credit card fraud detection model based on distance sum," in International joint conference on artificial intelligence, Hainan Island,China, 2009.

10. S. k. A. K. M. Ayushiagarwal, "Credit card fraud detection: A case study," in IEEE, New Delhi, India, 2015. 7. K. T. B. V. Sam Maes, "Credit cards fraud detection using bayesian and neural networks," p. 7, August 2002.

11. P. K. D. K. R. D. A. A. ThurayaRazoogi, Credit card fraud detection using fuzzy logic and neural networks, Society for modelling and simulation International(SCS), 2016.

12. E. D. Y. Sahin, "Detecting credit card fraud by decision trees," in Proceedings of the international multiconference of engineers and computer science, Hong Kong, 2011.