International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(RESEARCH ARTICLE)

Check for updates

Predicting fraud in credit card transactions

Yasmin Akter Bipasha [1, 2, *]

[1] *College of Business, Westcliff University, Irvine, CA 92614, USA.*
[2] *Bangladesh University of Professionals, Mirpur Cantonment, Dhaka-1216, Bangladesh.*

International Journal of Science and Research Archive, 2025, 15(02), 1167-1177

Publication history: Received on 17 April 2025; revised on 22 May 2025; accepted on 25 May 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.15.2.1552

Abstract

The exponential growth of internet-based services has led to an increase in credit card fraud, posing significant financial risks to users and institutions. This study shows the application of supervised machine learning algorithms—specifically Decision Tree and Random Forest classifiers—for effective detection and prediction of fraudulent credit card transactions. Using a large, simulated dataset of 555,719 transactions with both legitimate and fraudulent cases, we addressed the severe class imbalance through an under sampling technique. Our results demonstrate that the Random Forest model outperforms the Decision Tree, achieving an accuracy of 95.80%, sensitivity of 95.80%, precision of 99.58%, and F1 score of 97.49%.

Keywords: Machine Learning; Decision Tree; Random Forest; Credit Card; Fraud Detection and Prediction.

1. Introduction

In the last decade, there has been an exponential growth of the Internet. This has sparked the proliferation and increase in the use of services such as e-commerce, tap and pay systems, online bills payment systems etc. As a consequence, fraudsters have also increased activities to attack transactions that are made using credit cards. There exists a number of mechanisms used to protect credit cards transactions including credit card data encryption and tokenization [1]. Although such methods are effective in most of the cases, they do not fully protect credit card transactions against fraud.

Machine Learning (ML) is a sub-field of Artificial Intelligence (AI) that allows computers to learn from previous experience (data) and to improve on their predictive abilities without explicitly being programmed to do so [2]. In this work we implement Machine Learning (ML) methods for credit card fraud detection. Credit card fraud is defined as a fraudulent transaction (payment) that is made using a credit or debit card by an unauthorized user [3]. According to the Federal Trade Commission (FTC), there were about 1579 data breaches amounting to 179 million data points whereby credit card fraud activities were the most prevalent [4]. Therefore, it is crucial to implement an effective credit card fraud detection method that is able to protect users from financial loss. One of the key issues with applying ML approaches to the credit card fraud detection problem is that most of the published work are impossible to reproduce. This is because credit card transactions are highly confidential. Therefore, the datasets that are used to develop ML models for credit card fraud detection contain anonymized attributes. Furthermore, credit card fraud detection is a challenging task because of the constantly changing nature and patterns of the fraudulent transactions [5]. Additionally, existing ML models for credit card fraud detection suffer from a low detection accuracy and are not able to solve the highly skewed nature of credit card fraud datasets. Therefore, it is essential to develop ML models that can perform optimally and that can detect credit card fraud with a high accuracy score.

* Corresponding author: Yasmin Akter Bipasha

Copyright © 2025 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution License 4.0.

2. Literature Review

Logistic regression is a technique commonly used to predict a binary outcome variable. This method does not require the explanatory variables to follow a normal distribution or be correlated [16]. The outcome variable in logistic regression is categorical, while the explanatory variables may be numerical or categorical. Many researchers have applied logistic regression to detect financial bankruptcies.

A decision tree is a non-linear classification method that splits a dataset into smaller subgroups using a set of explanatory variables. At each branch of the tree, the algorithm selects the variable that has the strongest relationship with the outcome variable, based on a predefined criterion [17]. Being non-parametric, decision trees do not assume unimodal training data and can handle a wide range of quantitative and qualitative data types. However, decision trees are prone to overfitting when applied to the entire dataset, which can reduce their predictive performance. Applications include spam email filtering and identifying individuals at risk of certain diseases in medical fields.

Random forests [18] enhance the bagging method by adding more randomness. They modify how classification or regression trees are built, using different bootstrap samples for each tree and selecting the best split from a random subset of variables at each node. The final prediction is the average output of all trees. The random Forest package in R was used to develop both bagging and random forest models [19]. Feature importance scores can be generated to assess each variable's impact. However, random forests may favor attributes with more levels in datasets containing qualitative variables. Practical applications include bioinformatics (analyzing complex biological data), image classification, and video segmentation.

The types of credit card fraud identified by [20] include bankruptcy fraud, counterfeit fraud, application fraud, and behavioral fraud. Depending on the type of fraud encountered, banks or credit card companies may implement different preventive strategies. For fraud detection in various jurisdictions, machine learning methods such as Logistic Regression, Naive Bayes, Random Forest, K-Nearest Neighbors, Gradient Boosting, Support Vector Machines, and Neural Networks have been applied by [21]. Feature importance was used to select key predictors, and Gradient Boosting achieved an accuracy of 95.9%, outperforming the other models.

A machine learning-based method using hybrid models with AdaBoost and majority voting strategies was developed by [22] for detecting credit card fraud. They introduced noise levels of 10% and 30% to test their hybrid models. A strong score of 0.942 was achieved by multiple voting methods with 30% added noise, making the voting approach the most effective in noisy environments. Similarly, [23] proposed two types of random forests to capture behavioral characteristics of both normal and fraudulent transactions. Data from a Chinese e-commerce platform was used to evaluate these models. Although the proposed random forests performed well on small datasets, issues such as class imbalance reduced their effectiveness on larger or more diverse datasets [3].

Ayorinde et al. [24] used practical approaches for detecting credit card fraud that impacts financial institutions. They tested several machine learning algorithms and identified the best-performing ones. Both under sampling and oversampling techniques were used for training. Among the models tested, Random Forest, XGBoost, and Decision Tree yielded the highest AUC values—100%, 99%, and 99%, respectively.

Machine learning techniques can help detect and classify fraudulent credit card transactions and may even prevent suspicious transactions from proceeding [25]. Fraud detection models are typically trained on historical transaction data labeled as fraudulent or genuine and are then used to predict the outcome of new transactions [26], [27]..

3. Data and Methods

3.1. Data

The data set comprised of simulated transactions of credit cards between January 1, 2020 and December 31, 2020, including both legitimate and fraudulent transactions in the western side of the United States of America [15]. Haris's (2020) sparkov data generation was implemented for the simulation. It includes transactions made to a pool of 800 businesses using the credit cards of 1000 customers. The dataset contains every purchase, the customer's name, the merchant, and the type of purchase, as well as information regarding whether or not the transaction was fraudulent. It contains 555 719 rows of observations, which has 23 columns of variables. 12 of these variables are qualitative data.

In the pre-processing stage, the data was cleaned and formatted to eliminate missing values since our analysis is based on complete data. By performing feature scaling, we kept all numeric explanatory variables within the same domain by using range transformation to compute all numeric variables to be in a range of 0 and 1. We also used under sampling on the imbalanced data to prevent biasing of the algorithms towards the majority class [28]. Values less than 5 and greater than 1250 were removed. Because the dataset in this study is significantly skewed, [29] used Synthetic Minority Oversampling Technique (SMOTE) to balance the data, however, we employed under sampling to handle the imbalance in the dataset. Here, in the minority class, this approach decreases the majority cases to equal or slightly equal to the minority class. Fig. 1 shows the under sampled data. Table 1, Table 2 shows the summary statistics of the types of variables used in the study.

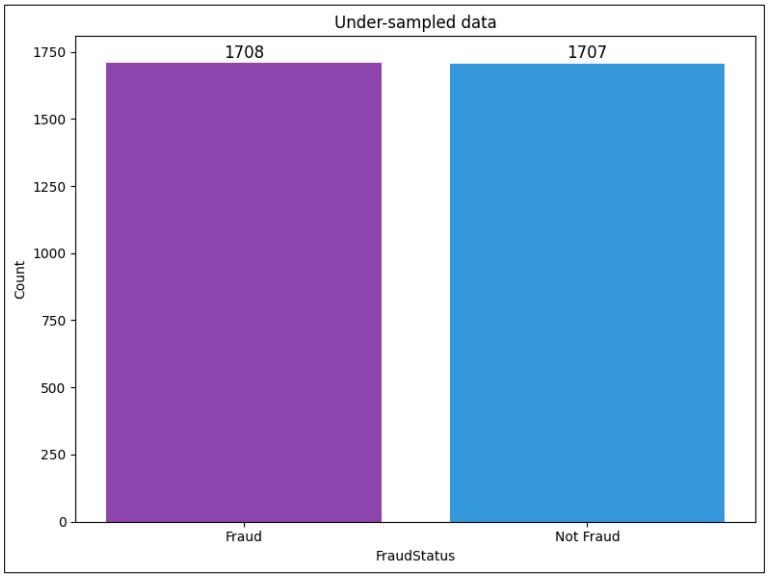

Figure 1 Under sampled data

Table 1 Basic Statistics for character variables

Name	Count	Unique	Top	Frequency
Transaction date and time	555 719	544 760	2020-12-19 16:02:22	4
Merchant	555 719	693	fraud_Kilback LLC	1859
Category	555 719	14	gas_transport	56 370
First	555 719	341	Christopher	11 443
Last	555 719	471	Smith	12 146
Gender	555 719	2	F	304 886
Street	555 719	924	444 Robert Mews	1474
City	555 719	849	Birmingham	2423
State	555 719	50	TX	40 393
Job	555 719	478	Film/video editor	4119
Date of birth	555 719	910	1977-03-23	2408
Transaction number	555 719	555 719	2da90c7d74bd46a	1

Table 2 Basic Statistics for numeric variables

Name	Count	Mean	Std	Min	25%	50%	75%
Unique identifier	555 719	277 859	160 422.4	0	138 929.5	277 859	416 788.5
Credit card number of customers	555 719	4 178 387	1 309 837	6 041 621	1 800 429	3 521 417	4 635 331
Amount	555 719	69.39	156.75	1	9.63	47.29	83.01
Zip	555 719	48 842.63	26 855.28	1257	26 292	48 174	72 011
Latitude	555 719	38.54	5.061	20.03	34.67	39.37	41.89
Longitude	555 719	−90.23	13.72	−165.67	−96.8	−87.48	−80.18
City population	555 719	88 221.89	300 390.9	23	741	2408	19 685
Time (s)	555 719	1 380 679	5 201 104	1 371 817	1 376 029	1 380 762	1 385 867
Merchant latitude	555 719	38.54	5.1	19.03	34.76	39.38	41.95
Merchant longitude	555 719	−90.23	13.73	−166.67	−96.91	−87.45	−80.27
Fraud status	555 719	0.0039	0.062	0	0	0	0

3.2. Methods

In this section, we discuss the supervised machine learning models such as Random Forest, and Decision Tree to classify fraudulent transactions.

3.2.1. Decision tree

Decision trees are non-parametric supervised learning techniques that can be employed for classification [30]. They generate decision rules with a tree-like structure using actual data attributes. Decision Trees evolved from the way humans make decisions [31, 32]. Graphically, they show information in a tree pattern that is easy to understand. The decision tree structure is made up of nodes, edges, and leaf nodes. According to [24], it consists of a set of branches/nodes that are connected by edges. The decision tree algorithm has the benefit of not needing feature scaling, being robust to outliers, and handling missing values automatically. It is quicker to train and is very good at resolving classification and prediction problems. The decision tree uses the following; the Gini index, information gain, and entropy as a metric for classification into two or more nodes.

Entropy is a measure of expected randomness or impurity in a dataset, typically ranging between 0 and 1 [33]. In the context of the Internet of Medical Things (IoMT) [34], entropy is crucial in decision-tree-based algorithms for analyzing medical sensor data. The formula for calculating entropy is

$$E(X) = -\sum_{i=1}^{n} p(x_i) \log_2 p(x_i) \quad (1)$$

$$E(X) = -p(Fraud) \log_2 p(Fraud) - p(Not\ \ Fraud) \log_2 p(Not\ \ Fraud) \quad (2)$$

3.2.2. Random Forest

Random Forest is a supervised machine learning algorithm [35, 36] that uses a group of decision tree models for classification and making predictions [37]. Each decision tree is a weak learner because they have a low predictive power. It is based on ensemble learning, which uses many decision tree classifiers to classify a problem and improve the accuracy of the model [38]. As a result, the random forest employs a bagging method to generate a forest of decision trees [52]. Given a dataset (X,Y) with N total observation where X being the predictor variables and Y the outcome variable, the random forest algorithm first creates Ki random variables (i=1,2,…,N) to form a vector and then converts each Ki random vector into a decision tree to obtain the dKi decision tree (dK1(X),dK2(X),…,dKN(X)). The final classification results are as follows:

$$D(X) = \arg\max\left\{\sum_i^N dK_i(X) = Fraud, \sum_i^N dK_i(X) = Not\ \ fraud,\right\}$$

Random forest typically does not require a feature selection procedure [39]. The drawback of this approach is how quickly it may identify data with a wide range of values and variables with numerous values as fraudulent. It is one of the financial sector's most accurate fraud detection algorithms, according to [40]. It is usually more uncertain when the Random Forest method begins to build the tree, so it is crucial to choose the most important feature out of all features for analysis, particularly in node splitting.

The entries in the confusion matrix (Table 3) are defined as the following: False positive (FP) is the total number of incorrect predictions classified as positive; False negative (FN) is the total number of incorrect predictions classified as negative; True positive (TP) is the total number of true predictions classified as positive; and True negative (TN) is the total number of true predictions classified as negative [41-59].

Table 3 Confusion Matrix

Predicted class	**Actual class**	
	Fraud (1)	**Not Fraud (0)**
Fraud (1)	True Positive (TP)	False Positive (FP)
Not Fraud (0)	False Negative (FN)	True Negative (TN)

4. Results

Table 4 shows the transaction status of the data. We observe that there are 0.4% of fraudulent transactions while the remaining 99.6% were true transactions.

Table 4 Transaction description

Description	**Fraud**	**Non-Fraud**
Total	2135	482 672
Percentage (%)	0.4%	99.6

As illustrated in fig. 2, most of the fraudulent transactions occurred in the shopping category (1.19%), followed by grocery (0.73%), miscellaneous (0.36%), transport (028%), and home care (0.16%). It is not surprising that home care transactions recorded fraud since not many transactions occur there.

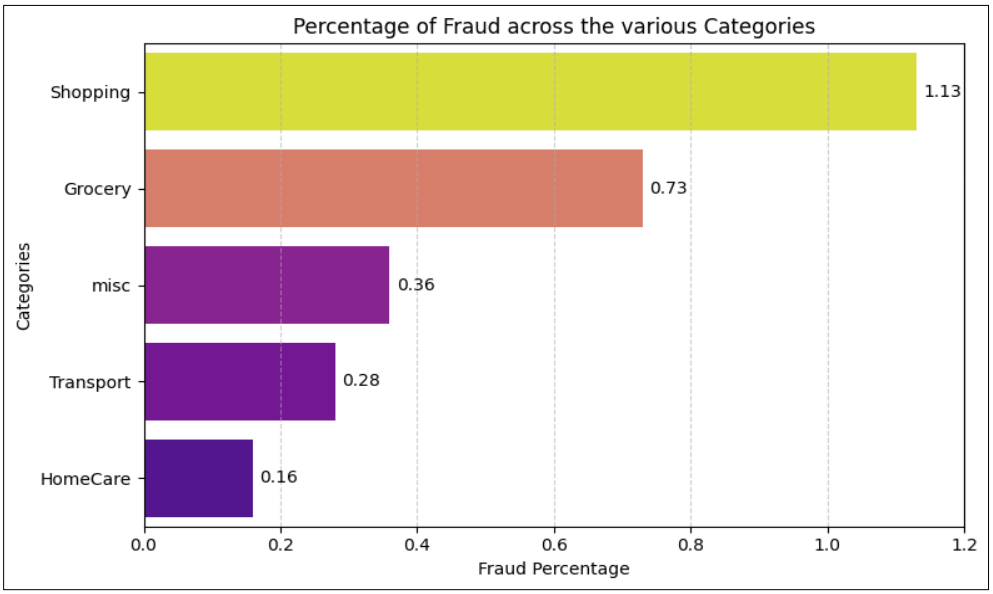

Figure 2 Fraudulent transaction across merchant categories

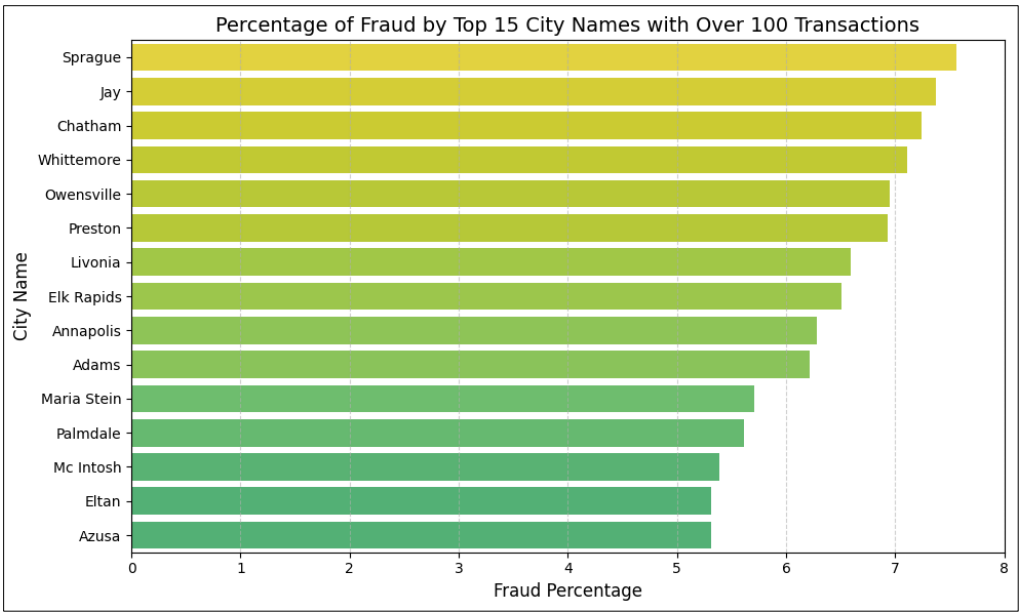

Figure 3 Percentage of fraudulent transactions among cities with over 100 credit card transactions

Most fraudulent credit card transactions affected customers in the cities of Jay and Chatham. From fig. 3, the cities of Sprague and Jay had the greatest percentage of fraudulent transactions, with a percentage of 7.56 and 7.37, respectively. The chart for the 15 cities with transactions above 100 and their percentages of fraudulent transactions is shown in fig. 3.

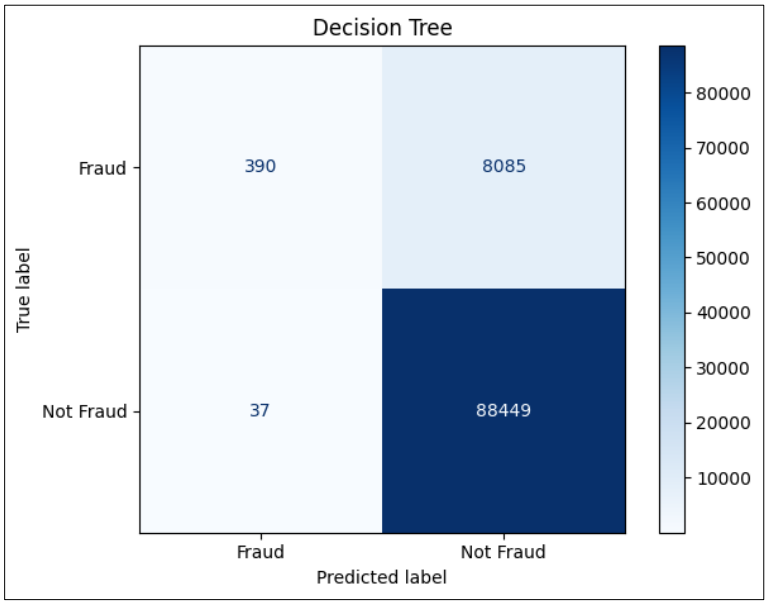

Figure 4 Confusion matrix of prediction using decision tree

Table 5 summarizes the results of the predictions of a confusion matrix when using the Decision Tree model. The model was able to correctly classify 390 fraudulent transactions out of the 427 total fraudulent transactions from the testing data as fraudulent, whereas 37 fraudulent transactions were labelled as not fraudulent. Once more, 8085 Not Fraud transactions were incorrectly classified as Fraud, whereas 88 449 Not Fraud transactions were correctly classified as Not Fraud. Table 5 shows the performance matrix of Decision Tree.

Table 5 Performance of Decision Tree

Metric measure	Estimate (%)
Accuracy	91.62
Sensitivity	91.62
Precision	99.54
F1 Score	95.23

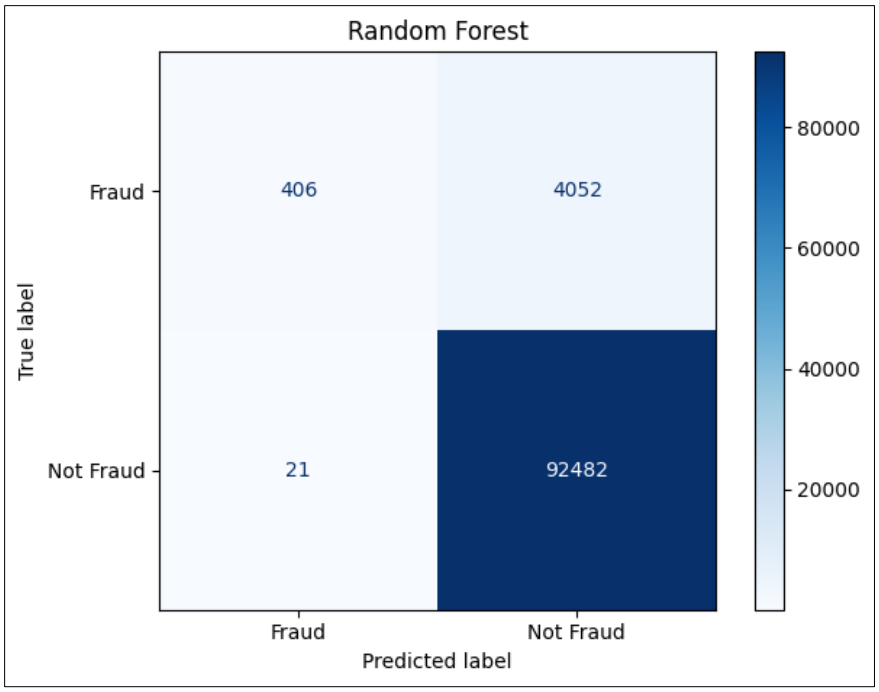

Figure 5 Confusion matrix of prediction using RF

Fig. 5 shows the output of the predictions in a confusion matrix. Out of the 427 total transactions from the Testing Data, the model was able to correctly classify 406 fraud transactions as fraud while 21 fraud transactions were classified as Not Fraud. Once more, 4052 not fraud transactions were incorrectly classified as fraud, whereas 92 482 not fraud transactions were appropriately classified as not fraud. Table 6 shows the performance matrix of Random Forest.

Table 6 Performance of Random Forest

Metric measure	Estimate
Accuracy	95.80
Sensitivity	95.80
Precision	99.58
F1 Score	97.49

5. Conclusion

In order to categorize online credit card transactions as either fraud or not, this study built two different classification models, Decision Tree, and Random Forest using supervised machine learning. To ensure that the model does not favour solely the majority class and prevent overfitting the model to the data, we balanced the dataset prior to generating the

models using the under sampling technique. With an accuracy value of 95.80%, the Random Forest model performed better than the other two models, making it the most suitable model for predicting fraudulent transactions,

Based on the data and analysis, it was determined that the majority of fraud cases occur between the hours of 20 (10 pm) and 5 (5 am). It can be concluded that banks will not be operating to monitor transactions at this time, and victims might be sleeping as well and the possibility of fraudsters to commit fraud is created by this.

The analysis revealed that cardholders over the age of 60 are most frequently the targets of fraudulent transactions. Adults over 60 seem to be more likely to report losses from particular sorts of fraud.

Based on the data and analysis performed, we recommend that the financial institutions should prioritize providing older clients with more in-person services. They must boost their security measures or over online services between the hours of 10 pm and 5 am.

As a matter of urgency, they should develop more robust and fraud-free systems. It is imperative that financial institutions embrace random forest model in predicting and detecting daily credit card fraud.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Iwasokun GB, Omomule TG, Akinyede RO. Encryption and tokenization-based system for credit card information security. Int J C,yber Sec Digital Forensics. 2018;7(3):283–93.

[2] Md. Hossain, Md. Bahar Uddin, "Digital twins in additive manufacturing", World Journal of Advanced Engineering Technology and Sciences, 2024, 13(02), 909-918.

[3] Maniraj SP, Saini A, Ahmed S, Sarkar D. Credit card fraud detection using machine learning and data science. Int J Eng Res 2019; 8(09).

[4] Dornadula VN, Geetha S. Credit card fraud detection using machine learning algorithms. Proc Comput Sci. 2019;165:631–41.

[5] Thennakoon, Anuruddha, et al. Real-time credit card fraud detection using machine learning. In: 2019 9th international conference on cloud computing, data science & engineering (Confluence). IEEE; 2019.

[6] E. Ileberi, Y. Sun, and Z. Wang, "A machine learning based credit card fraud detection using the GA algorithm for feature selection," Journal of Big Data, vol. 9, no. 1, Feb. 2022, doi: https://doi.org/10.1186/s40537-022-00573-8.

[7] F. Carcillo, Borgne, Y. Le, O. Caelen, Y. Kessaci, F. Oblé, Combining unsupervised and supervised learning in credit card fraud detection, Inform. Sci. 557 (2021) 317–331, http://dx.doi.org/10.1016/j.ins.2019.05.042.

[8] S. Xuan, S. Wang, Random forest for credit card fraud detection, 2018.

[9] V. Vlasselaer, Van, C. Bravo, O. Caelen, T. Eliassi-rad, L. Akoglu, M. Snoeck, B. Baesens, APATE : A novel approach for automated credit card transaction fraud detection using network-based extensions, Decis. Support Syst. 75 (2015) 38–48, http://dx.doi.org/10.1016/j.dss.2015.04.013.

[10] L.E. Faisal, T. Tayachi, S. Arabia, L.E. Faisal, O. Banking, The role of internet banking in society. 18 (13) (2021) 249–257.

[11] Dorphy, H. Hultquist, 2017 Financial Institution Payments Fraud Mitigation Survey, Federal Reserve Bank of Minneapolis, 2018.

[12] E. Kurshan, H. Shen, H. Yu, Financial crime & fraud detection using graph computing: Application considerations & outlook, in: 2020 Second International Conference on Transdisciplinary AI (TransAI), IEEE, 2020, pp. 125–130.

[13] Bharati, S., Robel, M.R.A., Rahman, M.A., Podder, P., Gandhi, N. (2021). Comparative Performance Exploration and Prediction of Fibrosis, Malign Lymph, Metastases, Normal Lymphogram Using Machine Learning Method. In:

Abraham, A., Panda, M., Pradhan, S., Garcia-Hernandez, L., Ma, K. (eds) Innovations in Bio-Inspired Computing and Applications. IBICA 2019. Advances in Intelligent Systems and Computing, vol 1180. Springer, Cham. https://doi.org/10.1007/978-3-030-49339-4_8

[14] B. Lebichot, Y.A.L. Borgne, L. He-Guelton, F. Oblé, G. Bontempi, Deep-learning domain adaptation techniques for credit cards fraud detection, in: INNS Big Data and Deep Learning Conference, Springer, Cham, 2019, pp. 78–88.

[15] https://www.kaggle.com/datasets/kartik2112/fraud-detection

[16] B.G. Tabachnick, L.S. Fidell, Using Multivariate Statistics, Harper Collins, New York, 1996.

[17] J.A. Michael, S.L. Gordon, Data Mining Technique for Marketing, Sales and Customer Support, John Wiley & Sons INC, New York, 1997, p. 445.

[18] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.

[19] A. Liaw, M. Wiener, Classification and regression by randomForest, R News 2 (3) (2002) 18–22.

[20] O. Citation, B. Systems, University of Huddersfield Repository Credit card fraud and detection techniques : a review

[21] A. Aditi, A. Dubey, A. Mathur, P. Garg, Credit Card Fraud Detection Using Advanced Machine Learning Techniques. (2022) 56–60. http://dx.doi.org/10. 1109/ccict56684.2022.00022.

[22] K. Randhawa, C.H.U.K. Loo, S. Member, Credit card fraud detection using AdaBoost and majority voting, IEEE Access 6 (2018) 14277–14284, http://dx. doi.org/10.1109/ACCESS.2018.2806420.

[23] L. Guanjun, L. Zhenchuan, Z. Lutao, W. Shuo, Random forest for credit card fraud, IEEE Access (2018).

[24] K. Ayorinde, Cornerstone : A Collection of Scholarly and Creative Works for Minnesota State University, Mankato a Methodology for Detecting Credit Card Fraud a METHODOLOGY for DETECTING CREDIT CARD FRAUD Kayode Ayorinde (T

[25] A.D. Pozzolo, G. Boracchi, O. Caelen, C. Alippi, Credit Card Fraud Detection : A Realistic Modeling and a Novel Learning Strategy. 29(8) (2018) 3784–3797.

[26] A. Dal Pozzolo, O. Caelen, Y.A. Le Borgne, S. Waterschoot, G. Bontempi, Learned lessons in credit card fraud detection from a practitioner perspective, Expert Syst. Appl. 41 (10) (2014) 4915–4928.

[27] G. Bontempi, Reproducible machine learning for credit card fraud detection - practical machine learning for credit card fraud detection - practical handbook foreword. May, 2021.

[28] A.D. Pozzolo, O. Caelen, R.A. Johnson, G. Bontempi, Calibrating probability with undersampling for unbalanced classification, 2015.

[29] R. Tyagi, R. Ranjan, S. Priya, Credit card fraud detection using machine learning algorithms. (2021) 334–341.

[30] B.T. Jijo, A.M. Abdulazeez, Classification Based on Decision Tree Algorithm for Machine Learning. 02 (01) (2021) 20–28. http://dx.doi.org/10.38094/ jastt20165.

[31] J.F.S. Iii, Evolving Fuzzy Decision Tree Structure that Adapts in Real-Time. (2005) 1737–1744.

[32] Tanvir Mahmud, "ML-driven resource management in cloud computing", World Journal of Advanced Research and Reviews, 2022, 16(03), 1230-1238.

[33] N. Freitas, Decision Trees, University of British Columbia, 2013.

[34] Tanvir Mahmud, "Applications for the Internet of Medical Things", International Journal of Science and Research Archive, 2023, 10(02), 1247-1254.

[35] Y. Liu, L. Hu, F. Yan, B. Zhang, Information Gain with Weight based Decision Tree for the Employment Forecasting of Undergraduates. (2013) 2–5. http: //dx.doi.org/10.1109/GreenCom-iThings-CPSCom.2013.417.

[36] Tanvir Mahmud, S A Sabbirul Mohosin Naim, "Predicting polycystic ovary syndrome using SVM ", International Journal of Science and Research Archive, 2024, 13(02), 4400-4408.

[37] T.R. Prajwala, A comparative study on decision tree and random forest using R tool, Int. J. Adv. Res. Comput. Commun. Eng. 4 (1) (2015) 196–199.

[38] E. Kabir, S. Guikema, B. Kane, Statistical modeling of tree failures during storms, Reliab. Eng. Syst. Saf. 177 (April) (2018) 68–79, http://dx.doi.org/10.1016/j. ress.2018.04.026.

[39] J.L. Speiser, A random forest method with feature selection for developing medical prediction models with clustered and longitudinal data, J. Biomed. Inform. 1172020 (2021) 103763, http://dx.doi.org/10.1016/j.jbi.2021.103763.

[40] N. Donges, A complete guide to the random forest algorithm, 2021, 2019. URL: https://builtin.com/data-science/random-forest-algorithm (: 25.05. 2021).

[41] Yasmin Akter Bipasha, "Market efficiency, anomalies and behavioral finance: A review of theories and empirical evidence", World Journal of Advanced Research and Reviews, 2022, 15(02), 827-839.

[42] Bharati S, Podder P, Mondal MRH, Prasath VBS. CO-ResNet: Optimized ResNet model for COVID-19 diagnosis from X-ray images. International Journal of Hybrid Intelligent Systems. 2021;17(1-2):71-85. doi:10.3233/HIS-210008.

[43] M. B. Hossain, K. Hoque, S. Abdi, E. Bazgir and M. A. Rahman, "Design and Simulation of a 1×2 Rectangular Microstrip Patch Antenna Array with Feeding Network," 2025 Fifth International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), Bhilai, India, 2025, pp. 1-7, doi: 10.1109/ICAECT63952.2025.10958936.

[44] Khandoker Hoque, Md Boktiar Hossain, Anhar Sami, Denesh Das, Abdul Kadir, Mohammad Atikur Rahman, "Technological trends in 5G networks for IoT-enabled smart healthcare: A review", International Journal of Science and Research Archive, 2024, 12(02), 1399–1410.

[45] Md Boktiar Hossain and Khandoker Hoque, "Machine Learning approaches in IDS", International Journal of Science and Research Archive, 2022, 07(02), 706-715.

[46] Halbouni, A., Gunawan, T. S., Habaebi, M. H., Halbouni, M., Kartiwi, M., & Ahmad, R. (2022). Machine learning and deep learning approaches for cybersecurity: A review. IEEE Access, 10, 19572-19585.

[47] Yasmin Akter Bipasha, "Blockchain technology in supply chain management: transparency, security, and efficiency challenges", International Journal of Science and Research Archive, 2023, 10(01), 1186-1196.

[48] Bharati S., Podder P., Mondal M.R.H. Diagnosis of polycystic ovary syndrome using machine learning algorithms; Proceedings of the 2020 IEEE Region 10 Symposium (TENSYMP); Dhaka, Bangladesh. 5–7 June 2020; pp. 1486–1489.

[49] Amit Deb Nath, Rahmanul Hoque, Md. Masum Billah, Numair Bin Sharif, Mahmudul Hoque . Distributed Parallel and Cloud Computing: A Review. International Journal of Computer Applications. 186, 16 (Apr 2024), 25-32. DOI=10.5120/ijca2024923547

[50] Mobasher Hasan, Jubair Bin Sharif, Md. Kwosar, Md. Faysal Ahmed, Daniel Lucky Michael . Maximizing Business Performance through Artificial Intelligence. International Journal of Computer Applications. 186, 54 (Dec 2024), 9-15. DOI=10.5120/ijca2024924252

[51] Md Bahar Uddin, Md. Hossain and Suman Das, "Advancing manufacturing sustainability with industry 4.0 technologies", International Journal of Science and Research Archive, 2022, 06(01), 358-366.

[52] Md Boktiar Hossain, Khandoker Hoque, Mohammad Atikur Rahman, Priya Podder, Deepak Gupta, "Hepatitis C Prediction Applying Different ML Classification Algorithm", International Conference on Computing and Communication Networks 2024 (ICCCNet 2024) (Aceepted)

[53] Md Maniruzzaman, Md Shihab Uddin, Md Boktiar Hossain, Khandoker Hoque, "Understanding COVID-19 Through Tweets using Machine Learning: A Visualization of Trends and Conversations", European Journal of Advances in Engineering and Technology, 10(5), 108-114.

[54] Khamparia, A., Mondal, R. H., Podder, P., Bhushan, B., de Albuquerque, V. H. C., & Kumar, S. (Eds.). (2021). Computational intelligence for managing pandemics (Vol. 5). Walter de Gruyter GmbH & Co KG.

[55] Rahmanul Hoque, Md Masum Billah, Amit Debnath, S. M. Saokat Hossain and Numair Bin Sharif, "Heart Disease Prediction using SVM", International Journal of Science and Research Archive, 2024, 11(02), 412–420.

[56] Hoque, M., Hasan, M. R., Emon, M. I. S., Khalifa, F., & Rahman, M. M. (2024, September). Medical image interpretation with large multimodal models. In CEUR workshop proceedings.

[57] Emon, M., Hoque, M., Hasan, M., Khalifa, F., & Rahman, M. (2024, September). Fingerprint identification of generative models using a multiformer ensemble approach. In CEUR workshop proceedings. https://ceur-ws. org/Vol-3497/.

[58] Emon, M. I. S., Hoque, M., Hasan, M. R., Khalifa, F., & Rahman, M. (2025, April). A novel vision transformer-based approach to detect generative model fingerprint. In Medical Imaging 2025: Imaging Informatics (Vol. 13411, pp. 336-342). SPIE.

[59] J. K. Afriyie et al., "A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions," Decision Analytics Journal, vol. 6, no. 100163, p. 100163, Mar. 2023, doi: https://doi.org/10.1016/j.dajour.2023.100163.