Global Transitions Proceedings 3 (2022) 31–37

Contents lists available at ScienceDirect

Global Transitions Proceedings

journal homepage: http://www.keaipublishing.com/en/journals/global-transitions-proceedings/

Exploratory analysis of credit card fraud detection using machine learning techniques

M J Madhurya [a], H L Gururaj [a,*], B C Soundarya [a], K P Vidyashree [a], A B Rajendra [a]

[a] *Vidyavardhaka College of Engineering, Mysuru, India*

ARTICLE INFO

Keywords:
Class imbalance
Data-driven model
Data prediction
Illegitimate
Malicious
Vector machine

ABSTRACT

In today's world, a lot of processes are carried over the Internet to make our lives easier. But, on the other hand, many unauthorized and illegitimate activities that take place over it are causing major trouble for the growth of the economy. One of them being the fraud cases that misguide people and lead to financial losses. Major frauds reported recently occur through the malicious techniques that are made to work on Credit cards that are used for financial transactions over online platforms. Hence, it is the need of the hour to investigate this problem. Several companies have started their study in this regard and have formulated data driven models that use various Machine Learning algorithms and models on datasets to analyse false activity. Several techniques used are Support Vector Machine, Gradient Boost, Random Forest and their mixtures. In this comparative study, the anomaly of class imbalance and ways to implement its solutions are analysed to prove certain results. The effectiveness of the algorithms varies on the set of data and the instance in which it is used. They prove that all algorithms despite of all the calculations show certain imbalance at some point in the study The limitations have also been evaluated and highlighted to help in future. In this study, it is found that although logistic regression had more accuracy but when the learning curves were plotted it signified that the majority of the algorithm under fit while KNN has the ability only to learn. Hence KNN is better classifier for the credit card fraud detection.

1. Introduction

The word 'Fraud', it could be understood as the act of intentional deception and dishonesty intended towards personal gain. Now, with the Internet taking over our lives, many people and businesses have become the target for fraudulent activities. Several reports claim that growth of commercial fraud attempts has risen in 2018 compared to 2016. Frauds in those years have unstripped each other by a whopping 83%. The E-commerce Fraud Index has claimed that fraud rate in stores have risen from 0.06% in 2016 to 0.23% in 2017. 10% of all frauds are considered exclusively of Credit cards that have resulted in huge financial losses that worry companies. Since much of the transactions are digitized there has been an increase in the number of cards that are active, and their transaction data has been multiplying more than ever. Therefore, the amount of data to be examined during the detection process has become voluminous. The main tools used by researchers are ML Algorithms, Neural Networking models, Classification and Clustering techniques. Many researchers are also working on early or pre-detection of credit card frauds. Other research scholars have also investigated feasible and efficient methods and ways for Fraud detections. ML and other correlated approaches are usually used, for example, ANN (Artificial Neural Network), the method of rule induction system, Logistic Regression (LR), Decision Trees (DT), and the Support Vector Machine etc. ML algorithms are AI techniques that have the ability to solve various problems from diverse disciplines and fields that usually possess large amounts of data. Though there have been many ideas and solutions proposed to prevent and detect fraud, there is still a big need to apply and analyze the strength of ML algorithm.

1.1. Clustering

Here, the large group of data is divided into smaller and similar ones based on the similarities that they have in their nature and form clusters. Items in different clusters may not have the property of other cluster elements as pictorially represented in Fig. 1.

1.2. Classification and Methods

When certain values are given as an input from a huge set of data, these algorithms find a common interest or a conclusion on its basis, it means these methods try to extract one or more output from input that is given. ML algorithms are useful while performing these activities. The Fig. 2 below shows the classifications applied in Machine Learning.

* Corresponding author.
E-mail address: gururaj1711@vvce.ac.in (H.L. Gururaj).

https://doi.org/10.1016/j.gltp.2022.04.006

Available online 6 April 2022

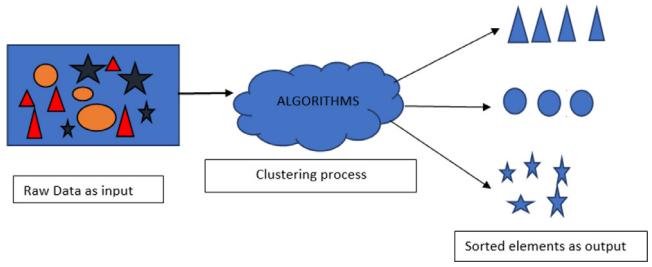

Fig. 1. Clustering technique.

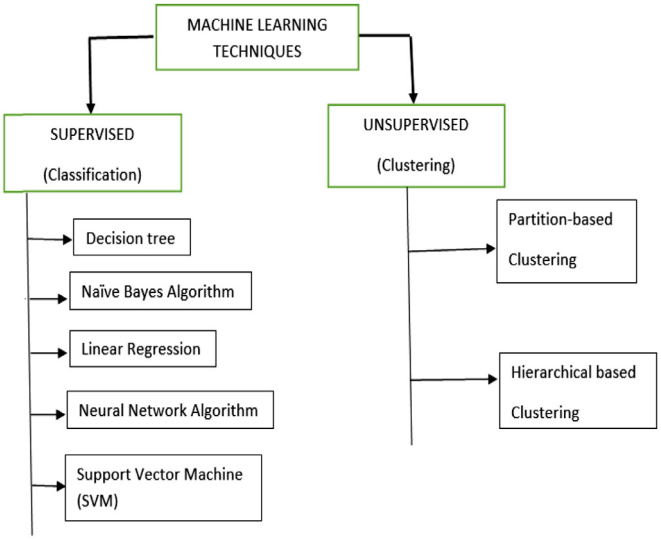

Fig. 2. Classification methods of Machine Learning.

1.3. Construction of references

This is a supervised learning technique that finds out the output directly depending on the input data that is given by the user. This is mostly used in continuous data that cannot have any particular discrete values. ML algorithms are used in such predictions. The main aspect behind this paper is to analyze the performance, accuracy, and efficiency of various Machine Learning classifiers while using it on the prediction analysis and preventive analysis of certain algorithms. Key factors such as additive techniques like oversampling, binary classification etc. It is observed that traditional classifier algorithms of ML while paired with such techniques give better results and increase efficiency of the process. The remaining part of the paper is organized and divided into sections pertaining to the research. In Section 2, there is a summary on the related works of the problem. In Section 3, the architecture and methodological study with details is given by inferring various classifiers. In Section 4, the results along with graphical representations have been established to show the analyzed test results and their values. In the end, in section 5, we draw some conclusions and throw light on future scope of the problem and its limitations to be focused upon in further studies.

2. Related work

Credit card transactions are either classified as fraudulent or legitimate transactions and are mainly a binary classification problem. Basically, data mining classification comprises problems such as Fraud detection that is used to figure out credit card transactions as fraudulent or legitimate. Some Additional techniques and factor methods apart from data mining which are involved in fraud detection are Web-services based collaborative schemes in which the private bodies like

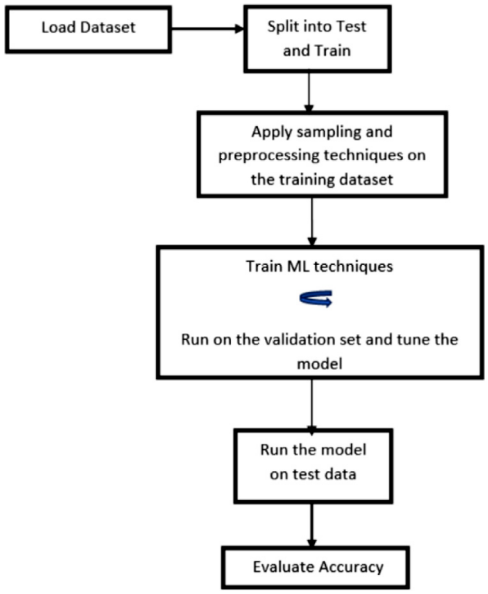

Fig. 3. The flow process diagram for developing a machine learning model.

banks can share the information about the fraud patterns and frequencies for the enhancement of the fraud detection capability and to reduce the financial loss. The basic procedure in developing any Machine Learning model [1] is given below in Fig. 3.

Numerous Machine Learning techniques have been developed and used in the various experimental studies to approach the problem of fraud detection in Credit cards.

In the [1], a data driven approach to set up fraud alerts has been proposed that runs on select features like Oversampling under sizes and SMOTE technique. A few authors in their studies [2] have come across the comparison among various models and their resultant analysis, like XGBOOST, Random Forest, Decision Trees etc. These are the most widely used techniques so far in detection of frauds. There has also been a study of new techniques like Adaboost and Majority Voting approaches that add or enhance the ML algorithm performance. [3,4]

Feature Selection (FS) with optimization used in Artificial Neural Network (ANN) for the selection of the required features while the implementation of the algorithms has been seen [5]. When more than one valid parameter is present, it becomes important to select the best effective feature. Since most models are irrelevant in transaction sequencing, they cannot learn by information at a single level, hence a new structured sequenced learning ensemble classifier that improves performance is also seen [6,7]. In another paper [8] by S. Venkata Suryanarayana et al. Many classifications and their metrics and performance have been analysed. This gives an idea as to how many metrics can be considered while finding a proper algorithm. Adaptive features selection processes by comparative study of 5 different techniques can be analysed [9], these adaptive features make it easier to bifurcate and eliminate the unimportant ones. The figure below shows the accuracy and precision found in [10,11].

The organized complete study on the application of Random Forest has been discussed by Priya Gupta et al in the paper that in detail analyses the key factors of RF and its limitations [12]. Also, and the concept of real time deep learning and binary data classification by various methods has been discussed in the paper cited [13] which comparatively analyses these techniques. A strong and new tool of bidirectional Long short-term memory (BiLSTM) and bidirectional Gated recurrent unit (BiGRU) is looked upon by Hassan Najadat and others who have also applied various other six strategies like Ada Boost etc [14]. to help enhance the performance [15]. The Fig. 4 represents the normal credit card transaction.

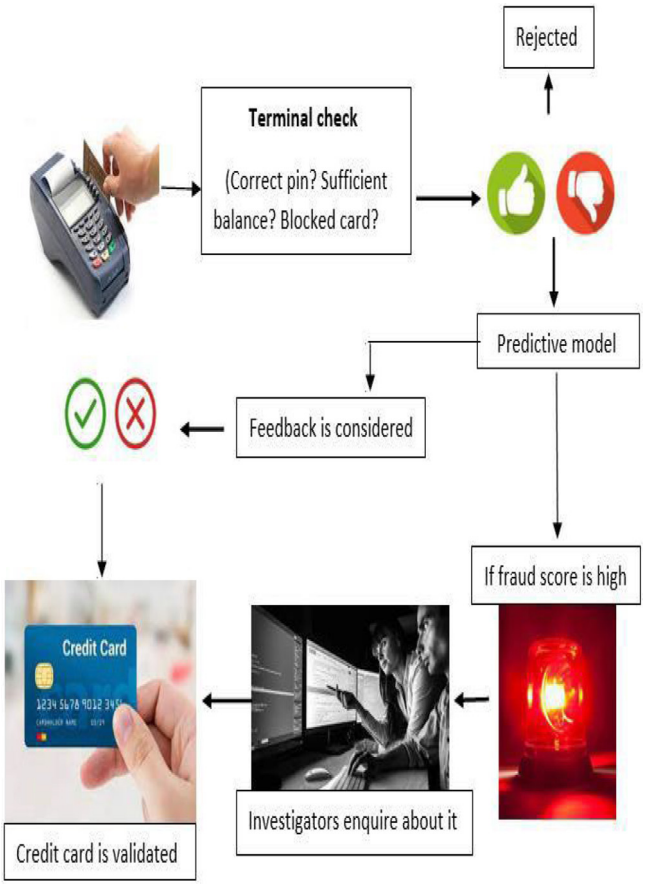

Fig. 4. Working of a credit card fraud detection model [5].

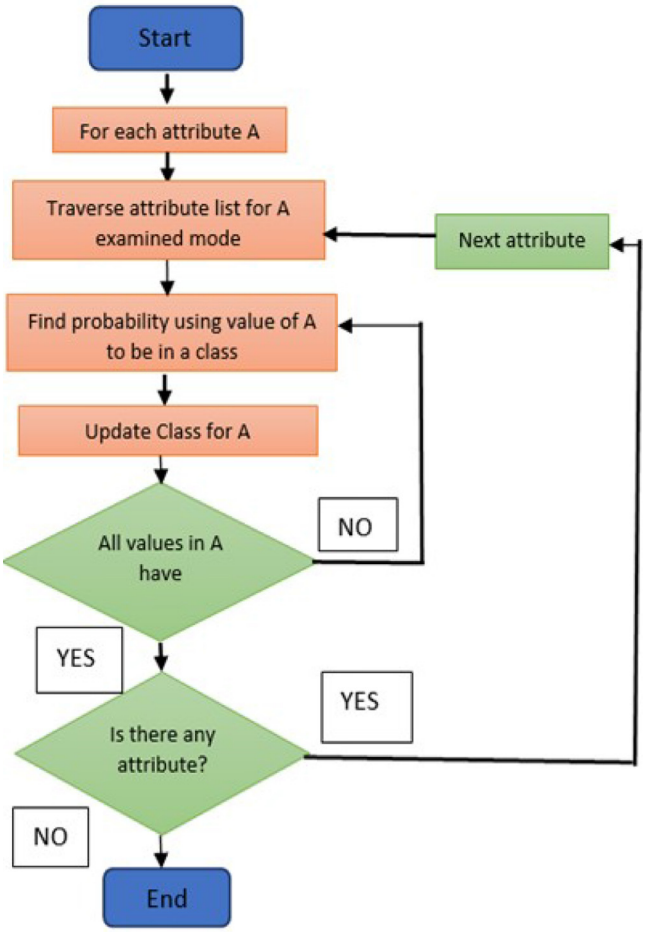

Fig. 5. Naïve Bayes algorithm.

The Novel strategies that effectively address the skew distribution of data are assessed by ML technique along with the implementation of the API (Application Programming Interface) module to decide if a transaction is fraudulent or legitimate are studied [16,17].

In the study by Aadhya Kaul and others it is seen it that compares the various techniques for the most suitable method, a certain ML criterion has been used on the dataset and F1-score of those are calculated to analyse it, [18] and Naive Bayes is introduced in arbitrary classification. Data processing techniques and their abilities are compared with those of the Machine Learning techniques and the prediction is done based on it. [19] also discusses the supervised based classification of dataset using Normalization and Principal element analysis and is seen 95% accuracy.

The hybrid approach that is made of combining the strength and credibility of three sub-methods [20], the GridSearchCV for Hyper-Parameters Optimization (HPO), the Recursive Feature Elimination (RFE) for the selection of useful predictive features, and the Synthetic Minority Oversampling (SMOTE) to overcome the imbalanced or disproportionate data problem is also proposed [21]. An auto-encoder based deep learning technique and restricted Boltzmann Machine (RBM) are implemented in hidden layers to find patterns [22] and anomalies in the huge set of data has been proposed by Apapan Pumsiriratin and others in their paper for the detection processes. The results show the mean squared error and area under the normal curve. This paper strictly focused on the comparison between various classifiers and techniques in Machine Learning and their performance accuracy.

3. Discussions

Fraud detection in simple words is a simple binary classification problem in which any particular transaction or exchange will be either classified as fraud or legit only. In this study, a few standard classification techniques like Naive Bayes, K-Nearest Neighbor and Logistic Regression methods, Random Forest Classifiers, Decision Trees. For effective usage of these algorithms, different stages are included such as gathering the data, cleaning data, researching and visualization of data and training the classifier algorithms and finally evaluating the result.

3.1. Naïve Bayes algorithm

It is a theory based on two assumptions. Initially, all features in a given entry that needs to be bifurcated contribute equally. Secondly, all the given attributes are statistically independent of each other, which implies that the values that attribute present do not show us any points about the attributes. This might not be true in all cases and Bayes rule is used in such a situation to find out if it's either fraudulent or legit. The class that is associated with the higher probability is taken as the predicted class for instance. Refer Fig. 5.

Naive Bayes has a base of restrictive independence among the various characters present in the dataset and the resultant classifier is based on restrictive probabilities of the matching options and is provided in Eq. (1).

$$\begin{aligned} &P[C(\mathrm{i})|f(\mathrm{k})] \times P[(\mathrm{i})] \\ &/P[f(\mathrm{k})](\mathrm{i})P[f(\mathrm{k})|C(\mathrm{i})] \\ &= \Pi P[f(\mathrm{k})|C(\mathrm{i})]k \\ &= 1\ldots, n : i = 1, 2\mathrm{etc}. \end{aligned} \tag{1}$$

Here, unknown probabilities are identified only with the already existing known values by making use of Bayesian concepts which will require prior knowledge to ensure logics. The flowchart given below in

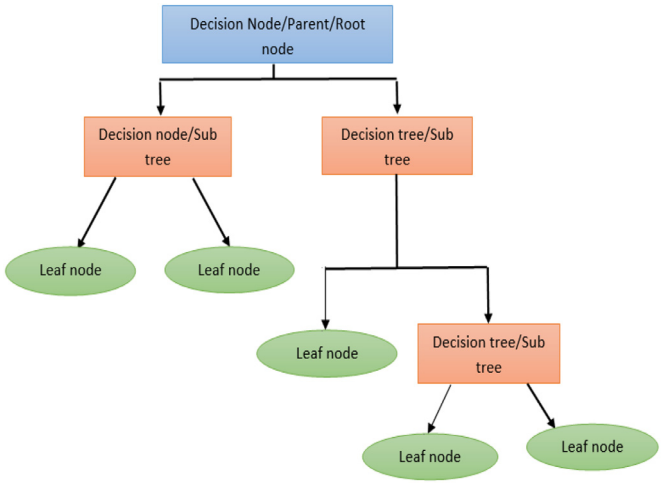

Fig. 6. Structure of a decision tree.

Fig. 5 [23] represents the structural working of this algorithm with various predictions.

3.2. Decision tree algorithm

In this method, there are two types of namely regression trees and classification trees. Here, a training dataset is used to construct a decision tree. This decision tree has separate nodes that form the structure of the tree, the topmost node of the tree is known as the root node. The other non-leaf nodes represent the test done on the attribute, each resulting branch denotes the outcome of the test and each leaf node on the tree denotes a class label.

These leaf nodes also show the classes that are returned if reached as the final prediction by the model. So, one can find out the prediction by properly traversing through the decision tree. A few of the decision tree algorithms include the C4.5, CART, ID3. This algorithm manages the constant set of data and uses the divide and solve approach to solve the main problem into subproblems through its repeated usage. The structure of a Decision Tree is as follows in Fig. 6 [12].

3.3. K-NN algorithm

This works on a simple logic that it plots all the existing training instances and later classifies the unlabelled instance based on the idea of their nearest neighbour present. Unlike the decision trees, here the instances are directly used to analyse. But it is also known that here, all existing algorithms are already instance-based as they are built on training models. Here, in this case, the unlabelled instance is classified and divided based on calculating the distances between the instances and by using the metric. The one which has the majority class is labelled for unlabelled class. Let us visualize how this algorithm works in steps depicted in Fig. 7.

- First, load from the data and then assign a certain value of K into the required group of neighbours.
- Secondly, we find the Euclidean distance between the test data set and the training data set. This will make the instance into a well-organized collection.
- Now, we must sort that ordered collection of indices and then order them in ascending order of the distance.
- The value of K is now initialized and the labels for chosen K entries.
- In the end, the values of mean and mode for K labels for classification and regression.

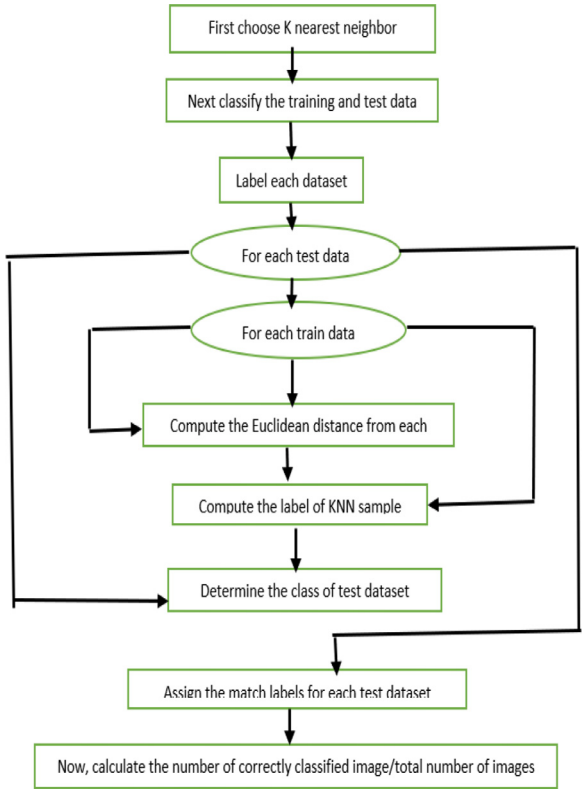

Fig. 7. KNN flowchart.

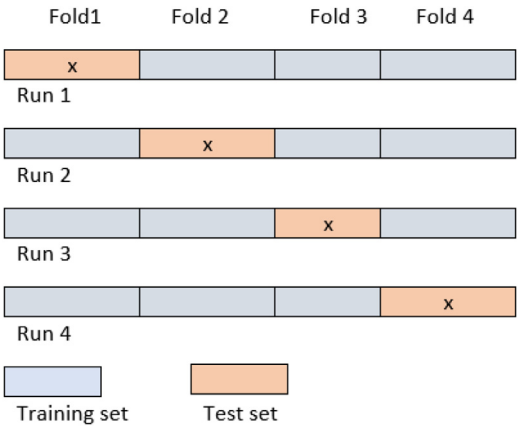

Fig. 8. Fold validation process.

3.4. Random forest classifiers

This algorithm is nothing but a Bayes classifier of Random Forest which is actually a simple implementation of Decision trees [12]. It could also be considered as a part of the Logistic Regression process [6]. It is a new approach to ensemble-tree based algorithms and is a learning algorithm. It is mainly selected from a series of randomly selected subsets of the taken training set for the experiment. So, as the trail goes on, objects of the class are decided upon the votes of all the other trees present in it. Let us understand the working of the Random Forest classifier algorithm in steps and depicted in Fig. 8.

- Before any of the test processes, the Python libraries are imported, and the required data set is loaded into the data frame.
- Now, that data is segregated into train and test dataset accordingly.
- The Random Forest Regression model is later applied on the training data that exists.

- After these procedures, the test outcome is gauged and predicted, and the suitable confusion matrix is made.

3.5. Support vector machine

SVMs are particularly depended on the structural risk minimization [6,12] unlike the other neural networks which focus on the empirical risk minimization. This technique was brought into light by Vapnik in 1992[12] to debug and solve only the binary classification problem, but now it is extended towards the non-linear regression also. These SVMs map a certain data to a pre-existing very high dimensional space via a particular kernel function and thereby finds the hyperplane that maximizes the margin between any two classes. The Solution SVM problems are based on those data points that exist particularly at the margin. Such points are called support vectors [6].

3.6. Artificial neural network

Artificial neural networks have been in existence and developed overtime by many scientists which are inspired by biological neural networks (neurons). These networks are then utilized to and approximate those functions/factors that can be largely determined by an input that are unknown. An artificial neural network is constructed in the nodes one by one in a stacked manner. They are stacked between a coating of the target vector and the feature vector [12]. ANN is developed to have a similar kind of interpretation just like our human brain that learns from various activities.[6] ANN has been successfully implemented in many of the model is seen.

4. Result and discussions

This paper is aimed at looking for a suitable algorithm to tackle the large amounts of data that are taken as the input for a fraud detection model and therefore a factual comparison of the Machine Learning techniques has been done on a credit card dataset considered.

4.1. F1 strategy

To estimate the following models and evaluate their results, the dataset is broken into training and testing data. Let us investigate one conclusion drawn on the hold of accuracy and F1-score which was calculated using the confusion matrix. Now, let us gauge the algorithms based on the other three parameters to get a better understanding of their efficiency discussed in Section 4.4.

4.2. Data analysis and pre-processing

The raw dataset taken for the study was sorted and pre-processed for the sole intention of improving the performance of the classifiers and reducing their training and operating time. If not, the data would require lots of time in between to sorted based on their common features. The pre-processing also includes the work of investigating the dataset feature space and handling the imbalance nature of the dataset [6].

4.3. Performance metrics

Many of the parameters can be used while comparing all the techniques and to report their performances including the confusion set matrix, Sensitivity, Specificity, False positive rate and balanced classification rate or even the Matthews Correlation coefficient. A confusion matrix is a table showing all the possible instances or the no. of instances that are classified correctly/ incorrectly in each of the prescribed classes. [6,12] Table represents the confusion matrix of a binary classifier. In the problem of fraud detection, positive means the legitimate transactions and negative represents the fraudulent transactions.

The three parameters discussed here are

Table 1
F1 score table.

Method	Accuracy	F1-Score
Logistic Regression	78.83	.015
Decision Tree	72.66	0.369
Random Forest Classifier	80.16	0.446
Naïve Bayes (Gaussian)	61.5	0.443
Naïve Bayes (Bernoulli)	75.83	0.491
K- Neighbour's Classifier	72.5	0.239
ANN-DL	77.63	0.44

- Specificity:

It is considered as the number of frauds that get predicted into the actual total number of fraud cases as described in equation (2).

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP}) \tag{2}$$

- Sensitivity:

This is understood as the number or the count of legit predictions compared to the sum/total number of legit transactions. But, in fraud detection, the most significant feature is the specificity or fraud detection rate. It is taken as having a value of recall means a lowest financial loss to the company as shown in equation (3).

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \tag{3}$$

- Accuracy:

This is a parameter which gives the overall accuracy [6] of the proposed system. It gives the total number of predictions to the total number of cases considered as shown in Equation (4).

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{TN} + \text{FP} + \text{FN}) \tag{4}$$

Sometimes the correctness of the model can be very misleading as in case of Credit card frauds as the number of fraudulent transactions are less compared to the total sum in whole. This makes the dataset totally imbalanced. Also, selecting the right metric depends on the goal or business objective that we are looking for. Sometimes one strategy may help one achieve customer satisfaction while others might have a higher ability to prevent the financial losses.

4.4. Experimental results

For our understanding and convenience let us take up four models and train and test them using Weka which stands for "Waikato Environment and Knowledge Analysis".[6] It is a workbench, or a platform used for Machine Learning that has the capacity to implement many of the data mining techniques. It can also help apply the various early and preprocessing and sampling techniques. Weka was developed in the Java language in New Zealand by the University of Waikato. [6,12]

In this experiment we have used 0-fold,5-folds,10-folds,15-folds, 20-folds cross validation processes. It is made so to ensure the equal representation of all data as training and test data. Then, the average of these responses is taken to find out the result of a specific parameter as depicted in Fig. 8.

Now, coming to the dataset, Table 2 shows the output result of the Decision Tree algorithm. Table 3 represents the performance of K nearest neighbour while Table 4 illustrates the performance of the Neural network algorithm. And finally, we have Table 5 representing the performance of the Logistic Regression.

Now, let us analyse the graphical representation of the values with 15 folds and 20 folds and compare the result depicted in Figs. 9 & 1 respectively.

Table 2
. This table gives the recorded values of accuracy, sensitivity and specificity at various folded experiments using the Decision Tree algorithm.

2.Results of Decision Trees across different fraud rates			
Folds used	Accuracy measured	Sensitivity found	Specificity found
0	0.94119706	94.03%	76%
5	0.94293	94.68%	71%
10	0.94476	94.70%	74%
15	0.9417	94.94%	65%
20	0.94776	94.77%	75%

Table 3
This table interprets the recorded values of accuracy, sensitivity and specificity at various folded experiments using the K-Nearest Neighbour algorithm.

3. Results of K-Nearest Neighbour across different fraud rates			
Folds used	Accuracy measured	Sensitivity found	Specificity found
0	0.958897945	96.50%	76.90%
5	0.95838	96.70%	74.44%
10	0.95838	96.70%	74.40%
15	0.9585	96.71%	74.50%
20	0.95838	96.70%	74.44%

Table 4
This table represents the recorded values of accuracy, sensitivity and specificity at various folded experiments using the Neural Network algorithm.

4.Results of Neural Network across different fraud rates			
Folds used	Accuracy measured	Sensitivity found	Specificity found
0	0.769688	99.30%	69.70%
5	0.84369	97.70%	44.70%
10	0.93995	95.70%	68.50%
15	0.9372	96.40%	56.31%
20	0.92813	96.61%	54.33%

Table 5
This table accounts the recorded values of accuracy, sensitivity and specificity at various folded experiments using the Logistic Regression algorithm.

Results of Logistic Regression across different fraud rates			
Folds used	Accuracy measured	Sensitivity found	Specificity found
0	0.963198	96.90%	82.00%
5	0.96238	96.80%	80.49%
10	0.963198	96.80%	80.40%
15	0.9624	96.85%	80.62%
20	0.96238	96.85%	80.49%

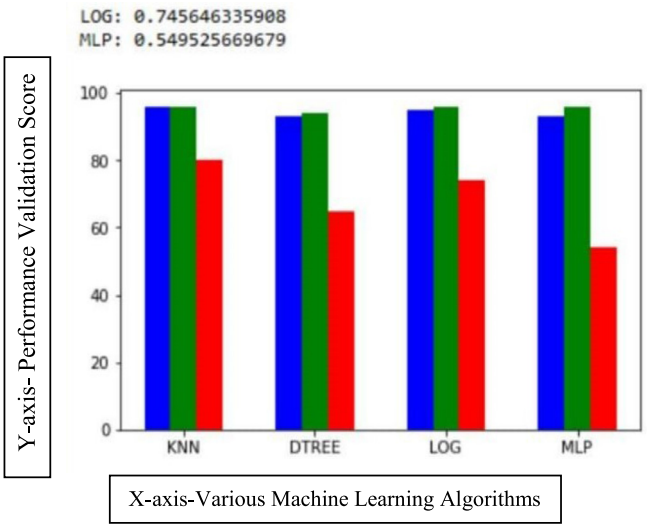

Fig. 9. Graph 1 with 15 folds.

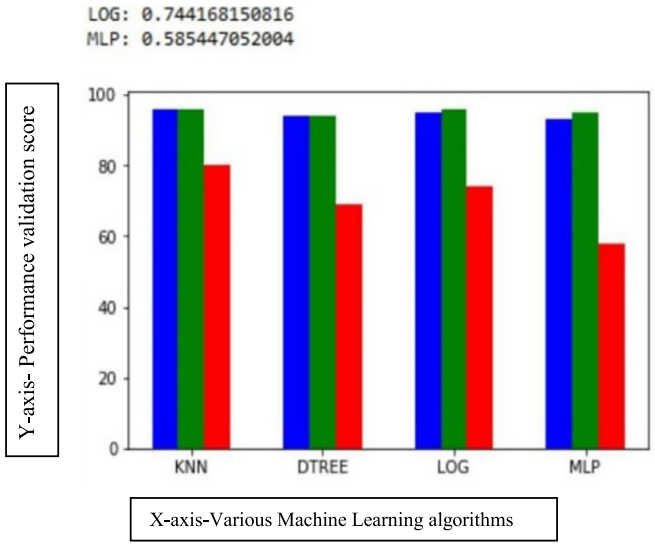

Fig. 10. Graph 2 with 20 folds.

5. Conclusion

Though the above used techniques and algorithms are proven to be efficient and accurate while being implemented individually or in combinations, there could be a chance of having dissimilar and irregular dataset. Hence, some pre-processing and sampling algorithms must be added to the raw dataset to classify it before being subjected to any of the Random Forest, Linear Regression, AdaBoost or ANN techniques used to analyse it. Failure of such models could happen mainly because credit card transactions are very confidential and hence one must take care while developing a model such that it doesn't leak any key data while in the process. Also, the total number of frauds is always lesser than the total data considered hence the variation of the subset is high and may vary at different instances causing a model to fail in a particular situation. The real limitation of any Machine Learning Algorithm is the fact that real-time actions must take place. Or while using predictive analytic methods, the algorithm must be adaptive in nature in order to tackle the imbalance or variations found in huge datasets from various resources. So, there doesn't exist any data mining technique that can be universally better in all cases and therefore studies in the future work must look upon these factors in order to increase the prediction accuracy. The limitation of this paper are as follows on which work can be done in future:

1. The performance of other machine learning algorithm can be checked for credit card fraud detection.
2. The accuracy of XG Boost, Random Forest etc. machine learning algorithm should also be tested more subjectively on other data sets for credit card fraud detection.
3. The performance of other algorithms can also be tested on other data sets of different domains and varied patterns.

Acknowledgments

No financial support was received to perform the research work of this manuscript.

References

[1] Vinod Jain, Mayank Agrawal, Anuj Kumar, "Performance Analysis of Machine Learning Algorithms in Credit Cards Fraud Detection, 2020 at 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2022.
[2] Roopesh Akula in "Fraud identification of credit card using ML techniques at, Int. J. Comput. Artif. Intell. 1 (2) (2020) 31–33.

[3] A. Krishnaiah, P.B. Divakarachari, Automatic Music Mood Classification using Multi-class Support Vector Machine based on Hybrid Spectral Features, 2022.
[4] Gurumurthy Krishnamurthy Arun*, Kaliyappan Venkatachalapathy, Intelligent feature selection with social spider optimisation based Artificial Neural Network Model for Credit card Fraud detection, in: | Arun & Venkatachalapathy, 11, IIOABJ, 2020, pp. 85–91. ||||.
[5] T.G. Nguyen, T.V. Phan, D.T. Hoang, T.N. Nguyen, C. So-In, Efficient SDN-based traffic monitoring in IoT networks with double deep Q-network, in: International conference on computational data and social networks, Springer, Cham, 2020, pp. 26–38.
[6] Xurui Li, Wei Yu, Tianyu Luwang Jianbin Zheng, Xuetao Qiu, Jintao Zhao, Lei Xia Yujiao Li in " Transaction Fraud detection using GRU-Centered Sandwich-structured Model at, in: Proceedings of the 2018 IEEE 22nd International Conference on Computer Supported Cooperative Work in Design, 2022.
[7] K. Yu, L. Lin, M. Alazab, L. Tan, B. Gu, Deep learning-based traffic safety solution for a mixture of autonomous and manual vehicles in a 5G-enabled intelligent transportation system, IEEE Trans. Intell. Transp. Syst. 22 (7) (2020) 4337–4347.
[8] S. Venkata Suryanarayana *, G.N. Balaji, G. Venkateswara Rao in "Machine learning approaches for credit card fraud detection" at, Int. J. Eng. Technol. 7 (2) (2018) 917–920.
[9] Ajeet Singh, Anurag Jain, Adaptive Credit Card Fraud Detection Techniques Based on Feature Selection Method, University Grants Commission (UGC), Delhi, India, 2022 in fellowship research problem at.
[10] Vaishnave Jonnalagadda, Priya Gupta, Eesita Sen in "Credit card fraud detection using Random Forest Algorithm" in Jonnalagadda Vaishnave et al.; International Journal of Advance Research, Ideas and Innovations in Technology, Volume 5, Issue 2.
[11] B.D. Parameshachari, K.M. Keerthi, T.R. Kruthika, A. Melvina, R. Pallavi, K.S. Poonam, Intelligent Human Free Sewage Alerting and Monitoring System, in: 3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC 2021), Atlantis Press, 2021, pp. 480–486.
[12] Youness Abakarim, Mohamed Lahby, Abdelbaki Attioui, "An Efficient Real Time Model For Credit Card Fraud Detection Based On Deep Learning", SITA'18, 2018 October 24–25Rabat, Morocco.
[13] Hassan Najadat, et al., Credit Card Fraud Detection Based on Machine and Deep Learning, 2020 11th International Conference on Information and Communication Systems (ICICS), 2022.
[14] R.K. Dash, T.N. Nguyen, K. Cengiz, A. Sharma, Fine-tuned support vector regression model for stock predictions, Neural Comput. Appl. (2021) 1–15.
[15] Anuruddha Thennakoon, Chee Bhagyani, Sasitha Premadasa, Shalitha Mihiranga, Nuwan Kuruwitaarachchi in "Real-time Credit Card Fraud Detection Using Machine Learning, 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), 2022.
[16] Z. Guo, K. Yu, A. Jolfaei, A.K. Bashir, A.O. Almagrabi, N. Kumar, Fuzzy detection system for rumors through explainable adaptive learning, IEEE Trans. Fuzzy Syst. 29 (12) (2021) 3650–3664.
[17] Aadhya Kaula, Mihika Chhabraa, Prachi Sachdevaa, Rachna Jaina, Preeti Nagratha in "Credit Card Fraud Detection Using Different ML and DL Techniques" in their research works.
[18] G.B. Rajendran, U.M. Kumarasamy, C. Zarro, P.B. Divakarachari, S.L. Ullo, Land-use and land-cover classification using a human group-based particle swarm optimization algorithm with an LSTM Classifier on hybrid pre-processing remote-sensing images, Remote Sensing 12 (24) (2020) 4135.
[19] Hamzah Ali Shukur, Sefer Kurnaz in, Credit card fraud detection using machine learning methodology, Int. J. Comput. Sci. Mob. Comput. 8 (3) (2019) 257–260 Vol.Issuepg.
[20] Z. Guo, K. Yu, Y. Li, G. Srivastava, J.C.W. Lin, Deep learning-embedded social internet of things for ambiguity-aware social recommendations, IEEE Trans. Netw. Sci. Eng. (2021).
[21] Naoufal Rtayli*, Nourddine Enneya in "Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization, J. Inf. Secur. Appl. 55 (2020) 102596.
[22] D.L. Vu, T.K. Nguyen, T.V. Nguyen, T.N. Nguyen, F. Massacci, P.H. Phung, A convolutional transformation network for malware classification, in: 2019 6th NAFOSTED conference on information and computer science (NICS), IEEE, 2019, pp. 234–239.
[23] Apapan Pumsirirat, Liu Yan, Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine" in (IJACSA), Int. J. Adv. Comput. Sci. Appl. 9 (1) (2018).