Scientific Journal of Impact Factor (SJIF): 5.71

e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 8, Issue 03, March -2021

A Comparative Study of Machine Learning Techniques for Fraud Detection in Imbalanced Credit Card Datasets

Simran Sharma[1], Dr. N. C. Barwar[2]

[1] *Computer Science Department, M.B.M. Engineering College Jodhpur*
[2] *Professor and Head Department of Computer Science, M.B.M. Engineering College Jodhpur*

Abstract - *Along with the great increase of internet and e-commerce, the use of credit card is an unavoidable one. Due to the increase of credit card usage, the frauds associated with this have also increased. Fraudsters are continuously trying to find new ways and tricks to misuse the card and transparency of online payment. To detect such frauds, comparing the usage pattern and current transaction of a user over the past transactions, then classify it as either fraud or a legitimate transaction. Thus, to overcome these fraud activities we need a powerful fraud detection technique. To detect outliers, different machine learning algorithms such as logistic regression, Random forest, Naive Bayes, Support Vector Machine, KNN, Neural Network Algorithm are used. However, credit card dataset is imbalanced and the classification model can't apply directly on the imbalanced dataset because prediction may incline toward the majority cases so the resulted prediction can be wrong. So, dataset need to be converted into balanced dataset which is done by sampling methods. In this study, classification algorithms were applied on balanced and imbalanced dataset after that calculate the accuracy for each algorithm to measure the performance of algorithms and then compare the result of different machine learning algorithms to determine which algorithm give best result for identifying fraud transactions.*

Keywords- *Credit card; fraud detection; imbalanced dataset; resampling methods; logistic regression; random forest; naïve bayes; support vector machine; KNN; neural network*

I. INTRODUCTION

Credit card becomes the most popular method of payment for both online and offline so the cases of fraud associated with this have also increased. Illegal use of credit card or credit card information to make purchase without the knowledge of the cardholder is referred to as credit card fraud. The purpose may be to purchase goods without paying, or to transfer unauthorized funds from an account. There are several types of frauds: application fraud, electronic imprints, card not present, counterfeit fraud, lost and stolen fraud, card id theft, account takeover and many more. Considering only the credit card frauds, they can be of two kinds:

a) Offline Credit Card Fraud
b) Online Credit Card Fraud

Offline credit card frauds are those where an individual's credit card is lost or stolen. If any attacker or hacker, hacks the details and use it to commit illegal actions is referred as online frauds. Thus, to overcome these fraud activities we need a powerful fraud detection system. Figure 1.1 shows the overall system framework. A fraud Detection system is run on the bank of credit card issued. Each incoming transaction is sent to the FDS for verification. Fraud detection system receives the original card holder details and it verifies that the transaction is actual transaction or not. The type of items or products they purchased in that transaction are not known to the fraud detection system. This always tries to find any abnormality in the transaction depending on the profile of the card holder who doing shopping, shipping address and billing address etc. If the fraud detection system verifies that the transaction to be not valid or it to be malicious, it raises an alarm and the credit card transaction was declined by the issued bank. Then the investigators investigate to the card holder and provide feedback to the fraud detection system. The aim of FDS is to distinguish the fraudulent transaction and nonfraudulent transaction. By historic data, user's pattern and spending behavior used to check and verify that the transaction is fraudulent or not [7].

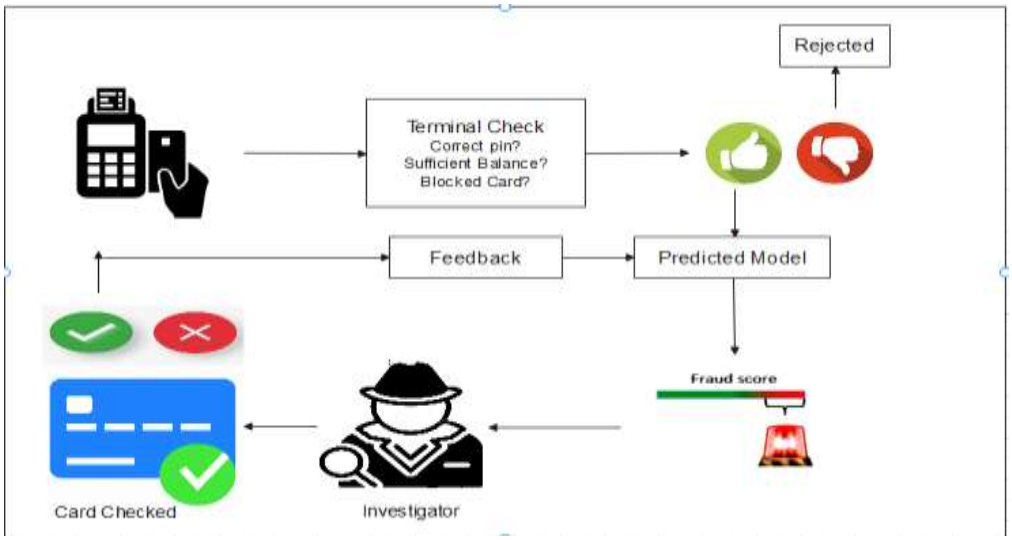

Figure 1.1: credit card fraud detection process

II. FRAUD DETECTION TECHNIQUES

Machin learning models are used in detecting credit card fraud. In our work, we have applied seven ML classifiers in the dataset which are: Logistic Regression, Decision Tree, KNN, Support Vector, Naïve Bayes, Random Forest and Neural Network.

1. Logistic Regression: Logistic Regression is a supervised classification method. This method returns the probability of binary response variable that is predicated from the independent variable of dataset that is logistic regression predict the probability of an outcome which has two values either zero or one, yes or no and false or true. Logistic regression has similarities to linear regression but as in linear regression a straight line is obtained, LR produces logistic curves which plots the values between zero and one [19].

2. Decision Tree: This is used as both classification and regression problem. It is a tree where the root and each of its internal nodes are labeled with a question about an independent variable. The interconnections from each internal node represent outcome to the given question and the leaf node represents a prediction of a solution to the problem. Rather than solving a huge complex problem, decision tree solves multiple easy sub problems and then combines the solution [16].

3. Support Vector Machine: SVM are a set of supervised learning method and the aim of the SVM is to design a hyperplane that classifies all the data vectors into different classes. There could be many hyperplanes but we have to find an optimal hyperplane that leaves the maximum margin from among the classes of data. The points closest to the hyperplane in the different classes are known as support vectors and these support vectors are used to predict the classes of new incoming data points. A new incoming data point is classified as to which class it belongs on the basis of which side of hyperplane it falls on the vector space. To train our machine we feed supervised data i.e data with results already known [16].

4. Random Forest: Random forests work as a large collection of decision trees. It is an ensemble learning model for classification and prediction. In this technique, the given large training dataset is divided into many random subsets. Since every data subset is randomly made, each subset is known as random tree and all random trees are collectively forming a random forest. For each subset, a decision tree is being constructed at training time. Then, a tuple is given as input for every decision tree and each decision tree in turn outputs the class. The class for the particular tuple/instance is predicted based on the most returned label by the decision trees. Hence polling is conducted among the decision trees to predict the class label for the given instance. The mostly voted label or the most output of decision trees will become the label for that particular instance. This is an ensemble process just to improve the accuracy of classifier model. Random forest has advantage over decision tree as it corrects the habit of overfitting to their training set [2].

5. K-NEAREST NEIGHBOUR: KNN Algorithm has been used in several anomaly detection techniques. It is one of the most used supervised algorithms for both classification and regression predictive problems. The performance of KNN algorithm is depends on three main factors [5]:

- The distance metrics.
- The distance rules.

- The value of K.

The distance metrics used to locate the nearest neighbours of any incoming data point. The distance rule used to classify the new data point into a class by comparing its features with that of data points in its neighbourhood. And the value of K decides the number of neighbours used to classify the new sample. There are various phases of KNN algorithm such as training phase, testing phase, classification phase. In training phase, the algorithm only stores the feature vectors and corresponding class labels. In testing phase, decisions are made by the algorithm on the basis of training data set. In classification phase, a single number is given to k, which decides how many neighbors influence the classification. The value of k can be large or small. If k=1, then it is called nearest neighbor algorithm. If value of k is large, it reduces the effect of noise on classification.

6. NAÏVE BAYES: Naive Bayes is based on Bayesian theory, which choose the decision based on highest probability. Bayesian probability estimates unknown probabilities from known values. It is also allows prior knowledge and logic to be applied to uncertain statements. The principle it follows is that every feature is independent of the remaining features. The Naïve Bayes is a supervised learning method which uses a training data set with known target classes to predict the classes of future instances [9].

7. NEURAL NETWORK: NN is inspired by the way the biological nervous system such as brain process information. it uses the processing of the brain as a basis to develop algorithms that can be used to model complex patterns and prediction problems. NN acquires a large collection of units that are interconnected in some pattern to allow communication between the units. These units, also referred to as nodes or neurons, are simple processors which operate in parallel. Every neuron is connected with other neuron through a connection link. Each connection link is associated with a weight that has information about the input signal. This is the most useful information for neurons to solve a particular problem because the weight usually excites or inhibits the signal that is being communicated. Each neuron has an internal state, which is called an activation signal. Output signals, which are produced after combining the input signals and activation rule, may be sent to other units [1].

III. DATASET DESCRIPTION

The Dataset is obtained from Kaggle datasets which contains credit card transactions done by European credit cardholders in September 2013. This dataset presents transactions that occurred in two days, out of 284,807 transactions 492 were frauds and remaining 284,315 were labeled as nonfraudulent. The data is highly imbalanced, the fraud transactions account only for 0.172% of total transactions. Some data pre-processing has already been done on the data. The original data are not exposed due to the high confidentiality. The data set contains a total of 30 numerical input variables out of which 28 variables are the result of PCA. Features from V1, V2, V3……,V28 have obtained by PCA. Features such as Time and Amount have not been transformed using PCA. Time depicts the time between first transaction and every other transaction in the dataset. Amount is the transaction Amount. Class is the target class for binary classification and it takes value 1 for fraud and 0 for nonfraudulent [8].

IV. PROPOSED METHODOLOGY

In this research work, credit card transaction dataset is used and this dataset contains a dependent variable that classifies that the new transaction is either fraudulent or not. In this different machine learning algorithm such as logistic regression, Random forest, Naive Bayes, Support Vector Machine, KNN and Neural Network Algorithm are used for detecting the fraud in credit card data set. The accuracy for each algorithm is calculated to measure the performance of algorithms and then compare the result of different machine learning algorithms to determine which algorithm give best result for identifying fraud transactions. python is use to design for the experimentation with machine learning algorithms.

In this research work, there are 2 Experimental Setup is used to measure the performance of machine learning algorithms for the credit card fraud dataset.

I. Compare different machine learning algorithms performance using credit card fraud data set for 30 attributes with 1 target attribute and 284,807 instances.

II. Compare different machine learning algorithms performance using credit card fraud data set for 30 attributes with 1 target attribute and 284,807 instances with sampling methods.

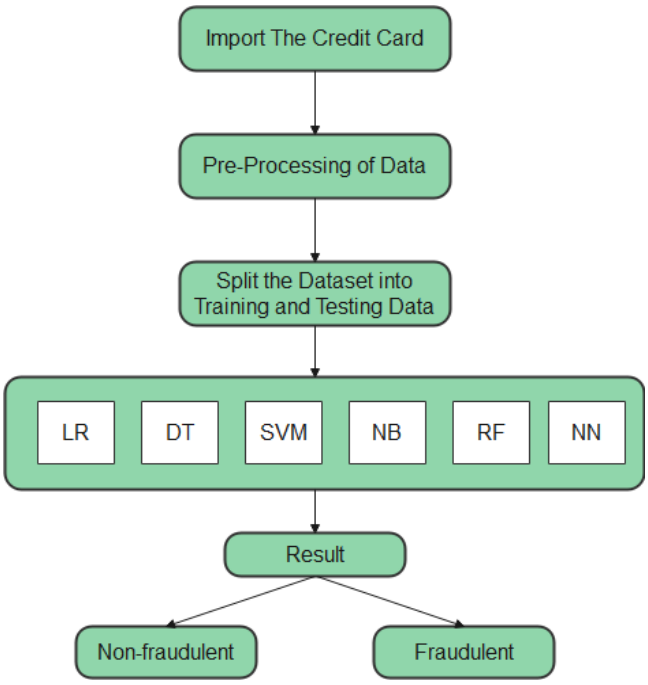

Figure 1.2: Work Flow of System Architecture

The proposed model is designed to predict that the transaction is fraudulent or not. In our architecture first machine read the data. The next step is data pre-processing where the dataset will be analysed and all the null values and duplicate values will be removed from the dataset. After that split the entire dataset into 70:30 ratio where 70% transaction are considered as a training set and the remaining 30% transaction are testing set. To detect outliers, different machine learning algorithms such as logistic regression, Random forest, Naive Bayes, Support Vector Machine, KNN, Neural Network Algorithm are used. Now the performance analysis will be done on the obtained confusion matrix. The calculated accuracy for each algorithm is used to measure the performance of algorithms and then compare the result of different machine learning algorithms to determine which algorithm give best result for identifying fraud transactions.

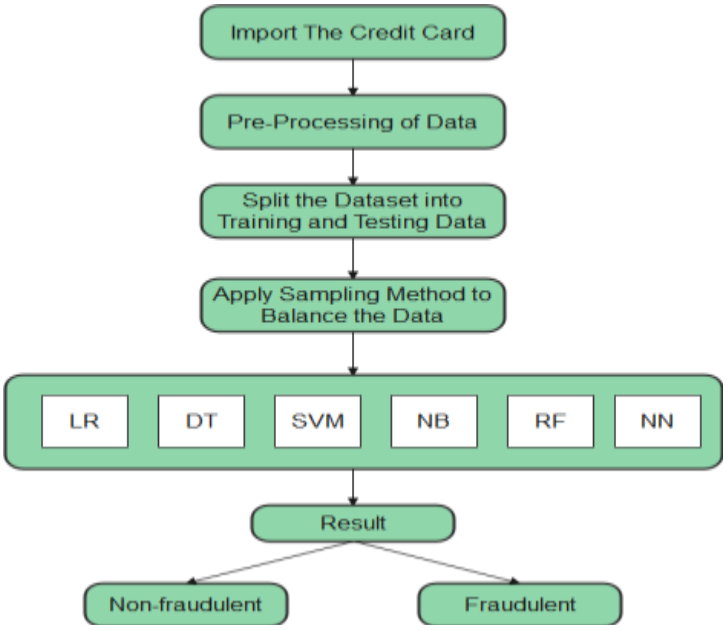

Figure 1.3: Work Flow of System Architecture with Sampling Method

In second case as shown in figure 1.3 the only difference would be that we will apply sampling methods to the training data set. Our dataset is unbalanced dataset and the classification models can't be applied directly on the unbalanced dataset so it need to be converted into balanced dataset which is done by sampling methods. In the previous case we applied classification model on the imbalanced dataset where non fraudulent cases are higher than the fraudulent cases so the prediction may incline toward the majority cases so the resulted prediction can be wrong. Further we will apply our machine learning algorithm on this balanced data to see which algorithm will perform better on the provided data. Our aim here is to reduce the error with minimum number of components as possible. We used sampling method with different algorithms and try to increase the accuracy. With different algorithm we get different results and accuracy.

V. RESULTS

After visualization we scaled the time and amount features then split the dataset into two part i.e. 30% for the testing data and remaining for training data. we will apply our first algorithm i.e. logistic regression and the algorithm gives the test accuracy of 99.92%. Decision Tree gives the test accuracy of 99.92. Naïve bayes algorithm gives 97.84% of test accuracy. Support vector machine gives the test accuracy of 99.94%. K-mean accuracy score is 79.46%. Random forest gives the test accuracy score of 99.95%. Neural network gives the test accuracy score of 99.94% in 05 epochs. Autoencoder gives the test accuracy score of 97.97% in 50 epochs.
In the first experimental setup where we apply classification algorithm on imbalanced dataset and we got maximum accuracy in Random forest algorithm, with accuracy score of 99.95% and the neural network is the second-best algorithm with 99.94% accuracy.

Table 1.1 Classification Outcome of Techniques

Techniques	Class	Precision	Recall	F-1 Score	Support
Logistic Regression	0	1.00	1.00	1.00	85296
	1	0.88	0.62	0.73	147
	Avg.	0.94	0.81	0.86	85443
Decision Tree	0	1.00	1.00	1.00	85296
	1	0.80	0.76	0.78	147
	Avg.	0.90	0.88	0.89	85443
Random Forest	0	1.00	1.00	1.00	85296
	1	0.95	0.76	0.85	147
	Avg.	0.97	0.88	0.92	85443
Naïve Bayes	0	1.00	0.98	0.99	85296
	1	0.06	0.84	0.12	147
	Avg.	0.53	0.91	0.55	85443
Neural Network	0	1.00	1.00	1.00	85296
	1	0.85	0.81	0.83	147
	Avg.	0.92	0.90	0.91	85443
SVM	0	1.00	1.00	1.00	85296
	1	0.95	0.69	0.80	147
	Avg.	0.98	0.84	0.90	85443
K-Mean	0	1.00	1.00	1.00	85296
	1	0.84	0.31	0.46	147
	Avg.	0.92	0.66	0.73	85443

Table 1.2 Different classification algorithm with their accuracy scores.

Classification algorithm	Accuracy scores
Logistic regression	99.92%
Decision Tree	99.92%
K-mean	79.46%
Neural Network	99.94%
Random Forest	**99.95%**
Support Vector Machine	99.93%
Naïve bayes	97.84%

Table 1.3 Classification Outcome of Classification Techniques using RUS Method

Techniques	Class	Precision	Recall	F-1 Score	Support
Logistic Regression	0	1.00	0.97	0.98	85296
	1	0.05	0.91	0.09	147
	Avg.	0.52	0.94	0.54	85443
Decision Tree	0	1.00	0.92	0.96	85296
	1	0.02	0.93	0.04	147
	Avg.	0.51	0.93	0.50	85443
Random Forest	0	1.00	0.98	0.99	85296
	1	0.07	0.90	0.13	147
	Avg.	0.54	0.94	0.56	85443
Naïve Bayes	0	1.00	0.98	0.99	85296
	1	0.06	0.86	0.11	147
	Avg.	0.53	0.91	0.55	85443
Neural Network	0	1.00	1.00	1.00	85296
	1	0.46	0.84	0.60	147
	Avg.	0.73	0.92	0.80	85443
SVM	0	1.00	0.99	1.00	85296
	1	0.14	0.86	0.24	147
	Avg.	0.57	0.93	0.62	85443
K-Mean	0	1.00	1.00	1.00	85296
	1	0.84	0.31	0.46	147
	Avg.	0.92	0.66	0.73	85443

Table 1.4 Classification Results using RUS with their Accuracy Scores.

Classification algorithm	Accuracy scores
Logistic regression	96.80%
Decision Tree	99.05%
Random Forest	98.00%
Neural Network	**99.80%**
Naïve bayes	97.53%
SVM	99.03%
K-Mean	96.87%

Table 1.5 Classification Outcome of Classification Techniques using ROS Method

Techniques	Class	Precision	Recall	F-1 Score	Support
Logistic Regression	0	1.00	0.98	0.99	85296
	1	0.07	0.93	0.12	147
	Avg.	0.53	0.95	0.56	85443
Decision Tree	0	1.00	1.00	1.00	85296
	1	0.81	0.76	0.79	147
	Avg.	0.91	0.88	0.89	85443
Random Forest	0	1.00	1.00	1.00	85296
	1	0.94	0.78	0.86	147
	Avg.	0.97	0.89	0.93	85443
Naïve Bayes	0	1.00	0.97	0.99	85296
	1	0.06	0.86	0.10	147
	Avg.	0.53	0.92	0.55	85443
Neural Network	0	1.00	1.00	1.00	85296
	1	0.60	0.83	0.69	147
	Avg.	0.80	0.91	0.85	85443
SVM	0	1.00	0.99	1.00	85296
	1	0.14	0.88	0.23	147
	Avg.	0.57	0.93	0.61	85443
K-Mean	0	1.00	1.00	1.00	85296
	1	0.84	0.31	0.46	147
	Avg.	0.92	0.66	0.73	85443

Table 1.6 Classification Results using ROS Method with their Accuracy Scores.

Classification algorithm	Accuracy scores
Logistic regression	97.74%
Decision Tree	99.92%
Random Forest	**99.95%**
Neural Network	99.87%
Naïve bayes	97.44%
SVM	99.01%
K-Mean	97.87%

Table 1.7 Classification Outcome of Classification Techniques using SMOTE Method

Techniques	Class	Precision	Recall	F-1 Score	Support
Logistic Regression	0	1.00	0.98	0.99	85296
	1	0.06	0.92	0.11	147
	Avg.	0.53	0.95	0.55	85443
Decision Tree	0	1.00	1.00	1.00	85296
	1	0.55	0.72	0.62	147
	Avg.	0.77	0.86	0.81	85443
Random Forest	0	1.00	1.00	1.00	85296
	1	0.90	0.79	0.84	147
	Avg.	0.95	0.89	0.92	85443
Naïve Bayes	0	1.00	0.98	0.99	85296
	1	0.06	0.86	0.11	147
	Avg.	0.53	0.92	0.55	85443
Neural Network	0	1.00	1.00	1.00	85296
	1	0.40	0.84	0.54	147
	Avg.	0.70	0.92	0.77	85443
SVM	0	1.00	0.99	0.99	85296
	1	0.10	0.87	0.18	147
	Avg.	0.55	0.93	0.99	85443
K-Mean	0	1.00	1.00	1.00	85296
	1	0.84	0.31	0.46	147
	Avg.	0.92	0.66	0.73	85443

Table 1.8 Classification results using SMOTE method with their accuracy scores.

Classification algorithm	Accuracy scores
Logistic regression	97.49%
Decision Tree	99.77%
Random Forest	**99.95%**
Neural Network	99.78%
Naïve bayes	97.54%
SVM	98.66%
K-Mean	98.87

In our second experimental setup where we apply classification algorithm on balanced dataset and we get different results. In RUS, Neural Network has highest accuracy rate which is 99.80%. In ROS, Random Forest give the accuracy of 99.95%. In SMOTE, Random Forest give the accuracy of 99.95%.

VI. CONCLUSION

Due to the rapid growth of electronic commerce technology, the use of credit cards has dramatically increased. As usage of 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 increased. Thus, to overcome these fraudulent transactions there is a need to find a powerful fraud detection technique. The major task of today is to build an accurate, precise and fast detecting credit card FDS that can detect not only frauds happening over the internet like site cloning and phishing but also signals an alarm when the tampered credit card is being used. There are some Machine learning technique like Logistic regression, Decision Tree, Support Vector, KNN, Neural Network, Random forest and neural network were used to detect the fraud in credit card system. In this study, I applied machine learning technique to predict whether a credit card transaction is fraudulent or not. For this, we use European dataset which contain 284,807 transaction out of which only 492 are fraudulent. The data is highly imbalanced, fraud transaction account only for 0.172% of total transactions. In this study, there are two experiment setups is used to measure the performance of machine learning algorithm or credit card fraud dataset. In first experiment, compare different machine learning performance using credit card fraud dataset for 30 attributes with 1 target attribute and 284,807 instances. When this unbalanced data is directly applied on the classification method then the model gives biased result toward the minority samples. As a result, it tends to misrepresent a fraudulent transaction as a genuine transaction. In this the comparison result revealed that, random forest gives the maximum test accuracy of 99.95% and the neural network is the second-best algorithm with 99.94% accuracy. In second experiment, compare different machine learning algorithm perform using credit card fraud dataset for 30 attributes with 1 target attribute and 284,807 instances with sampling methods. When we apply classification algorithm on balanced dataset then we get different results. In RUS, Neural Network has highest accuracy rate which is 99.80%. In ROS, Random Forest give the accuracy of 99.95%. In SMOTE, Random Forest give the accuracy of 99.95%.

VII. ACKNOWLEDGEMENT

I would like to thank N. C. Barwar Professor and Head Department of Computer Science, M.B.M. Engineering College Jodhpur, for his motivation without which I would not be able to write this paper. I am grateful to my family and friends for their support.

VIII. REFERENCES

[1] Ghosh and D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network", Proceedings of the 27th Hawaii Interntional Conference on System Sciences (1994), vol.3 pp.621-630.
[2] Nur Farhana Hordri and Siti Sophiayati Yuhaniz, "Handling Class Imbalance in Credit Card Fraud using Resampling Methods", International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 9, No. 11, 2018.
[3] Ankit Mishra, Chaitanya Ghorpade, "Credit Card Fraud Detection on the Skewed Data Using Various Classification and Ensemble Techniques" ,2018 IEEE.
[4] Rimpal R.Popat, Jayesh Chaudhary, "A Survey on Credit Card Fraud Detection Using Machine Learning ", Proc. Of the 2nd International Conference on Trends in Electronics and Informatics (ICOEI) IEEE Conference Record,2018.
[5] C.Sudha,T.Nirmal Raj , "Credit Card Fraud Detection In Internet Using KNN Algorithm", IPASJ International Journal of Computer Science(IIJCS),Volume-5,2017.
[6] John Richard D.S Kho and Larry A. Vea, "Credit Card Fraud Detection Based on Transaction Behavior", Proc. of the 2017 IEEE Region 10 Conference (TENCON), Malaysia, November 5-8, 2017.
[7] V.Bhusari and S.patil. "Study of Hidden Markov Model in Credit Card Fraudulent Detection." International Journal of Computer Applications Volume 20– No.5, April 2011.
[8] Ayushi agarwal,Shiv Kumar,Amit kumar Mishra "Credit card fraud detection: A case study," in IEEE, New Delhi, India, 2015.
[9] Yashvi Jain, Namrata Tiwari, Shripriya Dubay, Sarika Jain, "A Comparative Analysis of Various Credit Card Fraud Detection Techniques", International Journal of Recent Technology and Engineering (IJRTE), Volume -7,2019.
[10] Mrs.Vimala Devi. J, Dr.Kavitha,K.S" Fraud Detection in Credit Card Transactions by using Classification Algorithms "International Conference on Current Trends in computer ,Electrical ,Electronics and Communication (ICCTCEEC-2017).
[11] Mitchell, T," Machine Learning", McGraw Hill, 1997.
[12] Lecture Notes- "Deep Learning": Prof. Prabir Kumar Biswas, IIT Kharagpur from Swayam online learning.
[13] Lecture Notes- "Python for Data Science": Prof. Ragunathan Rengasamy, IIT Madras from Swayam online learning.
[14] Lecture Notes- "Practical Machine Learning with Tensorflow": By Prof. Ashish Tendulkar, Prof. Balaraman Ravindran, Google, IIT Madras from Swayam online learning.
[15] Krishna Modi and Reshma Dayma, "Review on Fraud Detection Methods in Credit Card Transactions", International Conference on Intelligent Computing and Control,2017.
[16] Y. Sahin and E. Duman, "Detecting Credit Card Fraud by Decision Trees and Support Vector Machines" , International Multi conference of Engineers and Computer Scientists, volume 1,March,2011,Hong-Kong.
[17] Lakshmi S V S S, Selvani Deepthi Kavila, "Machine Learning For Credit Card Fraud Detection System", International Journal of Applied Engineering Research ISSN,Volume 13,2018.
[18] John O., Adebayo O. Adetunmbi, and Samuel A. Oluwadaren Awoyemi, "Credit card fraud detection using machine learning techniques: A comparative analysis.", International Conference on Computing Networking and Informatics (ICCNI), pp. 1-9, 2017.
[19] Y. Sahin and E. Duman, "Detecting Credit Card Fraud by ANN and Logistic Regression ", In Innovations in Intelligent System and Applications(INISTA), 2011 International Symposium on (pp. 315-319) IEEE.
[20] Sahil Dhankhad, E.A. Mohammed, B.Far, "Supervised Machine Learning Algorithm for Credit Card Fraudulent Transaction Detection: A Comparative Study", IEEE International Conference on Information Reuse and Integration for Data Science,2018.
[21] Chandrahas Mishra, D.L. Gupta, Raghuraj Singh, "Credit Card Fraud Identification Using ANN",International Journal of Computer System,Volume 4,2017.
[22] N.Malini ,Dr.M.pushpa "Analysis on Credit Card Fraud Identification Techniques based on KNN and Outlier Detection" 3rd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEEICB17) 978-1-5090-5434-3©2017 IEEE.
[23] Himani Bhavsar, Mahesh H.panchal,"A Review on Support Vector Machine for Data Classification",International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 1, Issue 10, December 2012.
[24] Tanmay Kumar Behera, Suvasini Panigrahi"Credit Card Fraud Detection: A Hybrid Approach Using Fuzzy Clustering & Neural Network " 2015 Second International Conference on Advances in Computing and Communication Engineering.
[25] Xiaoguo Wang, Hao Wu, Zhichao Yi "Research on Bank Anti-fraud Model Based on K-means and Hidden Markov Model" 2018 3rd IEEE International Conference on Image, Vision and Computing.

[26] Tanmay Kumar Behera, Suvasini Panigrahi, "Credit Card Fraud Detection: A Hybrid Approach Using Fuzzy Clustering & Neural Network" 2015 Second International Conference on Advances in Computing and Communication Engineering.
[27] Raghavendra Patidar, Lokesh Sharma, "Credit Card Fraud Detection Using Neural Network", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1, Issue-NCAI2011, June 2011.

