

CREDIT CARD FRAUD DETECTION

A.DURGA DEVI¹, KELLA DIVYA SWATHI²

¹ Assistant Professor MCA, DEPT, Dantuluri Narayana Raju College, Bhimavaram, Andhra Pradesh

Email id:- adurgadevi760@gmail.com

²PG Student of MSc Computer Science, Dantuluri Narayana Raju College, Bhimavaram,

Andhra Pradesh

Email id:- swathikella2001@gmail.com

ABSTRACT

In our project, we mainly focused on credit card Error detection in real world. The motive is to find all the Error transactions that have taken place. In order to find whether a transaction is Error or not the model is trained from past credit card transactions that turned out to be Error. The models that we have used are Logistic Regression.

Billions of dollars of loss are caused every year by fraudulent credit card transactions. The design of efficient fraud detection algorithms is key for reducing these losses, and more and more algorithms rely on advanced machine learning techniques to assist fraud investigators. The design of fraud detection algorithms is however particularly challenging due to the non-stationary distribution of the data, the highly unbalanced classes distributions and the availability of few transactions labeled by fraud investigators.

1. INTRODUCTION

Nowadays the usage of credit cards has dramatically increased. As credit cards become the most popular mode of payment for both online as well as regular purchase, cases of Error associated with it are also rising. In this paper, we model the sequence of operations in credit card transaction processing using a Random Forest to show how it can be used for the detection of Errors. In 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 Error lent. 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.

Credit cards are used for purchasing goods and services with the help of virtual card and physical card where as virtual card for online transaction and physical card for offline transaction. In a physical-card based purchase, the cardholder presents his card physically to a merchant for making a payment. To carry out fraudulent transactions in this kind of purchase, an attacker has to steal the credit card. If the cardholder does not realize the loss of card, it can lead to a substantial financial loss to the credit card company. In online payment mode, attackers need only little information for doing fraudulent transaction (secure code, card number, expiration date etc.). In this purchase method, mainly transactions will be done through Internet or telephone. To commit fraud in these types of purchases, a fraudster simply needs to know the card details. Fraud is an offensive activity, carried out by an unauthorized person by cheating innocent. Credit card fraud involves stealing the essential credentials from the cardholder and using it unauthorized manner by the fraudsters either by using phone calls or SMS. This fraud in credit card may also happen using some software applications that are under the control of fraudsters. The credit card fraud detection takes place as: the user or the customer enters the necessary credentials in order to make any transaction using credit card and the transaction should get approved only upon being checked for ay fraud activity. For this

to happen, we first pass the transaction details to the verification module where, it is classified under fraud and non-fraud categories. Any transaction that is put under fraud category is rejected. Otherwise, the transaction gets approved.

2. LITERATURE SURVEY AND RELATED WORK

1. Fraudulent Detection in Credit Card System Using SVM & Decision Tree

With growing advancement in the electronic commerce field, fraud is spreading all over the world, causing major financial losses. In the current scenario, Major cause of financial losses is credit card fraud; it not only affects tradesperson but also individual clients. Decision tree, Genetic algorithm, Metalearning strategy, neural network, HMM are the presented methods used to detect credit card frauds. In contemplating system for fraudulent detection, artificial intelligence concept of Support Vector Machine (SVM) & decision tree is being used to solve the problem. Thus by the implementation of this hybrid approach, financial losses can be reduced to greater extent.

2. Machine Learning Based Approach to Financial Fraud Detection Process in Mobile Payment System

Mobile payment fraud is the unauthorized use of mobile transaction through identity theft or credit card stealing to fraudulently obtain money. Mobile payment fraud is a fast-growing issue through the emergence of smartphone and online transaction services. In the real world, a highly accurate process in mobile payment fraud detection is needed since financial fraud causes financial loss. Therefore, our approach proposed the overall process of detecting mobile payment fraud based on machine learning, supervised and unsupervised method to detect fraud and process large amounts of financial data. Moreover, our approach performed sampling process and feature selection process for fast processing with large volumes of transaction data and to achieve high accuracy in mobile payment detection. F-measure and ROC curve are used to validate our proposed model.

3. The Use of Predictive Analytics Technology to Detect Credit Card Fraud in Canada

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 were outlined.

4. BLAST-SSAHA Hybridization for Credit Card Fraud Detection.

This paper 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. EXISTING SYSTEM

Since the credit card fraud detection system is a highly researched field, there are many different algorithms and techniques for performing the credit card fraud detection system. One of the earliest system is CCFD system using markov model. credit card fraud detection(CCFD) is also proposed by using neural networks. The existing credit card fraud detection system using neural network follows the whale swarm optimization algorithm to obtain an inceptive value. It the uses BP network to rectify the values which are found error. All of these techniques has some serious disadvantages such as decreasing accuracy levels, lack of efficiency, sometimes classifying the normal transactions as fraud transactions and vice versa. These disadvantages are overcome in the upcoming proposed systems.

Disadvantages of Existing system

In existing System, a research about a case study involving credit card fraud www.jespublication.com 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. 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.

4. PROPOSED SYSTEM

The Credit card fraud detection system is initiated for detecting the fraud transactions from number of transactions made by the card holders. The transactions done by credit card holders are derived in the form of datasets.

Our goal is to implement machine learning model in order to classify, to the highest possible degree of accuracy, credit card fraud from a dataset gathered from Kaggle. After initial data exploration, we knew we would implement a logistic regression model for best accuracy reports.

Advantages

- The results obtained by the Logistic Regression Algorithm is best compared to any other Algorithms.
- The Accuracy obtained was almost equal to cent percent which proves using of Logistic algorithm gives best results.
- The plots that were plotted according to the proper data that is processed during the implementation
- The 'amount' feature is the transaction amount. Feature 'class' is the target class for the binary classification and it takes value 1 for positive case (fraud) and 0 for negative case (not fraud).
- Random forest ranks the importance of variables in a regression or classification problem in a natural way can be done by Random Forest.

5. METHODOLOGIES**MODULES****DATASET**

This paper utilizes the dataset provided by revolution analytics for the detection of the fraudulent credit card transaction from Kaggle. Dataset has 51149 legal transactions and 3312 fraudulent transactions. The dataset is divided as 60%, 20% and, 20% in the Train, Valid and Test set, respectively

DATA PREPROCESSING

For efficient implementation of the classification algorithm, data preprocessing is performed before feature selection. Under-sampling is performed to make the dataset balanced to avoid the biasing of the classification algorithm towards the majority class. Feature Selection is implemented on a balanced dataset.

FEATURE SELECTION

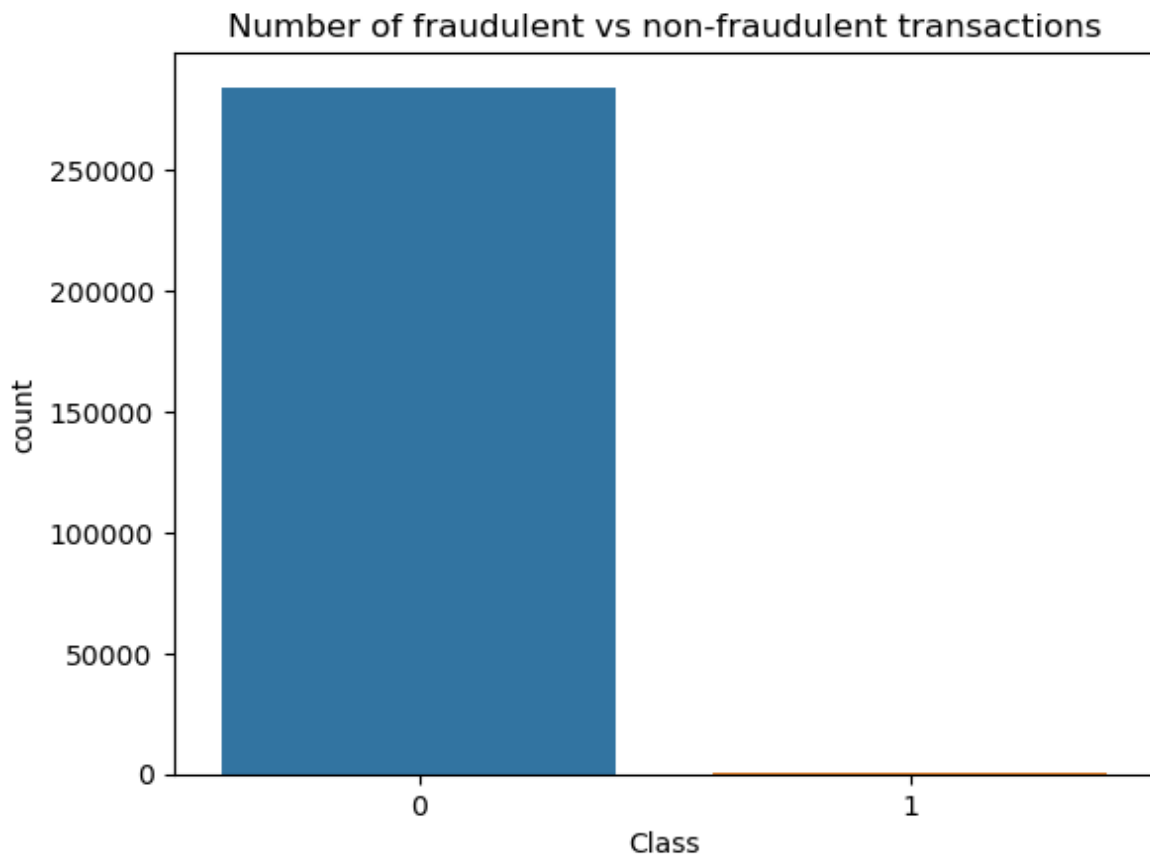
Feature selection methods are used to remove unnecessary, irrelevant, and redundant attributes from a dataset that do not contribute to the accuracy of a predictive model or which might reduce the accuracy of the model. In this paper seven feature selection techniques namely Select-K-best, Feature Importance, Extra tress classifier, Person's correlation, Mutual Information, Step forward selection and Recursive feature elimination are used.

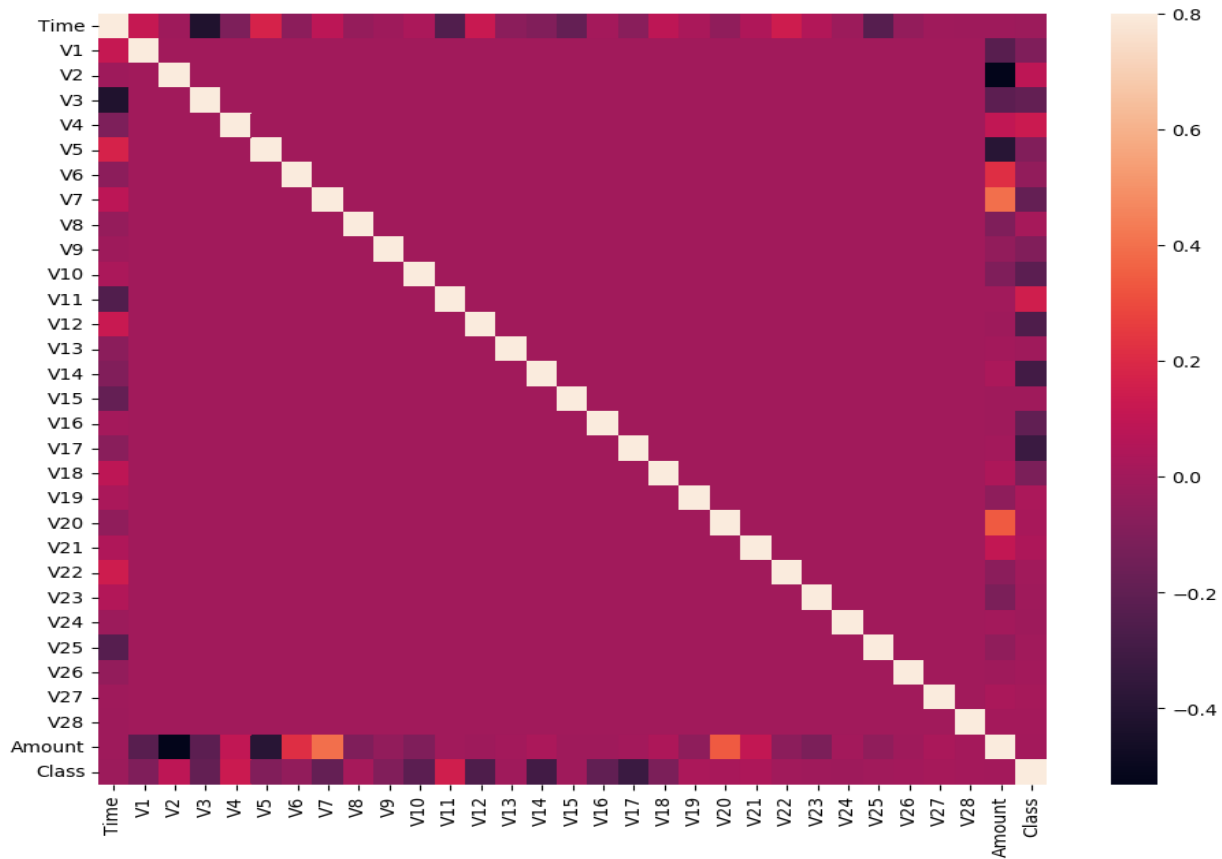
FEATURE IMPORTANCE

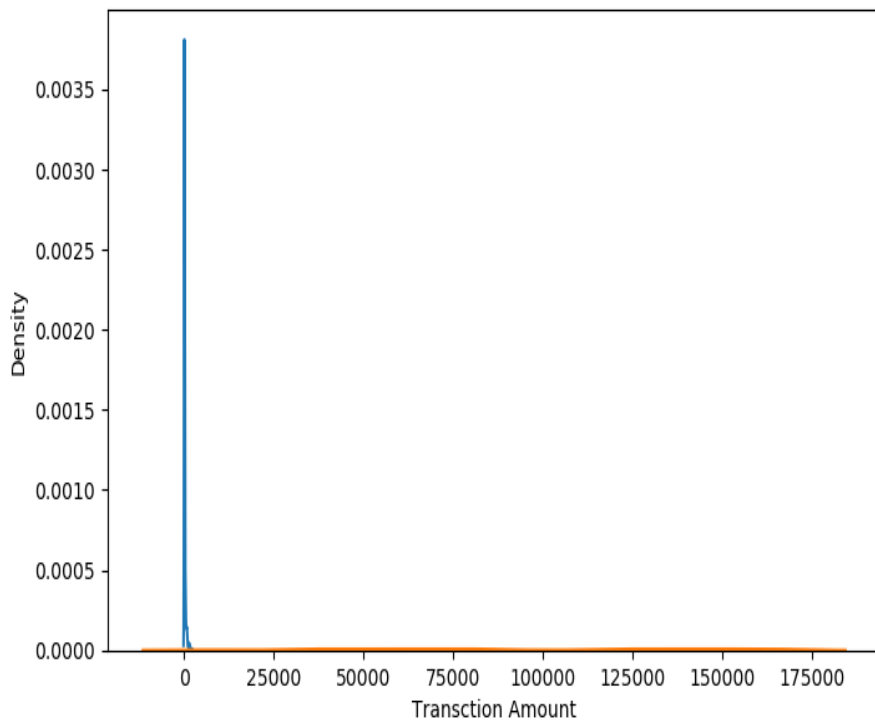
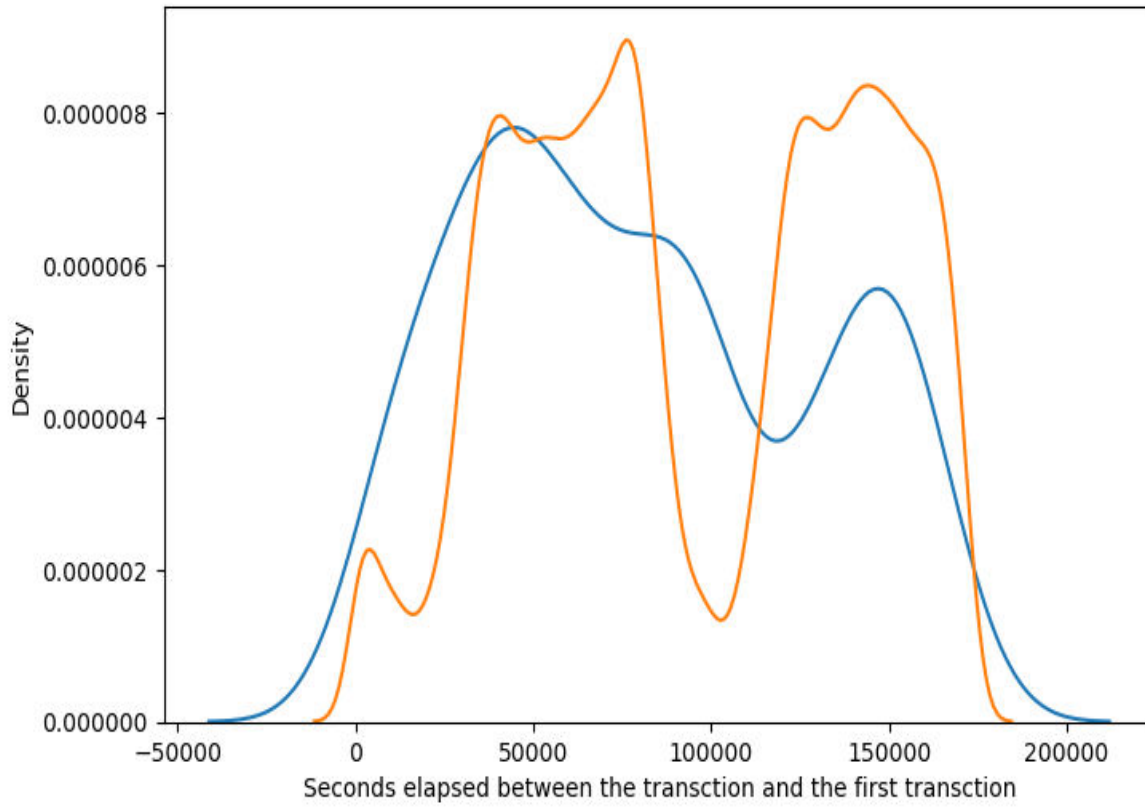
Feature importance is a class of techniques for assigning scores to input features to a predictive model that

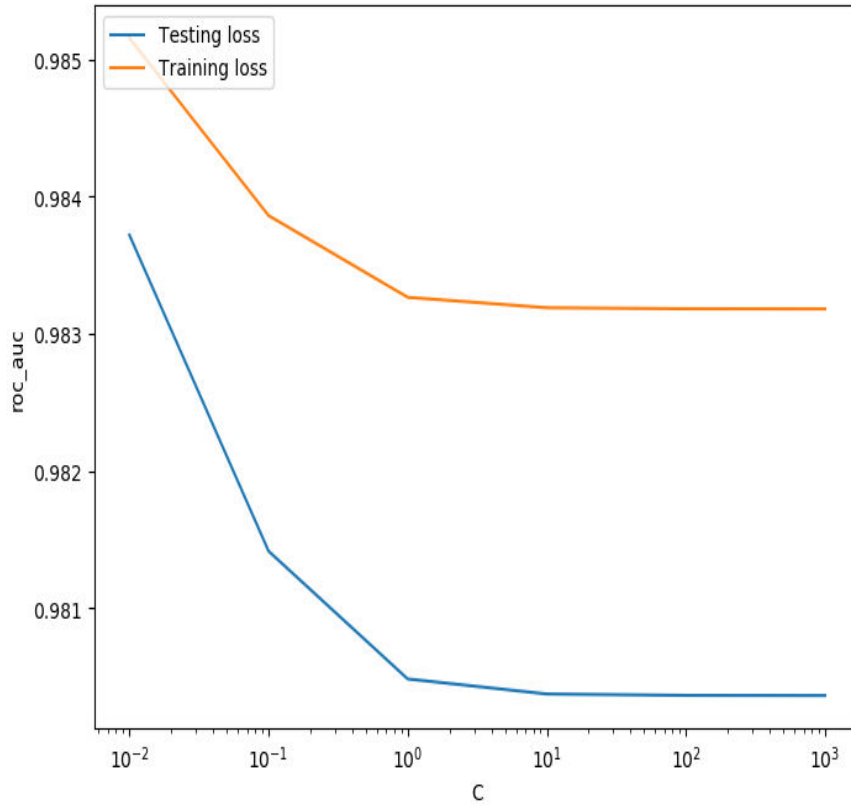
indicates the relative importance of each feature at the time of making a prediction. It reduces the number of input features. In this paper, feature importance is implemented using an extra tree classifier from the decision tree. Extra Trees is similar to Random Forest, it builds multiple trees and splits nodes using random subsets of features, but unlike Random Forest, Extra Tree samples without replacement and nodes are split on random.

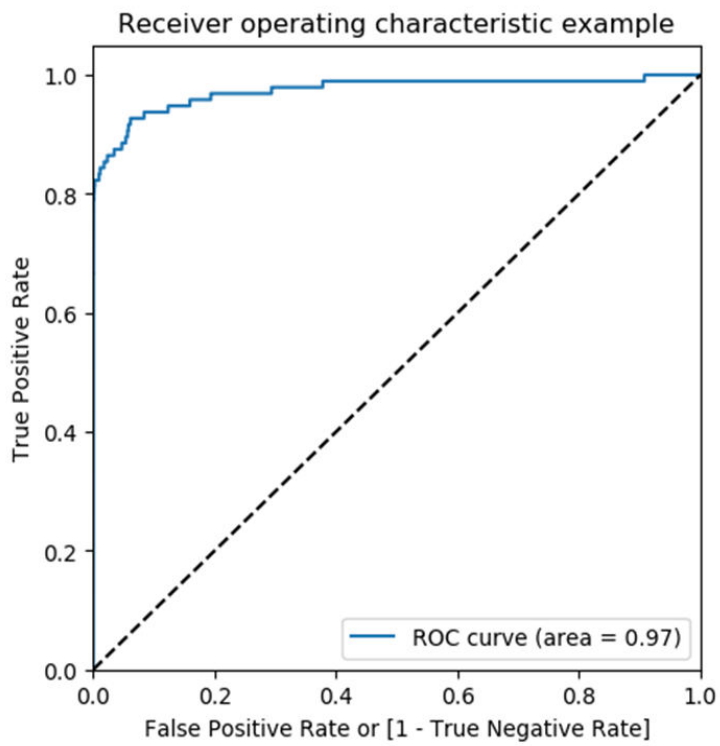
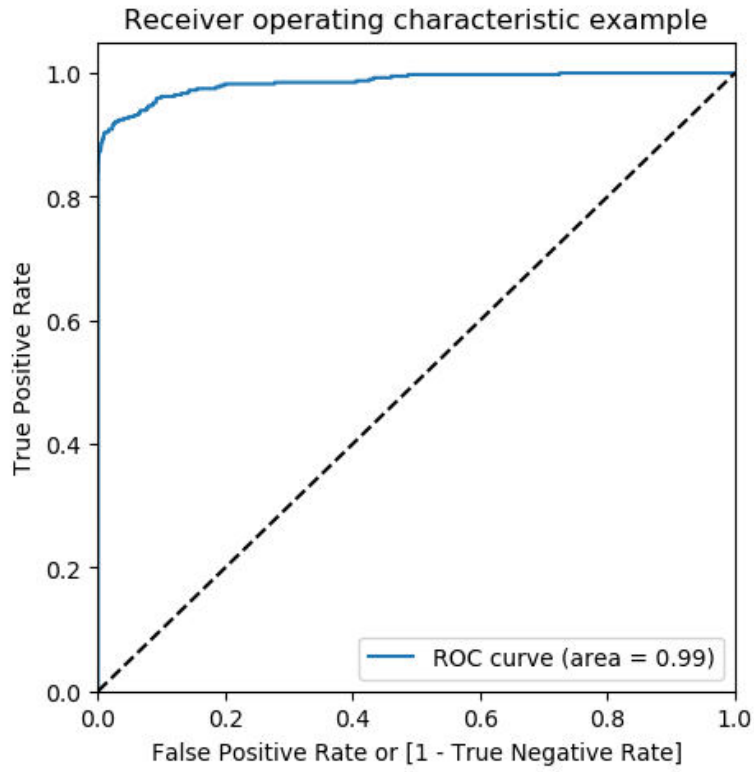
6. RESULTS AND DISCUSSION SCREEN SHOTS

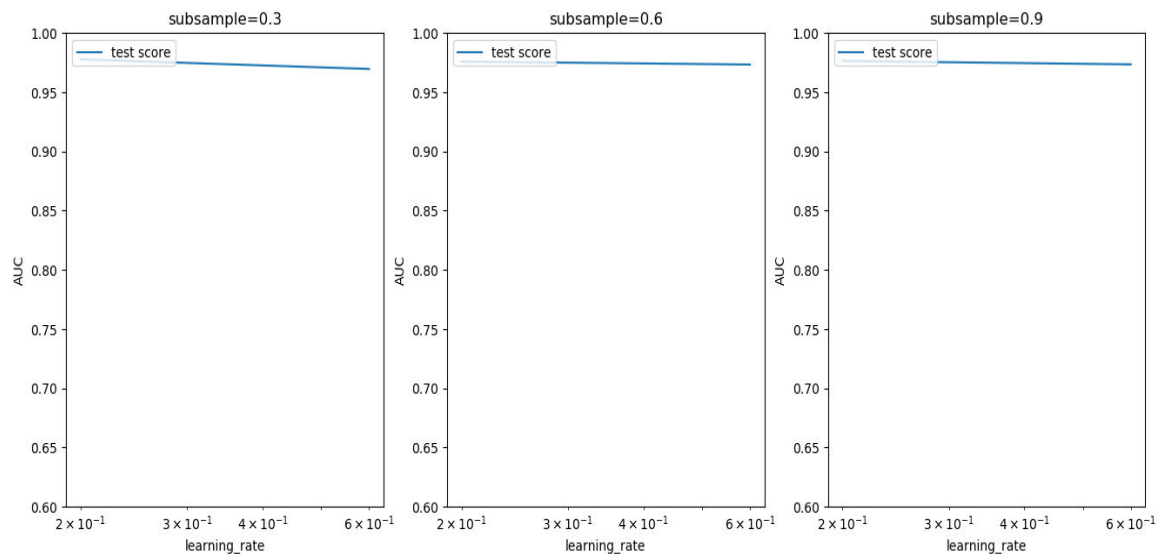












7. CONCLUSION AND FUTURE SCOPE

This machine learning fraud detection tutorial showed how to tackle the problem of credit card fraud detection using machine learning. It is fairly easy to come up with a simple model, implement it in Python and get great results for the Credit Card Fraud Detection task on Kaggle. Credit card fraud detection system using whale optimization algorithm and SMOTE (Synthetic minority optimization technique) aims in indentifying the fraud transactions occurring during the transactions made by the card holder. The system also aims to improve the convergence speed and solves the data imbalance.

In this research, we have proposed a method to detect the fraud in credit card transactions that is based on deep learning. We first compare it with machine learning algorithms such as k- Nearest Neighbor, Support vector machine etc. Finally we have used the neural network, even though tough to train the model which would fit fine to model for detecting a fraud in credit card Transactions. In our model, by using an artificial neural network (ANN) which gives accuracy approximately equal to 100% is best suited for credit card fraud detection. It gives accuracy more than that of the unsupervised learning algorithms.

From the above analysis, it is clear that many machine learning algorithms are used to detect the fraud but we can observe that the results are not satisfactory, so we would like to implement deep learning algorithms to detect credit card fraud accurately.

8. REFERENCES

1. A. Taha, S.J. Malebary, An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine, *IEEE Access* 8 (2020) 25579–25587, doi:10.1109/ACCESS.2020.2971354.
2. S. Makki, Z. Assaghir, Y. Taher, R. Haque, M. Hacid, H. Zeineddine, An experimental study with imbalanced classification approaches for credit card fraud detection, *IEEE Access* 7 (2019) 93010–93022, doi:10.1109/ACCESS.2019.2927266.
3. C. Jiang, J. Song, G. Liu, L. Zheng, W. Luan, Credit card fraud detection: a novel approach using aggregation strategy and feedback mechanism, *IEEE Internet Things J.* 5 (5) (Oct. 2018) 3637–3647, doi:10.1109/JIOT.2018.2816007.