Penerbit Akademia Baru

Journal of Advanced Research in Computing and Applications

Journal homepage:
https://akademiabaru.com/submit/index.php/arca/index
ISSN: 2462-1927

A Comparative Performance Analysis of Several Machine Learning Classifiers on The Credit Card Data

Nor Aishah Ahad[1,*], Friday Zinzendoff Okwonu[2], Yik Siong Pang[3], Olimjon Shukurovic Sharipov[4]

[1] Institute of Strategic Industrial Decision Modelling, School of Quantitative Sciences, College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia
[2] Department of Mathematics, Faculty of Science, Delta State University, P.M.B.1, Abraka, Nigeria
[3] School of Quantitative Sciences, College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia
[4] National University of Uzbekistan named after Mirzo Ulugbek, 100174 University Street 4, Almazar District, Tashkent Uzbekistan

ARTICLE INFO

Article history:
Received 1 November 2024
Received in revised form 1 December 2024
Accepted 11 December 2024
Available online 31 December 2024

Keywords:
Credit card; fraud; classification; mahalanobis distance; accuracy; efficiency

ABSTRACT

As customers tendencies learn rapidly about online transactions due to the outbreak of the Covid-19 pandemic, fraudulent transactions that mimic legitimate transactions will continue to increase unabated. Over the years, fraudulent transactions have been committed but the supervised and unsupervised machine learning classifiers and fraud detection techniques could not stop the fraud from occurring at the onset. However, fraudulent cases are detected after post-occurrence thereby recognizing these techniques as non-smart detection models because fraud could only be detected using past data sets. Various fraud detection or classifier techniques using machine learning techniques such as decision trees, random forests, k-means, and neural networks have been studied. Also, the possibility of overfitting due to the branch growth associated with the decision trees is a classifier performance problem. The application of the probability axiomatic metric (PAM) to determine the accuracy of the models is a classification evaluation problem because of the relatively moderate error rate. To overcome the above shortfalls, we proposed a threshold Mahalanobis distance (TMD) classifier and adopted BETH as a performance evaluation metric. The result revealed that TMD is comparable to neural network (NN), k-means, isolation forest, XGBoost, and random forest based on the credit card data set. The analysis showed that TMD and NN have higher efficiency compared to other classifiers. This study demonstrated that the TMD classifier is unique and could be applied to extract legitimate and illegitimate transactions from customer transaction data sets. The study concludes that TMD is robust and comparable to NN and the classification errors associated with BETH are minimal compared to PAM for both TMD and NN.

1. Introduction

The outbreak of the Covid-19 pandemic and the subsequent lockdown increased the ways people interact and purchase vital everyday household items online. As a result of the lockdown, many people migrated to online purchases which have consequences due to privacy information theft and

[*] *Corresponding author.*
E-mail address: aishah@uum.edu.my (Nor Aishah Ahad)

https://doi.org/10.37934/arca.37.1.5064

fraudulent purchase and request. People have been defrauded through online transactions and romance scams via credit card information theft.

Fraud classification (FC) is simply the process of separating legitimate and illegitimate transactions into different groups. Fraud classification is an unequal sample size problem because one group has majority items (legitimate transactions) and the other minority group (illegitimate transactions). As such FC problems consist of heavily skewed data sets. Fraudulent transactions are increasing as the rate of online transactions increases. The fraud classification problem is naturally skewed to the left (fraud) and legitimate transactions is skewed to the right (non-fraud). Credit card (CC) usage is increasing drastically, especially during the outbreak of Covid-19 and subsequent lockdown, surge in online and physical usage of CC for various degrees of transactions increased [1]. CC fraud is classified as financial fraud and other classes of fraud could be mentioned in healthcare [2], insurance [3], telecommunication and bankruptcy [4] to mention a few. Fraud classification models could be classified into the following machine learning classifiers, supervised [5], unsupervised [6], and hybrid [7].

Fraud often takes place when the fraud prevention mechanism is disrupted. Therefore, fraud detection becomes necessary to determine the antics of the fraudster to prevent future occurrences [8]. Fraud investigation consists of extremely large data sets due to the volume of online transactions taking place globally [9]. Therefore, detection and prevention are evolving because fraudsters continue to counter prevention mechanisms to remain active in criminality. Fraud can easily be detected when customers assign their cards to unique purchase requests, say using a card XZ for booking airline tickets, card YK for car purchases, and card DG for grocery only to mention a few. In this scenario, if card DG is used for the airline ticket, the card owner may reasonably assume that the card has been compromised. This is a simple way to detect a compromised credit card [8]. Another means to detect fraud is sudden destination and location variation in online transactions.

Credit card data is naturally n>p problem which originates from ungrouped data by natural occurrence, this is true because all the transactions are assumed to be legitimate until classifiers are applied to perform group separation. When classifiers are applied, the output often appears as $G_1(n_1) > G_2(n_2)$,, where G_k and $n_k, k = 1,2$ are the respective groups and sample sizes. In the unsupervised classification domain, which allows classifiers to separate objects into groups, the possibility of unequal sample size frequently occurs, and very impossible to have an equal sample size. Therefore, the process of transforming an unbalanced dataset into a balanced one negates the principles of unsupervised classifiers. It is possible to balance unequal data set for supervised classifiers because the structure of the data set is known. In most cases, the unsupervised classifiers classify ungrouped objects into the desired groups. Concurring with this concept, it is reasonable to conclude that the outcome of supervised classifiers often is unequal based on the sample size, but this concept is hidden in the classification domain, rather we focused mainly on the probability of correct classification instead of the sample size or number of the objects reclassified. Therefore, the output of the supervised classifier is reclassified and the sample size for the groups is unbalanced in most cases. Therefore, by classification, the supervised classifiers do confirmation classification while the unsupervised classifier separates the ungrouped objects into the respective groups based on the classifier decision benchmark. This implies that unsupervised classifiers are efficient and smart in grouping objects.

The existing fraud detection techniques often depend on supervised or unsupervised classifiers based on the nature of the data set. In which direct applications of the classifiers on the data set are considered. Therefore, this paper focused on a hybrid procedure in which the unclassed data set is classed by the unsupervised model, creating classes for the data set. This class creation allows the supervised classifiers to reclassify and confirm objects to the correct class, which is unsure in most existing classifiers. These hybrid classifiers enable data partition and reconfirmation of objects that are lacking in existing classifiers. This process can be described as stage partition and reconfirmation technique.

2. Literature Review

It is vital to observe that among the various fraud detection models, none was able to detect fraud at the onset rather detection was done post-fraud based on data comparison or grouping using the various classifiers [10]. Previous studies revealed that an attempt has been made to develop a prediction model that could detect and prevent fraud. However, these precognitive prediction models still depend on previous data which invariably revealed that purchases or transactions could be classified as legitimate or fraudulent transactions after the process has been completed [11].

As online transactions are increasing, fraud associated with them will also increase [12]. The significant increase in fraudulent activities can be attributed to advancements in technology and the prevalence of online transactions, leading to significant financial losses [13]. As such, fraudulent purchases could be classified as legitimate (non-fraud) or illegitimate (fraud). Stealing associated with credit cards is increasing globally and exponentially [8,14]. As a result, various researchers and financial organizations are seeking efficient mechanisms to protect and decrease the rate of fraudulent online purchases based on the applications of efficient supervised, unsupervised, and hybrid machine learning algorithms. The artificial neural network (ANN) [15-17], unsupervised learning neural network [1], decision trees [18] and other supervised learning techniques have been deployed to detect fraudulent transactions [19-26]. The support vector machine [27], Adaboost, frequent item mining [12,28], isolation forest [29], and logistic regression [30], machine learning and deep learning [13,31-32], profiling and neural network [33] and artificial immune system (AIS) [34], Glass algorithm [35] and several other supervised and unsupervised classifiers have been applied to detect fraud [3,36-38]. Genetic algorithm and scatter search algorithm [39-41] have also been applied to classify transactions as legitimate or illegitimate. Several methods have been advanced and compared based on performance analysis and the outcome revealed that some of these machine learning or hybrid methods performed comparably [42-43]. The performance of the machine learning algorithms depends on the evaluation metric as such underperformance, overfitting, accuracy, and efficiency are associated with classification problems. To remedy the shortfalls, we proposed an unsupervised threshold classification model based on the Mahalanobis distance and adopted the BETH evaluation metric [44] to determine performance accuracy, efficiency and to compare the associated performance errors of PAM and BETH.

This article is organized as follows. The next section contains the methodology. In this section, we described the threshold classification method using the Mahalanobis concept, the BETH metric was also mentioned followed by the performance errors based on the PAM and BETH evaluation metric. The data set description, discussion, and analysis of various machine learning classifiers on the credit card data set are presented in the following section. This section includes the performance analysis of five previous studies on credit card data classifications. The concluding remark is presented in the last section.

3. Materials and Methods

In recent years, prediction, classification, and other machine-learning techniques have been applied to detect and analyze fraud in different categories [45-46]. Generally, a fraudulent classification problem (FCP) is a problem of unequal sample size. This is a classical problem that could be considered before selecting the most suitable classifiers to apply. In the statistical concept, the fraudulent transactions would reveal irregular data patterns compared to the legitimate transactions that would reveal regular patterns. In this case, fraudulent transactions are considered a noise or influential observation. Also, fraudulent transactions may show masking effects in such a way that the detective would be unable to identify that the transaction was fraudulent. Therefore, for group classification problems, legitimate transactions tend to contain robust information that aligned with the central limit theorem thereby resulting in better classification results than the fraudulent transactions group with less information which may also lead to a high misclassification rate [47].

3.1 Threshold Mahalanobis Distance (TMD)

Let X_i $(i = 1,2, \dots, n)$ be $n \times p$ data set which is assumed to be the legitimate transactions. Then at the onset, all transactions are assumed to be legitimate which is categorized into one group. However, as credit card owners initiate likely fraud attempts, all the information in X_i are evaluated to detect whether illegitimate transactions have occurred. In this case, X_i is assumed to contain legitimate and illegitimate transactions. This problem can be modeled as follows.

$$\emptyset = \left(X_i - \left(\frac{\sum_{i=1}^{n} X_i}{n} \right) \right) \tag{1}$$

From Eq. (1) we obtain the following Eq. (2) and (3)

$$\Delta = \frac{\left(X_i - \left(\frac{\sum_{i=1}^{n} X_i}{n} \right) \right) \left(X_i - \left(\frac{\sum_{i=1}^{n} X_i}{n} \right) \right)^T}{n-1} = \frac{\emptyset\emptyset^T}{n-1} \tag{2}$$

and

$$\partial = inv(\Delta) \tag{3}$$

Therefore,

$$\begin{aligned} \pi^2 &= (\emptyset \times \partial \times \emptyset^T), \\ \pi &= \sqrt{(\emptyset \times \partial \times \emptyset^T)} \end{aligned} \tag{4}$$

Eq. (4) transforms the data set such that a threshold is required to detect if a fraudulent transaction is contained in X_i. To achieve this, $\chi_2^2(0.01)$ is applied as the threshold value. Hence the TMD classifier is described as

$$\delta = \begin{cases} \pi < \chi_2^2(0.01), \ nf \\ \pi > \chi_2^2(0.01), f \end{cases}, \tag{5}$$

where nf denotes nonfraudulent or legitimate transactions and f is a fraudulent or illegitimate transaction. Eq. (5) can be described simply as assigning legitimate transaction if Eq. (6) is true

$$\pi < \chi_2^2(0.01) \tag{6}$$

otherwise, assign illegitimate transaction if Eq. (7) is valid

$$\pi > \chi_2^2(0.01) \tag{7}$$

Therefore, based on Eq. (6) and (7), the TMD method could classify the transaction types into two groups. This concept may be similar to the outlier detection concept enunciated by Murad and Pinkas [48] and the concept of purchase behavioral changes discussed in Cortes *et al.*, [49] and Fawcett and Provost [50]. The TMD scheme may be similar to a workflow system in Gadi *et al.*, [34] by scoring.

3.2 Evaluation Performance

Faraji [43] observed that the group with the larger sample sizes would have higher classification accuracy than the minority group. Due to the unequal sample sizes from the two classes, that is, legitimate and illegitimate transactions, as such Faraji [43] observed that accuracy as a performance evaluation criterion may result in poor performance analysis. Therefore, several researchers including Faraji [43] suggested that precision and recall based on the confusion matrix (CM) be applied as the evaluation metric. However, the use of accuracy as a performance metric [29,51] has gained attention in the classification and prediction literature [52]. In this study, we suggest that group classification strictly depends on the weight of the data set, we also adhere to classification information from CM. To solve the associated overfitting problem, we adopted accuracy and applied the benchmark threshold method (BETH) to enhance the classifiers' performance. To evaluate the classifier's performance error, we compared the PAM with the BETH metric.

Overfitting is a major performance problem in the classification domain. It is also an associated problem with decision trees [51]. In many n>p, p>n problems, the evaluation metric is mainly based on PAM, that is,

$$\omega = \alpha - \mu, \tag{7}$$

where α denotes the BETH value and $\rho = 1$,

$$\varepsilon = \rho - \mu. \tag{8}$$

Eq. (9) and (10) describe BETH and PAM efficiencies,

$$eff(\beta) = \frac{\mu}{\alpha} \times 100 \tag{9}$$

$$eff(\gamma) = \frac{\mu}{\rho} \times 100 \tag{10}$$

For comparative analysis of previous studies on the credit card data set, this study will adopt the reporting of the performance analysis of previous studies and compare the performance analysis with the benchmark threshold method (BETH). The BETH solves overfitting problems which are common when accuracy is applied to analyze performance. Therefore, the comparison would reveal the robustness of the various studies based on the classifiers.

3.3 Data Set

The credit card data is well utilized data set in the classification domain. The credit card transactions data set was taken from the Kaggle website (https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud?resource=download) and has been analyzed by different researchers with different classifiers. It is an unequal sample size data set because the fraudulent transaction is in the minority and the legitimate transaction is majority respectively. Therefore, the data set is highly skewed. This data set consists of 284,807 transactions (Sample size, that is $n = 284{,}807, p = 28$) for two days period by credit card owners in Europe in September 2013 obtained by Dal Pozzolo *et al*. [53]. The $p = 28$ variables are based on the principal component analysis which is restricted due to confidentiality. However, in this study the amount was the focused of the fraud classification. Previous studies have identified 492 illegitimate transactions and 284315 legitimate, respectively. In this study, we adopt the procedure explained in Figure 1 by Carcillo *et al*. [10] and apply the untransformed transaction amount to perform classification in order to detect legitimate and illegitimate transactions by the proposed TMD classifier. In general, a compromised credit card is mainly associated with the amount stolen or expended from the credit card.

4. Results and Discussions

Before we proceed, let discuss the following metric of measurement or performance as follows. Accuracy simply implies the exactness of the predicted value by the classifier. The idea of recall is that when the value of recall is above 0.5, it means that the true positive is well identified. The contrary is true for a recall value less than 0.5. On the other hand, precision measures the exact positive prediction. It is arguably sound to note that a higher precision generates a lower recall, and a higher recall generates a lower precision. Recall and precision can easily be manipulated via parameter tuning of the model [29] whereas accuracy is not manipulatable via parameter tuning. Accuracy gives better performance measures than recall and precision, hence the derivation of the BETH metric was developed from accuracy to obtain robust performance. The rest of this section focused on the comparative performance of different classifiers based on five previous research on the credit card data set. We also compared the efficiency and the metric errors.

Fraud is ravaging the global space due to the rapid growth of online transactions. As online transactions grow exponentially so also fraudulent transactions grow exponentially. To solve these problems several fraud detection techniques have been proposed to detect fraud in real time but unfortunately, all classifiers are based on historical data to detect fraudulent transactions. The TMD proposed aligns with the existing classifiers but is hybrid in nature. Therefore, it can be applied to solve fraud problems in real time when the data is available. It has a wide range of applications in commerce, banking, telecommunication, romance scam, insurance scam and round tripping etc. But in this study, the focus is on credit card fraud which is an aspect of commercial fraud.

Table 1 is based on the results reported in Table 3 by Faraji [43]. The TMD was able to detect 284317 (99.83%) legitimate transactions and 490 (0.17%) fraudulent transactions. The TMD classifier revealed comparable performance with decision trees, XGBoost and random forest and outperformed the logistic regression and KNN. The average classifier efficiency is in Table 1 for BETH (0.9591) and PAM (0.9397). Figure 1a contains the classifiers' error with TMD with minimum error compared to other classifiers and Figure 1b compares the efficiency of the different classifiers.

Table 1
Comparative performance analysis of credit card data (Table 3) in Faraji [43]

Classifier	Accuracy(μ)	BETH(α)	$eff(\beta)$	$eff(\gamma)$	ERROR(ω)	BAS ERROR(ε)
Decision trees	0.9900	0.9951	0.9949	0.9900	0.0051	0.0100
XGBoost	0.9900	0.9951	0.9949	0.9900	0.0051	0.0100
Random forest	0.9900	0.9951	0.9949	0.9900	0.0051	0.0100
Logistic regression	0.8100	0.9377	0.8638	0.8100	0.1277	0.1900
KNN	0.8600	0.9482	0.9070	0.8600	0.0882	0.1400
TMD	0.9983	0.9992	0.9991	0.9983	0.0009	0.0017

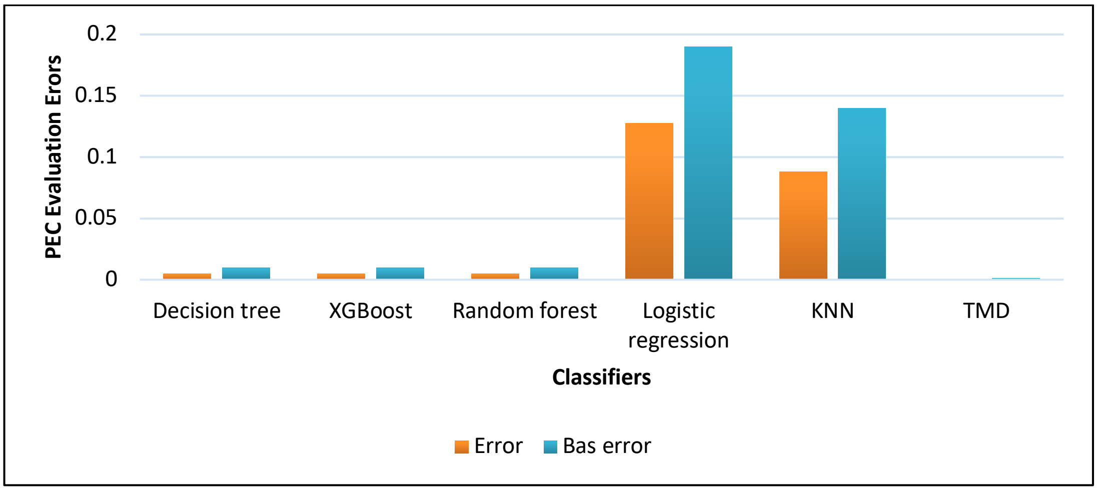

Fig. 1a. Comparative error analysis for credit card data using different classifiers

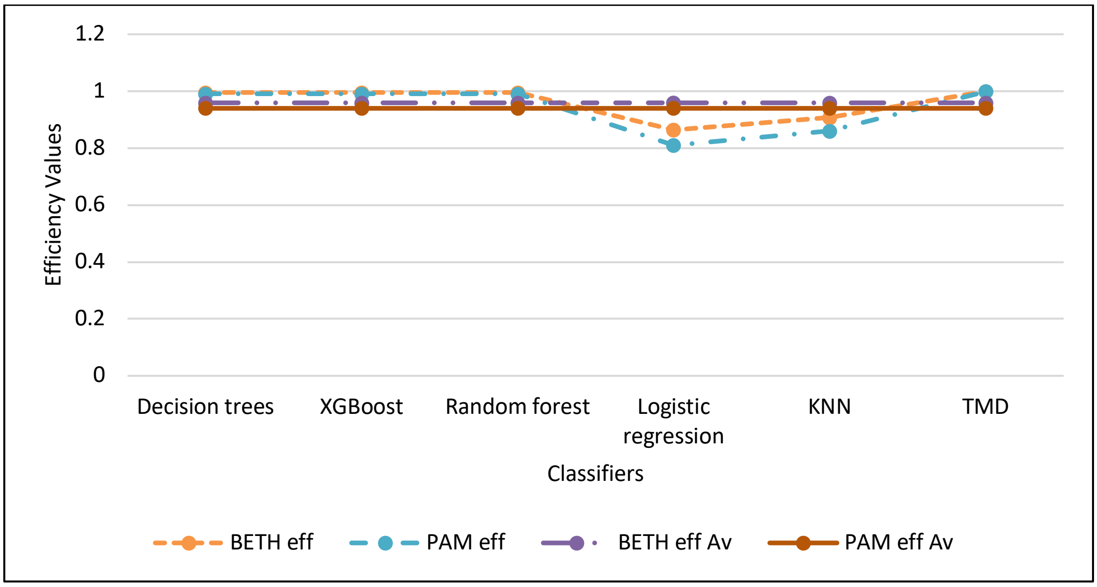

Fig. 1b. Comparative classifiers efficiency

Further, we shall consider the performance analysis of several classifiers discussed in Mishra and Ghorpade [47] in Table A with the TMD classifier for the credit card data. The result in Table 2 demonstrated that TMD outperformed the other methods. Figure 2a shows that TMD has minimum error compared to other classifiers and Figure 2b contains the classifiers efficiency values. In Figure 2b, we observed patterns different from Figure 1b.

Table 2
Comparative performance analysis of credit card data (Table A) in Mishra and Ghorpade [47]

Classifier	Accuracy(μ)	BETH(α)	$eff(\beta)$	$eff(\gamma)$	ERROR(ω)	BAS ERROR(ε)
Stacked	0.9570	0.9803	0.9762	0.9570	0.0233	0.0430
Logistic regression	0.8909	0.9567	0.9312	0.8909	0.0658	0.1091
Gradient boosted trees	0.9620	0.9824	0.9792	0.9620	0.0204	0.0380
Random forest	0.9738	0.9876	0.9860	0.9738	0.0138	0.0262
SVM	0.9411	0.9739	0.9663	0.9411	0.033	0.0589
TMD	0.9983	0.9992	0.9991	0.9983	0.0009	0.0017

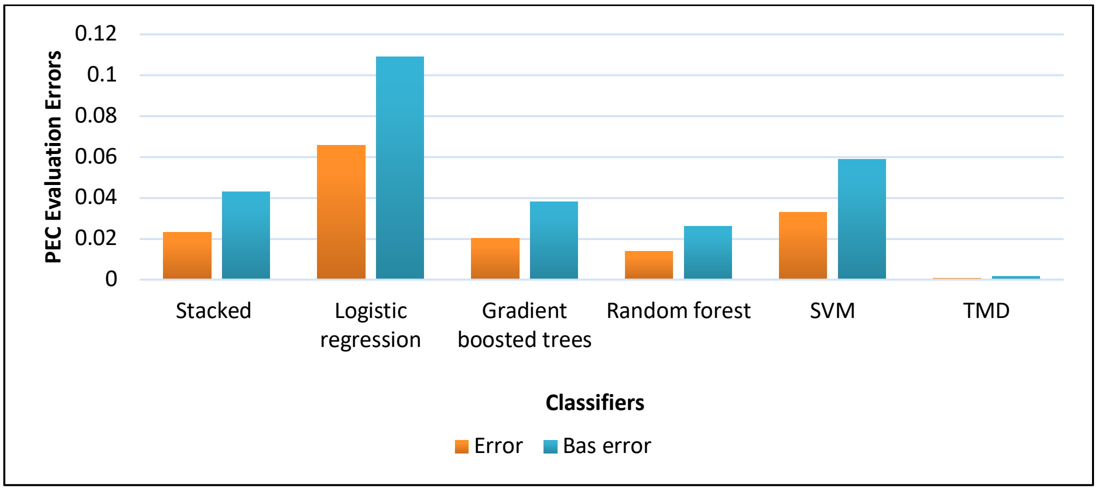

Fig. 2a. Comparative error analysis for credit cards data using different classifiers

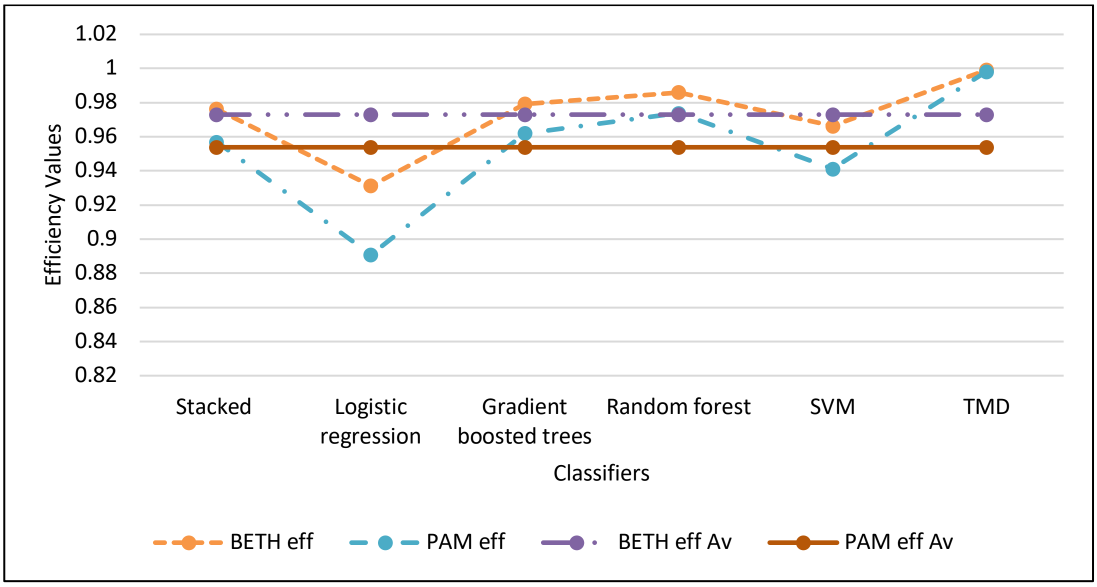

Fig. 2b. Comparative classifiers efficiency

Table 3 contains other results by Gowda [12] on different classifier's performance on credit card data. In Table 3, the K means and TMD performed equally followed by local outlier factor and isolation forest. In Figure 3a, K means and TMD has a smaller error rate than logistic regression, random forest, and SVM, while Figure 3b revealed the classifiers' efficiencies. Figure 3b demonstrated a unique pattern from Figure 1b and Figure 2b.

Table 3
Comparative performance analysis of credit card data (Table 1) in Gowda [12]

Classifier	Accuracy(μ)	BETH(α)	$eff(\beta)$	$eff(\gamma)$	ERROR(ω)	BAS ERROR(ε)
Local outlier factor	0.9967	0.9984	0.9983	0.9967	0.00165	0.0033
Logistic regression	0.9703	0.9852	0.9849	0.9703	0.01485	0.0297
K means	0.9982	0.9991	0.9991	0.9982	0.0009	0.0018
Random forest	0.9715	0.9856	0.9857	0.9715	0.01425	0.0285
Isolation forest	0.9967	0.9984	0.9983	0.9967	0.00165	0.0033
SVM	0.9849	0.9925	0.9923	0.9849	0.00755	0.0151
TMD	0.9983	0.9992	0.9991	0.9983	0.0009	0.0017

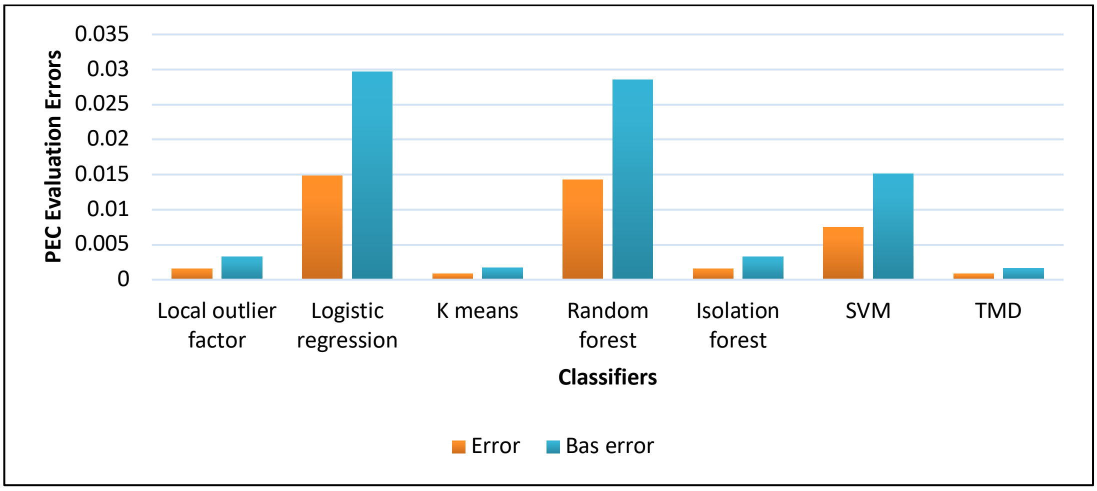

Fig. 3a. Comparative error analysis for credit cards data using different classifiers

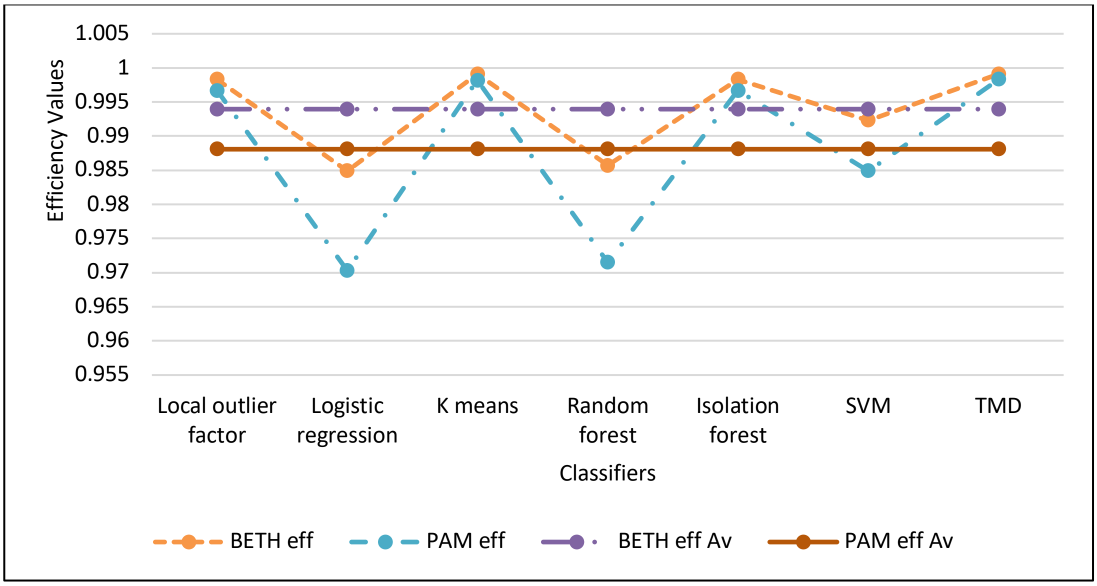

Fig. 3b. Comparative classifiers efficiency

Table 4 contains the comparative performance of various classifiers reported in Table 3 by Rai and Dwivedi [1]. The analysis revealed that the Neural network outperformed other classifiers. In Figure 4a, the Neural network classifier has the smallest error rate followed by TMD and Figure 4b has a unique shape indicating a different performance pattern from Figure 1b to Figure 3b.

Table 4
Comparative performance analysis of credit card data (Table 3) in Rai and Dwivedi [1]

Classifier	Accuracy(μ)	BETH(α)	$eff(\beta)$	$eff(\gamma)$	ERROR(ω)	BAS ERROR(ε)
Local outlier factor	0.9800	0.9904	0.9895	0.9800	0.0104	0.0200
Neural network	0.9987	0.9994	0.9993	0.9987	0.0007	0.0013
k-means clustering	0.9975	0.9986	0.9989	0.9975	0.0011	0.0025
Auto encoder	0.9700	0.9859	0.9839	0.9700	0.0159	0.0300
Isolation forest	0.9800	0.9904	0.9895	0.9800	0.0104	0.0200
TMD	0.9983	0.9992	0.9991	0.9983	0.0009	0.0017

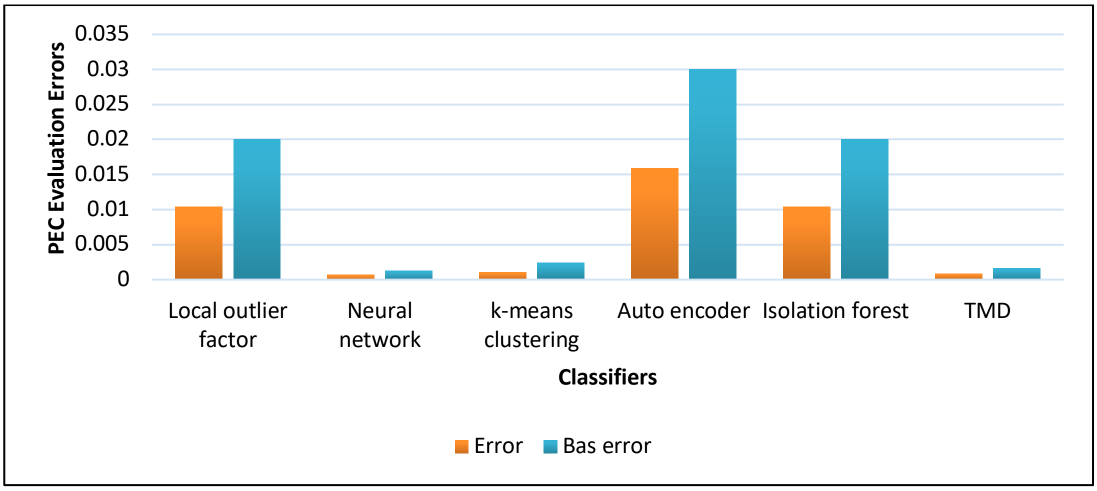

Fig. 4a. Comparative error analysis for credit cards data using different classifiers

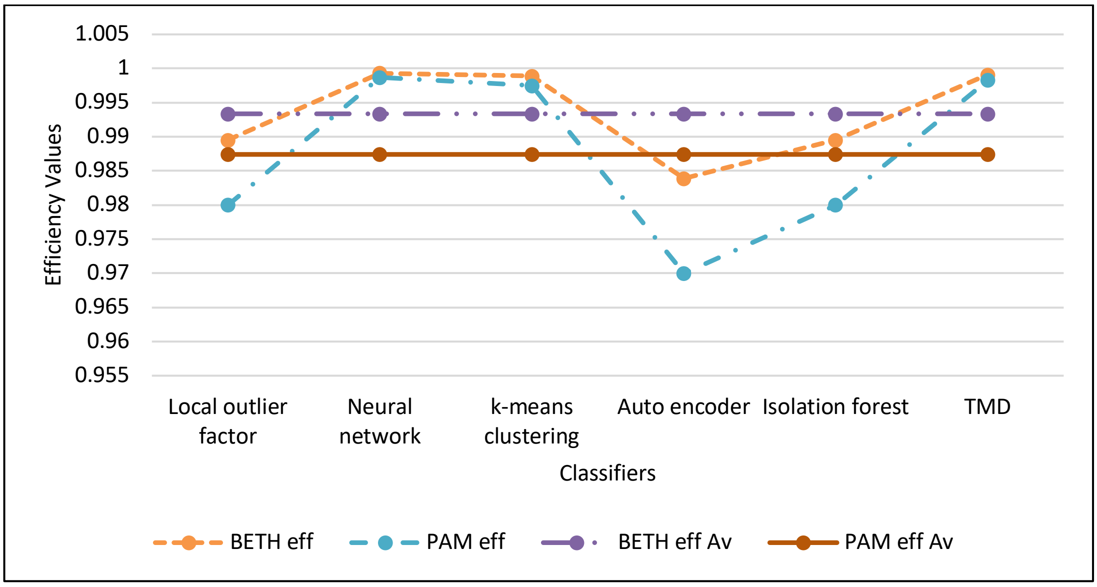

Fig. 4b. Comparative classifiers efficiency

Table 5 contains the performance of four classifiers reported in Table 1 by Khare and Yunus Sait [35]. The result showed that the TMD performed better followed by random forest and logistic regression. Figure 5a revealed that TMD has a minimum error rate compared to other classifies and Figure 5b demonstrated various patterns different from that of Figure 1b to Figure 4b. It showed that some of the classifiers performed above the average values.

Table 5
Comparative performance analysis of credit card data (Table 1) in Khare and Yunus Sait [35]

Classifier	Accuracy(μ)	BETH(α)	$eff(\beta)$	$eff(\gamma)$	ERROR(ω)	BAS ERROR(ε)
Logistic regression	0.9770	0.9890	0.9879	0.9770	0.0120	0.0230
Random forest	0.9860	0.9932	0.9928	0.9860	0.0072	0.0140
SVM	0.9750	0.9881	0.9867	0.9750	0.0131	0.0250
Decision trees	0.9550	0.9795	0.9750	0.9550	0.0245	0.0450
TMD	0.9983	0.9992	0.9991	0.9983	0.0009	0.0017

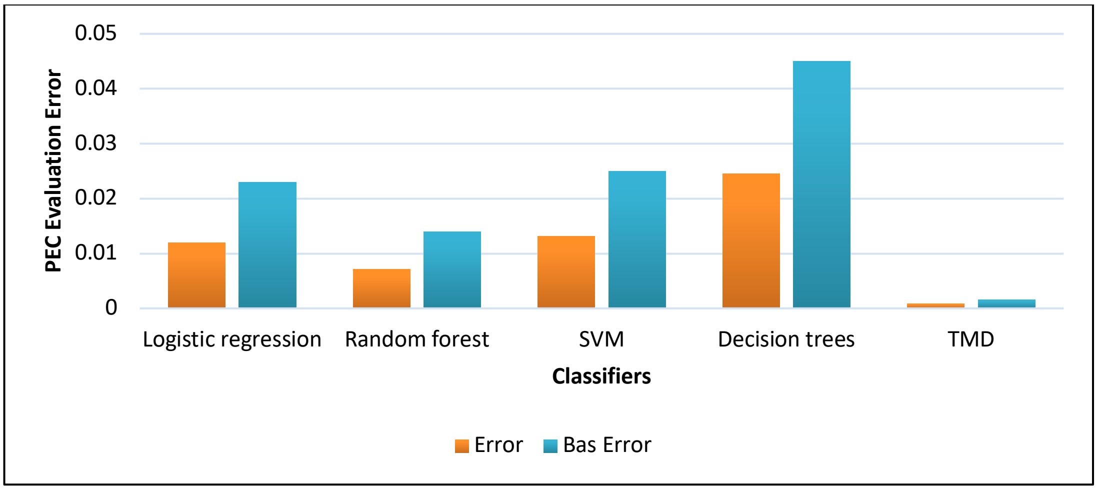

Fig. 5a. Comparative error analysis for credit cards data using different classifiers

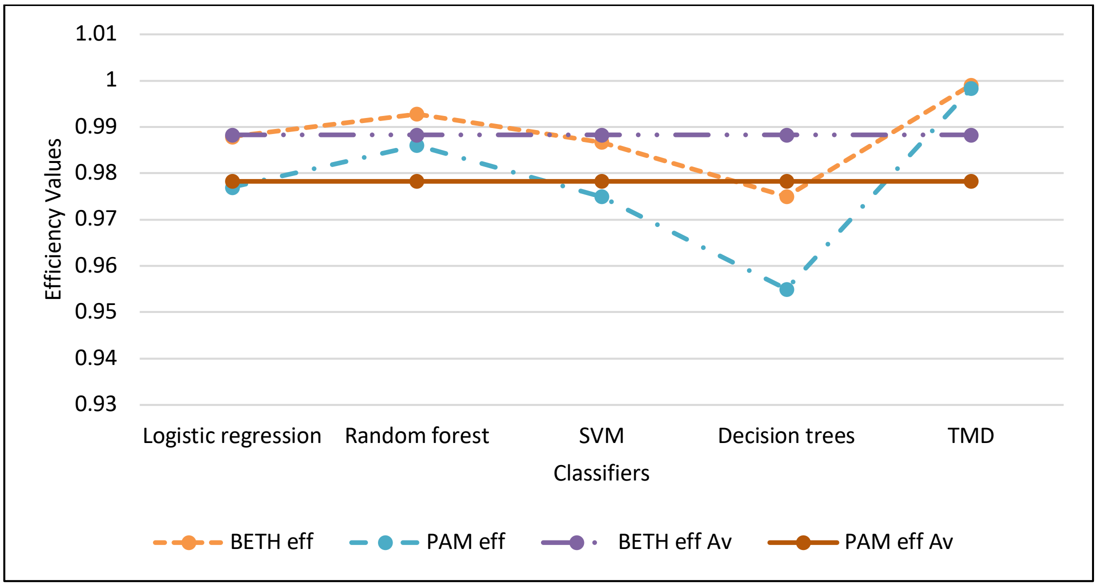

Fig. 5b. Comparative classifiers efficiency

Based on five previous studies on the credit card data set using different machine learning classifier, the neural network (NN) 99.87% [1] was superior to other classifiers followed by K means 99.82% [12], K mean clustering 99.75% [1], isolation forest, local outlier factor 99.67% [12], decision trees, XGBoost and random forest 99.00% [43] while KNN 86% and logistic regression 81% [43] are the least performed classifiers from the previous studied reported in this paper. Like the neural network classifier, the TMD method was able to detect 99.83% legitimate transactions and 0.17% fraudulent transactions. Therefore, based on this study, the performance difference between NN and TMD is 0.04%. Hence the performance is comparable for the credit card data set. The comparative error analysis indicates that for BETH and PAM evaluation metrics, the NN has 0.07% for BETH and 0.13% for PAM followed by TMD with 0.09% for BETH and 0.17% for PAM respectively. Therefore, NN has the smallest error rate followed by TMD. From the analysis, the BETH metric showed better efficiency than PAM, though the different classifiers demonstrated different performance patterns.

From the above comparative analysis, reeling on the different classifier's performance validates the data dependency theory which states that the performance of any classifier depends on the nature of the data set. This validation corroborates the effects and susceptibility of the classifier to outliers and class-unbalanced sample sizes. It is vital to state that an unbalanced data set and outliers enhance the underperformance of some classifiers while other classifiers are robust against outliers and unequal sample sizes. Therefore, these are some critical reasons for some classifiers outperforming other classifiers which was evident in this comparative analysis.

5. Conclusions

This study investigated different supervised and unsupervised machine learning classification models for fraud classifications using credit card data. The study revealed different performance analyses for the different classifiers. The outcome of this study based on five previous research demonstrated that the neural network outperformed other classifiers followed by the TMD. The NN and TMD error rate based on BETH is minimal compared to the error rate due to PAM. This study affirmed that the various machine learning classifiers indicated varying patterns in the credit card data set. Therefore, this study concludes that TMD is very suitable to do fraud classification.

Acknowledgement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

[1] Rai, Arun Kumar, and Rajendra Kumar Dwivedi. "Fraud detection in credit card data using unsupervised machine learning based scheme." In *2020 international conference on electronics and sustainable communication systems (ICESC)*, pp. 421-426. IEEE, 2020. https://doi.org/10.1109/ICESC48915.2020.9155615

[2] Li, Jing, Kuei-Ying Huang, Jionghua Jin, and Jianjun Shi. "A survey on statistical methods for health care fraud detection." *Health care management science* 11 (2008): 275-287. https://doi.org/10.1007/s10729-007-9045-4

[3] Gomes, Chamal, Zhuo Jin, and Hailiang Yang. "Insurance fraud detection with unsupervised deep learning." *Journal of Risk and Insurance* 88, no. 3 (2021): 591-624. https://doi.org/10.1111/jori.12359

[4] Bahnsen, Alejandro Correa, Aleksandar Stojanovic, Djamila Aouada, and Björn Ottersten. "Cost sensitive credit card fraud detection using Bayes minimum risk." In *2013 12th international conference on machine learning and applications*, vol. 1, pp. 333-338. IEEE, 2013. https://doi.org/10.1109/ICMLA.2013.68

[5] Khatri, Samidha, Aishwarya Arora, and Arun Prakash Agrawal. "Supervised machine learning algorithms for credit card fraud detection: a comparison." In *2020 10th international conference on cloud computing, data science & engineering (confluence)*, pp. 680-683. IEEE, 2020. https://doi.org/10.1109/Confluence47617.2020.9057851

[6] Van Engelen, Jesper E., and Holger H. Hoos. "A survey on semi-supervised learning." *Machine learning* 109, no. 2 (2020): 373-440. https://doi.org/10.1007/s10994-019-05855-6

[7] Srivastava, Nitish, and Ruslan Salakhutdinov. "Multimodal Learning with Deep Boltzmann Machines." *Journal of Machine Learning Research* 15, no. 84 (2014): 2949-2980.

[8] Bolton, Richard J., and David J. Hand. "Statistical fraud detection: A review." *Statistical Science* 17, no. 3 (2002): 235-255. https://doi.org/10.1214/ss/1042727940

[9] Leonard, Kevin J. "Detecting credit card fraud using expert systems." *Computers & Industrial Engineering* 25, no. 1-4 (1993): 103-106. https://doi.org/10.1016/0360-8352(93)90231-L

[10] Carcillo, Fabrizio, Yann-Aël Le Borgne, Olivier Caelen, Yacine Kessaci, Frédéric Oblé, and Gianluca Bontempi. "Combining unsupervised and supervised learning in credit card fraud detection." *Information Sciences* 557 (2021): 317-331. https://doi.org/10.1016/j.ins.2019.05.042

[11] Phua, Clifton, Damminda Alahakoon, and Vincent Lee. "Minority report in fraud detection: classification of skewed data." *Acm Sigkdd Explorations Newsletter* 6, no. 1 (2004): 50-59. https://doi.org/10.1145/1007730.1007738

[12] Gowda, Vikas Thammanna. "Credit Card Fraud detection using Supervised and Unsupervised Learning." *AIRCC Punlishing Corporation* (2021). https://doi.org/10.5121/csit.2021.111107

[13] Maheshwari, Vikash Chander, Nurul Aida Osman, and Norshakirah Aziz. "A Hybrid Approach Adopted for Credit Card Fraud Detection Based on Deep Neural Networks and Attention Mechanism." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 32, no. 1 (2023): 315-331. https://doi.org/10.37934/araset.32.1.315331

[14] Donepudi, Praveen Kumar. "Technology growth in shipping industry: an overview." *American Journal of Trade and Policy* 1, no. 3 (2014): 137-142. https://doi.org/10.18034/ajtp.v1i3.503

[15] Kumar, Jitendra, Ashutosh Kumar Singh, and Anand Mohan. "Resource-efficient load-balancing framework for cloud data center networks." *ETRI Journal* 43, no. 1 (2021): 53-63. https://doi.org/10.4218/etrij.2019-0294

[16] Dubey, Saurabh C., Ketan S. Mundhe, and Aditya A. Kadam. "Credit card fraud detection using artificial neural network and backpropagation." In *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 268-273. IEEE, 2020. https://doi.org/10.1109/ICICCS48265.2020.9120957

[17] Khan, Azeem Ush Shan, Nadeem Akhtar, and Mohammad Naved Qureshi. "Real-time credit-card fraud detection using artificial neural network tuned by simulated annealing algorithm." In *Proceedings of International Conference on Recent Trends In Information, Telecommunication and Computing, ITC*, pp. 113-121. 2014.

[18] Save, Prajal, Pranali Tiwarekar, Ketan N. Jain, and Neha Mahyavanshi. "A novel idea for credit card fraud detection using decision tree." *International Journal of Computer Applications* 161, no. 13 (2017). https://doi.org/10.5120/ijca2017913413

[19] Dong, Yuxin, Jianhua Yao, Jiajing Wang, Yingbin Liang, Shuhan Liao, and Minheng Xiao. "Dynamic fraud detection: Integrating reinforcement learning into graph neural networks." In *2024 6th International Conference on Data-driven Optimization of Complex Systems (DOCS)*, pp. 818-823. IEEE, 2024. https://doi.org/10.1109/DOCS63458.2024.10704268

[20] Al-Khater, Wadha Abdullah, Somaya Al-Maadeed, Abdulghani Ali Ahmed, Ali Safaa Sadiq, and Muhammad Khurram Khan. "Comprehensive review of cybercrime detection techniques." *IEEE Access* 8 (2020): 137293-137311. https://doi.org/10.1109/ACCESS.2020.3011259

[21] Chilaka, U. L., G. A. Chukwudebe, and A. Bashiru. "A review of credit card fraud detection techniques in electronic finance and banking." *Conic Res. Eng. J* 3 (2019): 456-467.

[22] Delamaire, Linda, H. A. H. Abdou, and John Pointon. "Credit card fraud and detection techniques: a review." *Banks and Bank systems* 4, no. 2 (2009).

[23] Mehndiratta, Sonal, and Kamal Gupta. "Credit card fraud detection techniques: a review." *Int. J. Computer Sci. Mobile Computing* 8, no. 8 (2019): 43-49.

[24] Lim, Kha Shing, Lam Hong Lee, and Yee-Wai Sim. "A review of machine learning algorithms for fraud detection in credit card transaction." *International Journal of Computer Science & Network Security* 21, no. 9 (2021): 31-40.

[25] Shirgave, Suresh, Chetan Awati, Rashmi More, and Sonam Patil. "A review on credit card fraud detection using machine learning." *International Journal of Scientific & Technology Research* 8, no. 10 (2019): 1217-1220.

[26] Sourabh, and Bhavna Arora. "A Review of Credit Card Fraud Detection Techniques." *Recent Innovations in Computing: Proceedings of ICRIC 2021, Volume 1* (2022): 485-496. https://doi.org/10.1007/978-981-16-8248-3_40

[27] Gyamfi, Nana Kwame, and Jamal-Deen Abdulai. "Bank fraud detection using support vector machine." In *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pp. 37-41. IEEE, 2018. https://doi.org/10.1109/IEMCON.2018.8614994

[28] Naik, Heta, and Prashasti Kanikar. "Credit card fraud detection based on machine learning algorithms." *International Journal of Computer Applications* 182, no. 44 (2019): 8-12. https://doi.org/10.5120/ijca2019918521

[29] Meneghetti, Lorenzo, Matteo Terzi, Simone Del Favero, Gian Antonio Susto, and Claudio Cobelli. "Data-driven anomaly recognition for unsupervised model-free fault detection in artificial pancreas." *IEEE Transactions on Control Systems Technology* 28, no. 1 (2018): 33-47. https://doi.org/10.1109/TCST.2018.2885963

[30] Kulkarni, Pallavi, and Roshani Ade. "Logistic regression learning model for handling concept drift with unbalanced data in credit card fraud detection system." In *Proceedings of the Second International Conference on Computer and Communication Technologies: IC3T 2015, Volume 2*, pp. 681-689. Springer India, 2016. https://doi.org/10.1007/978-81-322-2523-2_66

[31] Gandhar, Akash, Kapil Gupta, Aman Kumar Pandey, and Dharm Raj. "Fraud Detection Using Machine Learning and Deep Learning." *SN Computer Science* 5, no. 5 (2024): 1-10. https://doi.org/10.1007/s42979-024-02772-x

[32] Raghavan, Pradheepan, and Neamat El Gayar. "Fraud detection using machine learning and deep learning." In *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*, pp. 334-339. IEEE, 2019. https://doi.org/10.1109/ICCIKE47802.2019.9004231

[33] Hand, David J., and William E. Henley. "Statistical classification methods in consumer credit scoring: a review." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 160, no. 3 (1997): 523-541. https://doi.org/10.1111/j.1467-985X.1997.00078.x

[34] Gadi, Manoel Fernando Alonso, Alair Pereira do Lago, and Xidi Wang. "A comparison of classification methods applyed on credit card fraud detection." *Technical Report* (2016).

[35] Khare, Navanshu, and Saad Yunus Sait. "Credit card fraud detection using machine learning models and collating machine learning models." *International Journal of Pure and Applied Mathematics* 118, no. 20 (2018): 825-838.

[36] Carneiro, Nuno, Gonçalo Figueira, and Miguel Costa. "A data mining based system for credit-card fraud detection in e-tail." *Decision Support Systems* 95 (2017): 91-101. https://doi.org/10.1016/j.dss.2017.01.002

[37] Paruchuri, Harish. "Credit card fraud detection using machine learning: A systematic literature review." *ABC Journal of Advanced Research* 6, no. 2 (2017): 113-120. https://doi.org/10.18034/abcjar.v6i2.547

[38] Sharma, Pratyush, Souradeep Banerjee, Devyanshi Tiwari, and Jagdish Chandra Patni. "Machine learning model for credit card fraud detection-a comparative analysis." *Int. Arab J. Inf. Technol.* 18, no. 6 (2021): 789-796. https://doi.org/10.34028/iajit/18/6/6

[39] de Sá, Alex GC, Adriano CM Pereira, and Gisele L. Pappa. "A customized classification algorithm for credit card fraud detection." *Engineering Applications of Artificial Intelligence* 72 (2018): 21-29. https://doi.org/10.1016/j.engappai.2018.03.011

[40] Duman, Ekrem, and M. Hamdi Ozcelik. "Detecting credit card fraud by genetic algorithm and scatter search." *Expert Systems with Applications* 38, no. 10 (2011): 13057-13063. https://doi.org/10.1016/j.eswa.2011.04.110

[41] Sravya, Karanam, C. M. Kasthuri, Koramutla Ramesh Meghana, and A. S. Poornima. "Credit Card Fraud Detection using Machine Learning Algorithms-Study of Customer Behaviour." *11th International Conference on Advances in Computing, Control, and Telecommunication Technologies, ACT* 8, no. 2 (2020): 143-150.

[42] Dornadula, Vaishnavi Nath, and Sa Geetha. "Credit card fraud detection using machine learning algorithms." *Procedia computer science* 165 (2019): 631-641. https://doi.org/10.1016/j.procs.2020.01.057

[43] Faraji, Zahra. "A review of machine learning applications for credit card fraud detection with a case study." *SEISENSE Journal of Management* 5, no. 1 (2022): 49-59. https://doi.org/10.33215/sjom.v5i1.770

[44] O Okwonu, Friday Zinzendoff, Nor Aishah Ahad, Nicholas Oluwole Ogini, Innocent Ejiro Okoloko, and Wan Zakiyatussariroh Wan Husin. "Comparative performance evaluation of efficiency for high dimensional classification methods." *Journal of Information and Communication Technology* 21, no. 03 (2022): 437-464. https://doi.org/10.32890/jict2022.21.3.6

[45] Kumar, Kapil, and Vishal Bhatnagar. "Credit Card Fraud Transaction Detection System Using Neural Network-Based Sequence Classification Technique." *International Journal of Open Source Software and Processes (IJOSSP)* 12, no. 1 (2021): 21-40. https://doi.org/10.4018/IJOSSP.2021010102

[46] Tiwari, Pooja, Simran Mehta, Nishtha Sakhuja, Jitendra Kumar, and Ashutosh Kumar Singh. "Credit card fraud detection using machine learning: a study." *arXiv preprint arXiv:2108.10005* (2021).

[47] Mishra, Ankit, and Chaitanya Ghorpade. "Credit card fraud detection on the skewed data using various classification and ensemble techniques." In *2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, pp. 1-5. IEEE, 2018. https://doi.org/10.1109/SCEECS.2018.8546939

[48] Murad, Uzi, and Gadi Pinkas. "Unsupervised profiling for identifying superimposed fraud." In *Principles of Data Mining and Knowledge Discovery: Third European Conference, PKDD'99, Prague, Czech Republic, September 15-18, 1999. Proceedings 3*, pp. 251-261. Springer Berlin Heidelberg, 1999. https://doi.org/10.1007/978-3-540-48247-5_27

[49] Cortes, Corinna, Kathleen Fisher, Daryl Pregibon, and Anne Rogers. "Hancock: a language for extracting signatures from data streams." In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 9-17. 2000. https://doi.org/10.1145/347090.347094

[50] Fawcett, Tom, and Foster Provost. "Activity monitoring: Noticing interesting changes in behavior." In *Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 53-62. 1999. https://doi.org/10.1145/312129.312195

[51] Gamini, Prathima, Sai Tejasri Yerramsetti, Gayathri Devi Darapu, Vamsi Kaladhar Pentakoti, and Prudhvi Raju Vegesena. "A review on the performance analysis of supervised and unsupervised algorithms in credit card fraud detection." *International Journal of Research in Engineering, Science and Management* 4, no. 8 (2021): 23-26.

[52] Okwonu, Friday Zinzendoff, and Abdul Rahman Othman. "A model classification technique for linear discriminant analysis for two groups." *International Journal of Computer Science Issues (IJCSI)* 9, no. 3 (2012): 125.

[53] Dal Pozzolo, Andrea, Olivier Caelen, Yann-Ael Le Borgne, Serge Waterschoot, and Gianluca Bontempi. "Learned lessons in credit card fraud detection from a practitioner perspective." *Expert systems with applications* 41, no. 10 (2014): 4915-4928. https://doi.org/10.1016/j.eswa.2014.02.026