




# Machine Learning Classification Based Techniques for Fraud Discovery in Credit Card Datasets

Roseline Oluwaseun Ogundokun<sup>1</sup>, Sanjay Misra<sup>2</sup> ,  
Opeyemi Eytayo Ogundokun<sup>3</sup>, Jonathan Oluranti<sup>4</sup>, and Rytis Maskeliunas<sup>5</sup>

<sup>1</sup> Department of Computer Science, Landmark University, Omu Aran, Nigeria  
Ogundokun.roseline@lmu.edu.ng

<sup>2</sup> Department of Computer Science and Communication, Ostfold University College,  
Halden, Norway

<sup>3</sup> Directorate of Financial Services, Agricultural and Rural Management Training Institute,  
Ilorin, Nigeria

<sup>4</sup> Center of ICT/ICT, Covenant University, Ota, Ogun State, Nigeria  
jonatha.oluranti@covenantuniversity.edu.ng

<sup>5</sup> Kaunas University of Technology, Kaunas, Lithuania  
Rytis.Maskeliunas@ktu.lt

**Abstract.** The frequency of credit card-based online payment frauds has increased rapidly in recent years, forcing banks and e-commerce companies to create automated fraud detection systems that perform mining on massive transaction logs. Machine learning appears to be one of the most promising techniques for detecting illegal transactions since it uses supervised binary classification algorithms appropriately trained using pre-screened sample datasets to differentiate between fraudulent and non-fraudulent cases. This study aims to concentrate on machine learning (ML) methods thereby proposing a credit card fraud discovery scheme to detect fraud. The ML techniques employed are Decision Tree (DT) and K-Nearest Neighbor (KNN) ML classification techniques. The performance outcomes of the two ML classification techniques are evaluated depending on accuracy, precision, specificity, recall, f1-score, and false-positive rate (FPR). The area under the ROC curve (AUC) of the receiver operating characteristics (ROC) curve was similarly drawn built on the confusion matrix for both classifiers. The two classification techniques were evaluated and compared using the performance metrics mentioned earlier and it was demonstrated that the KNN technique outperformed that of the DT with a greater ROC curve value of 91% for KNN and 86% for DT. It was concluded that KNN is considered a better ML classification technique that can be employed to discover credit card fraudulent activities.

**Keywords:** Credit card fraud · Machine learning · Decision tree · K-Nearest Neighbor · Classification

# 1 Introduction

Credit cards (CC) have presently been progressively prevalent, especially with the upsurge of e-commerce. Credit card fraud (CCF) is a big challenge, even though CC purchases make many sorts of commercial dealings easier. CCF, not only costs fiscal establishments and banks lots of money, but it similarly leads to a lot of worry and tension in the existence of the persons who are impacted. According to current figures, the worldwide fiscal damage instigated by CC theft in 2018 was 27.85 billion dollars, up 16.2% from 23.97 billion dollars in 2017. If current trends continue, credit card theft will cost the economy more than \$35 billion by 2023 [1]. CCF lost to delivering and managing electronic transaction establishments that can be reduced or avoided with effective fraud monitoring and stoppage. Additionally, efficient fraud detection software may boost consumer confidence and minimize complaints. Machine learning is used in the majority of CCF discovery methods [2]. To tackle the problem of CCF, ML has several mature approaches [3, 4], which include supervised learning [5, 6], semi-supervised learning [2, 7], and unsupervised learning [7]. Despite considerable study [7–11], a flawless and competent resolution remains elusive [12].

Several obstacles and challenges facing fraud detection systems include [13–16]: (1) Due to imbrication of data, numerous transactions might appear to be deceitful when they are legitimate businesses. When a fraudulent business looks to be legitimate, the reverse occurs. (2) Data are unbalanced, such as credit card fraud detection data. This indicates that only a small proportion of all CC transactions are deceitful. (3) Adaptability: Structures must be capable of familiarizing to newfangled fraud types. To get the utmost results, an efficient fraud detection approach should be capable of dealing with these issues. A competent fraudster will always come up with new and innovative ways to carry out his task because successful fraud tactics lose their effectiveness over time as they become more widely recognized. Fraud examination (misappropriation discovery) and operator comportment examination are the two broad categories of CCF recognition approaches [15, 17] (Anomaly discovery). The first set of approaches deals with transaction-level supervised categorization. Based on past historical data, these techniques classify transactions as fraudulent or legitimate. This method has been shown to successfully detect the majority of previously identified fraud schemes (known fraud tricks). The second technique is based on account behavior and is based on unsupervised methods. A transaction is flagged as fraudulent in this manner if it deviates from the user's usual behavior (user profile). We don't anticipate fraudsters to act in similar means as the account holder or to be aware of the holder's comportment method. Different enough actions are identified as frauds when fresh behaviors are compared to this model. Even though operator comportment examination approaches are effective in detecting fresh scams, they are plagued by intensifying false alarm rates [17, 18].

Data mining (DM) procedures such as clustering examination; statistics like time series examination; ML technique such as neural network (NN); and artificial intelligence (AI) such as swarm optimization have all been used to implement existing fraud detection [13, 14, 16]. Conventional arithmetical approaches build classification models by approximating parameters to match the data, but ML approaches permit learning the model's specific organization from the data [19, 20].

Consequently, the framework of models acquired using arithmetical approaches are comparatively understandable, not difficult to commentate, and is likely to under-fit the data, whereas replicas created through ML techniques are often complex, difficult to elucidate, and likely to over-fit the data. Underfitting and overfitting data is a tradeoff between a model's descriptive power and frugality, whereas descriptive power leads to extreme forecasting accuracy and frugality typically ensuring the model's generalizability and interpretability. Recent research has demonstrated that data mining approaches based on artificial intelligence (AI) outperformed conventional arithmetical approaches for developing forecasting models [16, 21]. Clustering procedures are divided into hierarchical and partitioned procedures built on their abstraction structure [22]. In an endeavor to recuperate normal clusters that are accessible in the data, hierarchical clustering procedures build to order of divisions, whereas partitioned clustering procedures construct a solitary division of the data with a stated or predictable amount of non-overlapping clusters [22–27]. Among the different clustering algorithms, the K-means procedure is understandable and utmost extensively employed. The k-means technique is employed to reduce data grouping complications. This procedure is affected by the preliminary cluster centers, which are chosen at random. The sophisticated foraging behavior of honey bee swarms inspired the ABC algorithm [28]. ABC has several benefits over other optimization methods, including the use of fewer control parameters and the ability to handle both restricted and unrestrained situations [29]. ABC method was recently created to address clustering issues and has shown capable outcomes in terms of conversion speed and convergence to the optimum result [30, 31].

To improve the classification accuracy and detection rate as well as reduce the false positive rate of credit card discovery, this study hence intends to concentrate on machine learning (ML) methods thereby proposing a credit card fraud discovery scheme to detect fraud. The ML approaches employed are DT and K-Nearest Neighbor (KNN) ML classification techniques. The performance outcomes of the two ML classification techniques are evaluated depending on accuracy, precision, specificity, recall, f1-score, and false-positive rate (FPR). The ROC curve was drawn built on the confusion matrix for both classifiers.

The remaining segment of the article is pre-arranged as thus: Sect. 2 presented the review of related works. Section 3 discussed the materials and methods used for the execution of the study. Section 4 presented the results and implementation of the research and the article was concluded in Sect. 5 with the study conclusion presented.

## 2 Literature Review

There have been several prevailing investigations on CCF discovery approaches, for instance, a variety of research methodologies and fraud recognition strategies, with a focus on neural networks, DM, and distributed data mining. CCF is detected using a variety of methods. After conducting a literature review on several ways CCF recognition, it could be inferred that there are numerous additional approaches in Machine Learning that may be used to identify credit card fraud. SVM, DT, logistic regression (LR), gradient boosting (GB), KNN, and other Machine Learning procedures are employed to identify credit fraud and a few of the researches that have employed these ML techniques are discussed as follows:

SVM, artificial neural networks (ANN), Bayesian networks (BN), hidden Markov model (HMM), KNN, fuzzy logic (FL) system, and DT were among approaches [10] investigated by Jain, Tiwari, Dubey, & Jain [32]. They found that the techniques KNN, DT, and SVM offer an average degree of accuracy in their study. Among all the methods, FL and LR have the lowest accuracy. The detection rate of NNs, NB, fuzzy systems, and KNN had an extreme accuracy value. At the middle level, DT based on LR, SVM, gave a high detection rate (DR). ANN and Naive Bayesian Networks are two methods that outperformed each other across the board. Training costs for these techniques involved a lot of money. For all algorithms, there was a significant flaw. The disadvantage was that these algorithms may not produce consistent results in all situations. With one sort of dataset, they produced superior results, but with another, they produced bad results. Small datasets yield great results from algorithms like KNN and SVM, and raw and unsampled data yielded outstanding accuracy from methods like LR and FL systems.

Naik & Kanikar [33] researched in 2019 on a variety of algorithms such as NB, LR, J48, and Adaboost. Amid the classification procedures, NB was used. The Bayes theorem was used in this algorithm. The Bayes theorem determines the likelihood of an event occurring. The linear regression algorithm and the LR technique are quite alike. The linear regression method was employed to estimate or envisage values. For classification, LR was commonly employed. For the classification function, the J48 method was utilized to construct a DT. J48 is an ID3 extension (Iterative Dichotomiesier). Machine Learning's J48 is one of the most commonly utilized and studied domains and the constant and category variables are the focus of this method. Adaboost is a binary classification method that is one of the utmost frequently employed ML techniques. The procedure's main purpose is to improve the decision tree's performance. This was also how the regression was classified. The Adaboost algorithm uses fraud scenarios to distinguish between fraudulent and non-fraudulent transactions. According to the authors' findings, both the Adaboost and Logistic Regression yielded the greatest accuracy. Because they are both accurate, the time factor was used to select the superior algorithm. They determined that the Adaboost algorithm was effective at detecting CCF when the time component was taken into account.

Sahayasakila, Aishwaryasikhakolli, & Ysaswi [34] introduced the Whale Optimization Approaches (WOA) and SMOTE, which are two key algorithmic techniques (Synthetic Minority Oversampling Techniques). They primarily sought to enhance merging swiftness and resolve the challenge of data unevenness. The SMOTE and WOA techniques are used to solve the challenge of class imbalance. The SMOTE methodology separates all synthetic transactions, which are then re-sampled to ensure data correctness and optimization by utilizing the WOA method. The method similarly boosts the system's concurrence speed, dependability, and competency.

Navanushu & Yunus Sait [35] presented a study that employed DT, RF, SVM, and LR. They used an extremely skewed dataset to work on this sort of dataset. Accuracy, sensitivity, specificity, and precision are used to evaluate the performance of the study. The accuracy of LR was 97.7%, DT was 95.5%, RF was 98.6%, and SVM classifier was 97.5%, according to the results. They determined that, among all the algorithms employed in the study, the RF method has the maximum accuracy and is the best algorithm for detecting fraud. They also concluded that the SVM procedure has a data unevenness challenge and does not perform better in detecting CCF.

The main purpose of fraud detection is to identify the fraudulent actions and if this is done, it aids in characterizing the behavior of the fraudster in the specific fraud act and the historical dataset. Therefore, to detect new fraudulent activities and continually adopt the new credit card fraud activities, we proposed an ML-based classification technique. The goal of the proposed study is to increase the detection rate and accuracy and at the same time reduce the false positive rate (FPR) on CCF activities.

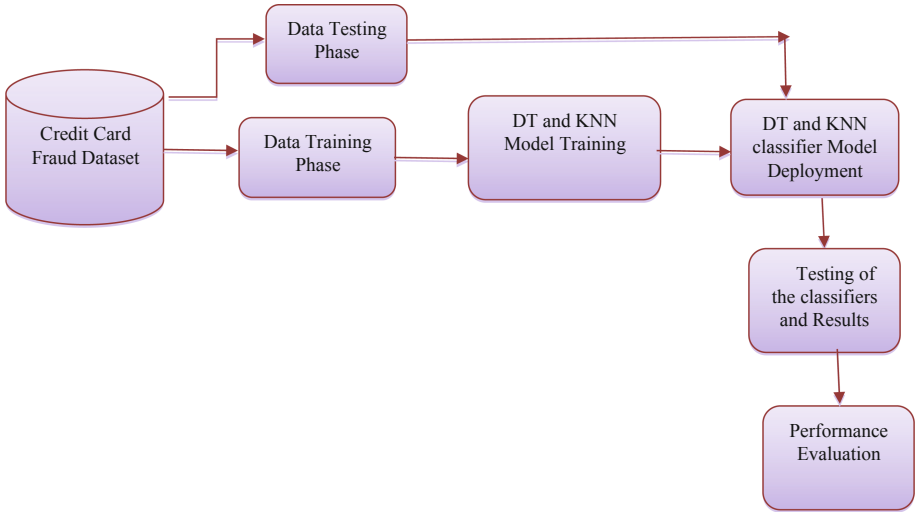
### 3 Materials and Methods

#### 3.1 Dataset

The credit card fraud dataset used for the implementation in this study was gathered from the Kaggle database repository. The dataset can be gotten from this link: <https://www.kaggle.com/rahulmakwana/creditcard-fraud-detection>. The dataset was uploaded to Kaggle by Rahul Makwana in 2020. The dataset comprises 31 numerical features. The dataset comprises 284807 transactions. The overall amount of the sample employed for testing was 85,442 since the test set of data accounted for 30% of the whole dataset.

#### 3.2 Proposed Method

This study employed two ML classification techniques which are DT and KNN. The core objective of this study is to classify the CCF dataset that has together with the fraud and non-fraud transactions in it by employing the two proposed ML classifiers. The classifiers' outcomes are thereby evaluated and likened with each other to establish the classifier that superlatively identifies CCF transactions. The proposed system block diagram for the CCF discovery is shown in Fig. 1. The CCF dataset was first gathered and was passed to the next phase which is the data preparation phase where the data was cleansed and normalized. The dataset is also structured and organized after which it was passed to the testing and training phases where the datasets were split into testing and training. The training dataset was later passed to the ML classifiers DT and KNN for classification after which the classified datasets are then evaluated to deduce the proposed system performance.



**Fig. 1.** Proposed system process flow

### 3.3 Performance Evaluation

The study employed numerous metrics of system classification performance commonly deduced in the literature. The accuracy of the positive (fraud) and negative (non-fraud) situations was measured using recall and specificity matrices. Naturally, there must be a balance between these true positives and true negatives. The various performance indicators are listed in Table 1 concerning the confusion matrix, where positive values correspond to fraud instances and negative values to non-fraud situations [36]. The study employed four performance metrics for the evaluation of the proposed system. The metrics are accuracy, precision, recall, false-positive rate (FPR), and AUC of the classifiers ROC.

**Table 1.** Confusion matrix

	Predicted positive	Predicted negative
Actual positive	True positives	False negatives
Actual negative	False positives	True negatives

## 4 Result and Discussions

Dissimilar measures for technique assessment were employed to evaluate which technique is superlatively appropriate for the challenge of identifying fraud transactions. Accuracy, recall, and precision are the most often used metrics for assessing the results

of ML procedures, but in this study, we utilized five performance measures: accuracy, precision, recall, FPR, and AUC. The confusion matrix for each of the ML classifiers was used to calculate all of the above metrics.

According to these metrics, the performance of the system was evaluated. Both classifiers were used to evaluate the approaches on original datasets, and the results revealed the optimum strategy for CCF discovery. For the implementation training and testing phases of the system, the system employed a 70:30 ratio. The overall amount of the sample employed for testing was 85,442 since the test set of data accounted for 30% of the whole dataset. The confusion matrix for DT is shown in Fig. 2 while the confusion matrix for KNN was displayed in Fig. 3. The AUC of the ROC for DT and KNN was shown in Figs. 4 and 5 respectively (Table 2).

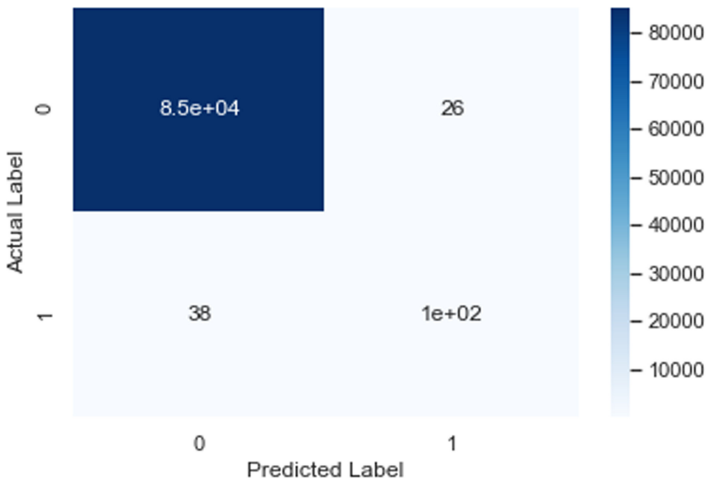


Fig. 2. Confusion matrix for DT

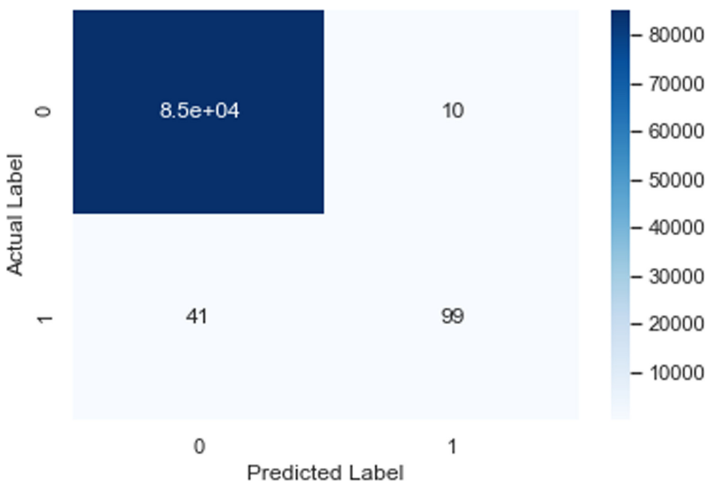
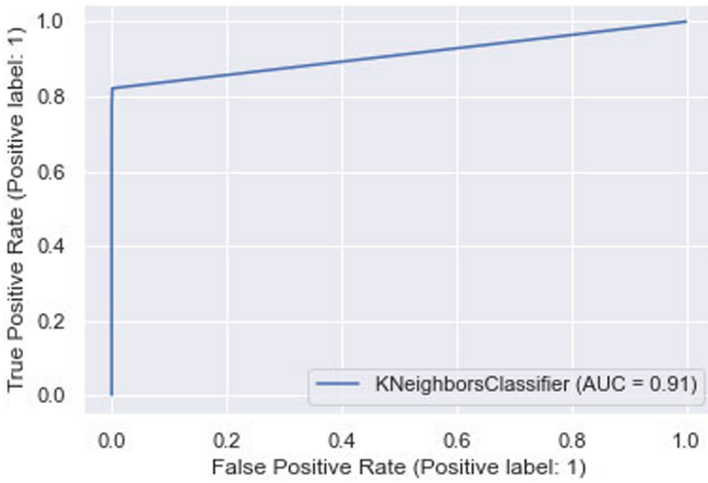


Fig. 3. Confusion matrix for KNN



**Fig. 4.** AUC for DT



**Fig. 5.** AUC for KNN

**Table 2.** Proposed classifiers confusion matrix

Classifiers	TP	TN	FP	FN
DT	85000	100	38	26
KNN	85000	99	41	10



#### 4.1 Discussion

Analyzing the developed system, findings reveal that accuracy is quite high for both classifiers although that of KNN surpassed DT, however, this does not indicate that the outcomes are perfect. Accuracy should be read with caution and it is best when combined with other measures (Varmedja, Karanovic, Sladojevic, Arsenovic, & Anderla, 2019). According to the results, traditional algorithms such as RF can provide similar outcomes to a basic NN (Varmedja, Karanovic, Sladojevic, Arsenovic, & Anderla, 2019). When the acquired discoveries are evaluated with those produced in this present investigation employing conventional algorithms [37] and [38], it is evident that the KNN classifier may upsurge the fraud discovery degree. It has been demonstrated in articles [39] and [40] that traditional procedures may be just as successful as deep learning (DL) methods. Although publications [41] and [42] recommended that DL approaches are superlative for this kind of challenge, it is up to the situation to decide which one to employ. DL technique, for instance, functions well with extra data and may be quickly espoused to other fields than conventional techniques. If there isn't a large proportion of data, though, it's usually best to baton with conventional ML methods. These ML techniques are also not difficult to comprehend and not expensive, both monetarily and computationally [43]. It is said that an algorithm with higher accuracy, precision, recall, f1-score, and AUC is an efficient and effective algorithm [44]. Therefore, in this study, the KNN ML classifier surpassed that of the DT in terms of accuracy of 99.94%, precision of 99.95%, recall of 99.99%, f1-score of 99.97%, NPV of 90.83%, and AUC of 91%. Table 4 displays a comparative analysis of the proposed system with existing systems and it was deduced that the projected system performance surpassed the existing system in terms of 99.94% accuracy, 99.99% recall, and 91% AUC over that of Rtayli & Enneya [45] having 99% accuracy, 95% recall and 81% AUC; Sailusha, Gnaneswar, Ramesh & Rao [46] having 99.91% accuracy, 99.97% recall and AUC was 94% which was higher than the proposed system and lastly Jain, Agrawal & Kumar [47] having a 99.93% accuracy, 99.97% recall but AUC wasn't used for their system evaluation (Table 3).

**Table 3.** Performance evaluation of the proposed system

Measures	DT (%)	KNN (%)
Accuracy	99.92	99.94
Precision	99.96	99.95
Sensitivity	99.97	99.99
Specificity	72.46	70.71
F1-score	99.96	99.97
False positive rate (FPR)	0.2754	0.2929
AUC	86	91

**Table 4.** Comparative analysis with state-of-the-art

Authors	Year	Method	Accuracy	Sensitivity	Specificity
Rtayli & Enneya [45]	2020	RFE, HPO and SMOTE	99%	95%	81%
Sailusha, Gnaneswar, Ramesh & Rao [46]	2020	Random Forest	99.91%	99.97%	94%
Jain, Agrawal & Kumar [47]	2020	Decision Tree	99.93	99.97	N/A
<b>Proposed System</b>	<b>2021</b>	<b>KNN</b>	<b>99.94</b>	<b>99.99</b>	<b>70.71</b>
		<b>DT</b>	<b>99.92</b>	<b>99.97</b>	<b>72.46</b>

## 5 Conclusion and Future Research Direction

Fraudulent credit card transactions are a major corporate issue. These types of scams can result in significant financial and personal losses. As a result, businesses are investing an increasing amount of money in creating new concepts and methods for detecting and preventing fraud. This paper's main objective was to evaluate two machine learning methods for detecting fraudulent transactions. This was determined using a variety of measures, including recall, accuracy, and precision. It is critical to have high recall, accuracy, and precision values for this type of situation. As a consequence of the comparison, it was discovered that the KNN technique produces the best results to that of the DT in terms of accuracy of 99.94%, the precision of 99.95%, recall of 99.99%, f1-score of 99.97%, NPV of 90.83% and AUC of 91%, i.e., it better identifies whether transactions are fraudulent or not. It was also discovered that DT outperformed KNN in terms of specificity 72.46% and FPR of 0.2754. It is therefore concluded that DT outperformed that of KNN because it has a higher specificity of 72.46% and at the same time a lower FPR of 0.2754 and the KNN classifier outperformed that of DT classifier in terms of sensitivity of 99.99% and accuracy of 99.94%.

To improve outcomes, more study should be done on alternative ML methods, for instance, genetic algorithms and several kinds of stacked classifiers, as well as comprehensive feature selection techniques.

## References

1. Tingfei, H., Guangquan, C., Kuihua, H.: Using variational autoencoding in credit card fraud detection. *IEEE Access* **8**, 149841–149853 (2020)
2. Salazar, A., Safont, G., Vergara, L.: Semi-supervised learning for imbalanced classification of credit card transactions. In: 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1–7. IEEE (July 2018)
3. Gao, J., Zhou, Z., Ai, J., Xia, B., Coggeshall, S.: Predicting credit card transaction fraud using machine learning algorithms. *J. Intell. Learn. Syst. Appl.* **11**(3), 33–63 (2019)
4. Yee, O.S., Sagadevan, S., Malim, N.: Credit card fraud detection using machine learning as a data mining technique. *J. Telecommun. Electron. Comput. Eng. (JTEC)* **10**(1–4), 23–27 (2018)

5. Roy, A., et al.: Deep learning detecting fraud in credit card transactions. In: 2018 Systems and Information Engineering Design Symposium (SIEDS), pp. 129–134. IEEE (April 2018)
6. Xuan, S., Liu, G., Li, Z., Zheng, L., Wang, S., Jiang, C.: Random forest for credit card fraud detection. In: 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), pp. 1–6. IEEE (March 2018)
7. Carcillo, F., Le Borgne, Y.A., Caelen, O., Bontempi, G.: Streaming active learning strategies for real-life credit card fraud detection: assessment and visualization. *Int. J. Data Sci. Anal.* **5**(4), 285–300 (2018)
8. Carcillo, F., Le Borgne, Y.-A., Caelen, O., Kessaci, Y., Oblé, F., Bontempi, G.: Combining unsupervised and supervised learning in credit card fraud detection. *Inf. Sci.* **557**, 317–331 (2021). <https://doi.org/10.1016/j.ins.2019.05.042>
9. Eshghi, A., Kargari, M.: Introducing a method for combining supervised and semi-supervised methods in fraud detection. In 2019 15th Iran International Industrial Engineering Conference (IIIEC), pp. 23–30. IEEE (January 2019)
10. Karimi Zandian, Z., Keyvanpour, M.R.: MEFUASN: a helpful method to extract features using analyzing social networks for fraud detection. *J. AI Data Min.* **7**(2), 213–224 (2019)
11. Tran, L., Tran, T., Tran, L., Mai, A.: Solve fraud detection problems by using graph-based learning methods. arXiv preprint [arXiv:1908.11708](https://arxiv.org/abs/1908.11708) (2019)
12. Morgan, J.P.: Payments Fraud and Control Survey. Kirchhain, Germany. <https://www.afp-online.org/publications-data-tools/reports/survey-research-economic-data/Index/>. Accessed 2016
13. Fashoto, S.G., Owolabi, O., Adeleye, O., Wandera, J.: Hybrid methods for credit card fraud detection using K-means clustering with hidden Markov model and multilayer perceptron algorithm. *Br. J. Appl. Sci. Technol.* **13**(5), 1–11 (2016)
14. Philip, N., Sherly, K.K.: Credit card fraud detection based on behavior mining. *TIST Int. J. Sci. Technol. Res.* **1**, 7–12 (2012)
15. Ishu, T., Mrigya, M.: Credit card fraud detection. *Int. J. Adv. Res. Comput. Commun. Eng.* **5**(1), 39–42 (2016)
16. Al-Khatib, A.: Electronic payment fraud detection techniques. *World Comput. Sci. Inf. Technol. J. (WCSIT)* **2**(4), 137–141 (2012)
17. Tripathi, K.K., Pavaskar, M.A.: Survey on credit card fraud detection methods. *Int. J. Emerg. Technol. Adv. Eng.* **2**(11), 721–726 (2012)
18. Dheepa, V., Dhanapal, R.: Behavior-based credit card fraud detection using support vector machines. *ICTACT J. Soft Comput.* **2**(4), 391–397 (2012)
19. Hastie, T., Trevor, H., Robert, T., Friedman, J.H.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York (2001). <https://doi.org/10.1007/978-0-387-21606-5>
20. Azeez, N.A., Asuzu, O.J., Misra, S., Adewumi, A., Ahuja, R., Maskeliunas, R.: Comparative evaluation of machine learning algorithms for network intrusion detection using Weka. In: Chakraverty, S., Goel, A., Misra, S. (eds.) *Towards Extensible and Adaptable Methods in Computing*, pp. 195–208. Springer, Singapore (2018). [https://doi.org/10.1007/978-981-13-2348-5\\_15](https://doi.org/10.1007/978-981-13-2348-5_15)
21. Oladele, T.O., Ogundokun, R.O., Kayode, A.A., Adegun, A.A., Adebisi, M.O.: Application of data mining algorithms for feature selection and prediction of diabetic retinopathy. In: Misra, S., et al. (eds.) *ICCSA 2019. LNCS*, vol. 11623, pp. 716–730. Springer, Cham (2019). [https://doi.org/10.1007/978-3-030-24308-1\\_56](https://doi.org/10.1007/978-3-030-24308-1_56)
22. Vaishali, V.: Fraud detection in credit card by clustering approach. *Int. J. Comput. Appl.* **98**(3), 29–32 (2014)
23. Siddhi, D., Vidhi, S., Jay, V.: Credit card fraud detection using a hybrid approach. *Int. J. Adv. Res. Comput. Commun. Eng.* **5**(5), 287–289 (2016)

24. Madhav, P., Anil, K., Varun, B.: Credit card fraud detection using an efficiently enhanced k-mean clustering algorithm. *Int. J. Eng. Comput. Sci.* **4**(2), 10367–10374 (2015)
25. Nadisha, A., Rakendu, R., Surekha, M.: A hybrid approach to detecting credit card fraud. *Int. J. Sci. Res. Publ.* **5**(11), 304–314 (2015)
26. Vadoodparast, M., Razak, H.: Fraudulent electronic transaction detection using a dynamic model. *Int. J. Comput. Sci. Inf. Secur.* **13**(2), 1–10 (2015)
27. Mortazavi, E., Ahmadzadeh, M.: A hybrid approach for automatic credit approval. *Int. J. Sci. Eng. Res.* **5**(8), 614–619 (2014)
28. Mohd, A., Yuk, Y., Wei, C., Noorhaniza, W., Ahmed, M.: ABC based data mining algorithms for classification tasks. *Can. Cent. Sci. Educ.* **5**(4), 217–231 (2011)
29. Rinkal, S., Samir, K., Hiteshkumar, N.: Artificial bee colony algorithm, a comparative approach for optimization algorithm and application: a survey. *Int. J. Fut. Trends Eng. Technol.* **4**(1), 17–21 (2014)
30. Faiza, A., Azuraliza, A.: A cluster-based deviation detection task using the artificial Bee colony algorithm. *Int. J. Soft Comput.* **2**(7), 71–78 (2012)
31. Deoshree, D., Snehlata, S.: Classification model using optimization technique a review. *Int. J. Comput. Sci. Netw.* **6**(1), 42–48 (2017)
32. Jain, Y., Tiwari, S.N., Jain, S.: A comparative analysis of various credit card fraud detection techniques. *Int. J. Recent Technol. Eng.* **7**(5S2), 402–407 (2019)
33. Naik, H., Kanikar, P.: Credit card fraud detection based on machine learning algorithms. *Int. J. Comput. Appl.* **182**(44), 8–12 (2019)
34. Sahayasakila, V., Aishwaryasikhakolli, D., Ysaswi, V.: Credit card fraud detection system using smote technique and whale optimization algorithm. *Int. J. Eng. Adv. Technol. (IJEAT)* **8**(5), 190–192 (2019)
35. Khare, N., Sait, Y.: Credit card fraud detection using machine learning models and collating machine learning models. *Int. J. Pure Appl. Math.* **118**(20), 825–838 (2018). ISSN 1314-3395
36. Abdulsalam, S.O., et al.: Performance evaluation of ANOVA and RFE algorithms for classifying microarray dataset using SVM. In: Themistocleous, M., Papadaki, M., Kamal, M.M. (eds.) *EMCIS 2020. LNBIP*, vol. 402, pp. 480–492. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-63396-7\\_32](https://doi.org/10.1007/978-3-030-63396-7_32)
37. Mishra, A., Ghorpade, C.: Credit card fraud detection on 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 (February 2018)
38. Navamani, C., Krishnan, S.: Credit card nearest neighbor-based outlier detection techniques. *Int. J. Comput. Tech* **5**(2), 56–60 (2018)
39. Kazemi, Z., Zarrabi, H.: Using deep networks for fraud detection in credit card transactions. In: 2017 IEEE 4th International Conference on Knowledge-Based Engineering and Innovation (KBEI), pp. 0630–0633. IEEE (December 2017)
40. Dhankhad, S., Mohammed, E., Far, B.: Supervised machine learning algorithms for credit card fraudulent transaction detection: a comparative study. In: 2018 IEEE International Conference on Information Reuse and Integration (IRI), pp. 122–125. IEEE (July 2018)
41. Wang, C., Wang, Y., Ye, Z., Yan, L., Cai, W., Pan, S.: Credit card fraud detection based on whale algorithm optimized BP neural network. In: 2018 13th International Conference on Computer Science & Education (ICCSE), pp. 1–). IEEE (August 2018)
42. Pumsirirat, A., Yan, L.: Credit card fraud detection using deep learning based on auto-encoder and restricted Boltzmann machine. *Int. J. Adv. Comput. Sci. Appl.* **9**(1), 18–25 (2018)
43. Deeplearningbook.org: Deep Learning (2019). <https://www.deeplearningbook.org/>. Accessed 11 Jan 2019
44. Su, T., Sun, H., Zhu, J., Wang, S., Li, Y.: BAT: Deep learning methods on network intrusion detection using NSL-KDD dataset. *IEEE Access* **8**, 29575–29585 (2020)

45. Rtayli, N., Enneya, N.: Enhanced credit card fraud detection based on SVM-recursive feature elimination and hyper-parameters optimization. *J. Inf. Secur. Appl.* **55**, 102596 (2020)
46. Sailusha, R., Gnaneswar, V., Ramesh, R., Rao, G.R.: Credit card fraud detection using machine learning. In: 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 1264–1270. IEEE (May 2020)
47. Jain, V., Agrawal, M., Kumar, A.: Performance analysis of machine learning algorithms in credit cards fraud detection. In: 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), pp. 86–88. IEEE (June 2020)