# Exploring Supervised Machine Learning Techniques for Detecting Credit Card Fraud: An Investigative Review

**Amit Patel[1*], Manish Patel[2], Pankaj Patel[3]**

M.Tech Scholar, Sankalchand Patel University, Visnagar, India1
Professor, CE Dept., SPCE, Sankalchand Patel University, Visnagar, India2
I/C HoD, CE Dept., SSPC, Sankalchand Patel University, Visnagar, India3

*amit86india@gmail.com[1], mmpatel.fet@spu.ac.in[2], pspatel.sspc@spu.ac.in[3]

**Abstract**: Given the current situation of the economy, credit card use has increased significantly. Users can make significant cash payments with these cards without carrying a lot of cash on them. They have simplified the process of conducting cashless transactions and enabled consumers to make payments of any kind with greater ease. While there are many benefits to using this electronic payment method, there are also some risks. In tandem with the expansion of the consumer base. A specific person's credit card information may be unlawfully acquired and used in fraudulent purchases. To tackle this issue, certain machine learning methods may be applied to gather information. This research offers a comparative analysis of many supervised learning method for identifying real from fake transactions. In this article, we have covered a variety of techniques for spotting credit card fraud.

**Keywords:** Credit Card, Credit Card Fraud, Machine Learning, Supervised Learning.

## I. INTRODUCTION

A fraud is an intentional deception carried out with the objective to achieve any kind of advantage, most commonly financial. It's an unfair practice that gets more widespread by the day. The sharp rise in the use of credit and debit cards as electronic means of payment has led to a rise in credit card fraud. Payments can be made using these cards online and off. It is possible that when making an online payment, the card does not need to be physically visible. Hackers or fraudsters may target the card data in such instances. An annual loss of millions of dollars has been suffered due to these types of frauds. Many different types of algorithms have been and are being developed to address this problem. Various detecting methods are being researched to find the best workable solution for this problem.

These days, although credit card transactions are very popular, they also have a unique set of issues. Recognizing fraud brings up a variety of issues. A transaction can be accepted or rejected in a matter of milliseconds, often less than a second. This means that identifying a fraudulent transaction requires a very fast and efficient method. The sheer number of comparable transactions occurring simultaneously is an additional cause for worry. This makes it challenging to keep an eye on every transaction separately and identify fraud. This implies that in order to distinguish between a legitimate transaction and a fraudulent one, an effective fraud detection system needs to be implemented. This type of system is intended to determine the specific ways in which each user uses their cards. As a result, the data works with both supervised and unsupervised machine learning methods currently in use. This work aims to evaluate an imbalanced dataset and identify the best supervised machine learning model for credit card fraud detection using a variety of models.

The rest of the text is structured as follows: Section II describes algorithms for machine learning. Section III contains related work. Section IV: Performance Metrics and Analysis of Various Algorithms, Section V contains Open Research Issues, while Section VI provides the conclusion.

## II. DIFFERENT MACHINE LEARNING ALGORITHMS

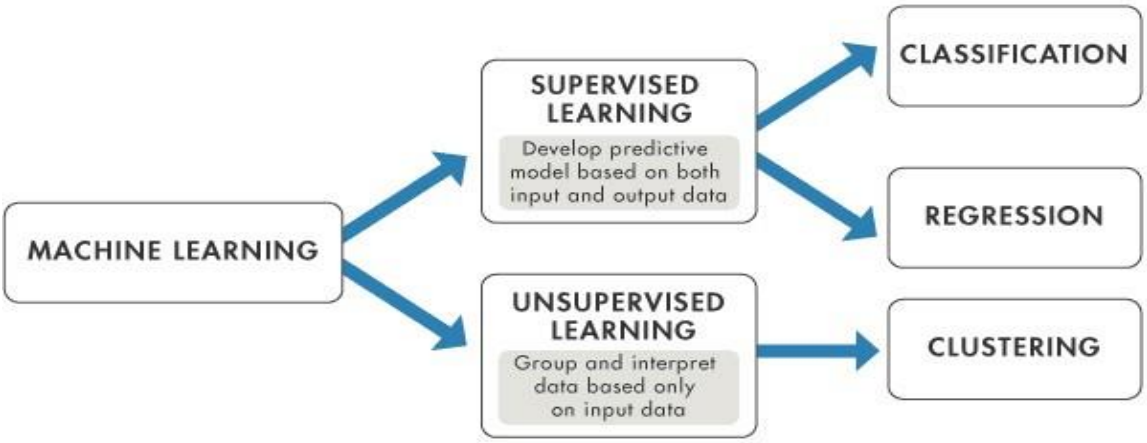


Fig. 1 Types of Machine Learning

Logistic Regression [3, 5]: One kind of classification technique used to forecast a binary output variable is called logistic regression. It is frequently utilized in machine learning applications, such spam filtering and fraud detection, where the output variable is either true or false. The goal of the logistic regression procedure is to determine whether the input features and the output variable have a linear relationship. A logistic function is then used to modify the output variable, yielding a probability value between 0 and 1.

Random Forest [1, 3]: This model is essentially a combining classifier that uses and combines many decision tree classifiers. The main goal of using many trees is to allow for adequate training of the trees, allowing each tree to contribute as a model. After the tree is generated, the majority is used to merge the result. It employs several decision trees, each reliant on a distinct dataset whose distribution is consistent across the tree. This particular technique has the capacity to balance errors in a class population of unequal data sets in an efficient manner. It may be used for both classification issues and regression challenges.

Decision Tree [3, 5]: This is one of the most often used approaches to predictive modeling. This is built like a tree, as implied by the model's name. This model may be used when performing a multi-dimensional analysis with several classes. The historical data, also known as the historical vector, is used to develop a model that may be used to predict the value of the output based on the input provided. A tree is made up of several nodes, each of which is associated with a distinct vector. Each leaf node, which represents a possible outcome, marks the conclusion of the tree.

KNN [1, 3]: The k-Nearest Neighbor model is one of the easiest to understand and most effective models. The class label of the training data components that are next to each other in this model determines the class label of the test datasets. One method for calculating how similar two items are to one another is the Euclidean Distance. It is also known as instance learning or a lazy model. It computes the number of "k," or the closest neighbors that need to be considered. For "k," a reasonable number should be chosen. The use of an appropriate distance measure is another need. Sometimes, the "Minkowski" distance is used. It is a generalized version of the Manhattan and Euclidean distances.

Naive Bayes [3]: It is a type of probabilistic classifier model, which suggests that it may predict several classes simultaneously. The Bayes Theorem forms its foundation. Multiple class predictions are feasible using probabilistic classifiers. Conditional probability serves as the foundation for the choice. Rather of relying on a single algorithm, this approach employs a collection of algorithms that share a similar idea. It is assumed in this model that every feature contributes to the output in an equal and unique way. Because it needs less training data than other models, this model offers a few advantages over them.

## III. RELATED WORK

### *A. Federated Learning and ANN [1]*

In [1], The authors have proposed a hybrid technique that combines neural networks (ANN) with federated learning architecture. It has been recognized as an effective technique for preserving anonymity while raising CCFD accuracy. According to the authors' proposed hybrid method, real-time datasets may be used to train the model in a way that protects privacy. The Federated Learning (FL) method based on ANNs can enhance the ML model's ability to detect fraudulent

transactions. The proposed hybrid technique can significantly alter the way CCFD functions and create new opportunities for the banking and financial industries with the use of real-world statistics.

Merits- Together, these strategies can help banks and other financial institutions leverage real-time datasets, which will help everyone involved in the process of developing a successful CCFD system.

Demerits-There still needs to be work done form gaining the confidence of banks and financial institutes to adopt this technology.

### *B. Clustering and similarity-based selection (SBS) [2]*

In [2], The minority class is disregarded and considered as noise by the algorithm since common methods like logistic regression benefit the majority class so much, which causes the two classes' distribution ratios to be unbalanced. The authors offer a framework that first clusters the dataset using fuzzy C-means and then resamples it using our proposed Similarity-Based Selection (SBS) method to assist improve the effectiveness and precision of the detection process. With an accuracy of 0.989, ANN was the most accurate, followed by LR (0.986), NB (0.984), and KNN (0.966).

Merits-Provide a framework (SBS) to use Fuzzy C-means to tackle the imbalanced class distribution problem. To highlight the effectiveness of the SBS technique and demonstrate its superiority, the experiment's performance is compared with that of alternative approaches.

Demerits- The authors want to carry out more study with the aim of improving the provided framework in order to get superior and ideal outcomes.

### *C. Decision Tree model [3]*

In [3], The authors tested the appropriateness of many supervised machine learning models to forecast the probability of a fraudulent transaction using an unbalanced dataset. Sensitivity, accuracy, and time were employed by the authors as the main variables in reaching their specific result. Since accuracy as a parameter does not provide an accurate response and is insensitive to data that is imbalanced, it was not utilized. The authors concluded that the Decision Tree mode is the most appropriate model for forecasting these types of frauds. Decision Tree (79.21%), Random Forest (78.22%), KNN (81.19%), LR (69.31%), and NB (85.15) all have accuracy rates.

Merits-Credit card fraud is a contemporary problem, and the authors concluded that the Decision Tree model is the most appropriate model for forecasting credit card fraud.

Demerits-The resampling approaches can be applied to the specific datasets being used by researchers in this sector. By lowering the dataset imbalance ratio, this method contributes to improved classification outcomes.

### *D. Optimized Light Gradient Boosting Machine [4]*

In [4], The authors have proposed a novel approach to detect credit card fraud using an optimized light gradient boosting machine (OLightGBM). Two real-world data sets were utilized by the author in several studies. The recommended method beat the other methods and produced the best results in terms of precision (97.34%), accuracy (98.40%), area under the receiver operating characteristic curve (AUC) (92.88%), and F1-score (56.95%). The writers employed OLightGBM.

Merits- By using this method, the imbalance ratio of the datasets is reduced, leading to improved classification outcomes.

Demerits- The new methods and algorithms may be utilized by researchers in this sector.

### *E. Convolutional Neural Networks (CNN) [5]*

In [5], Authors have proposed a model for credit card recognition problems that outperforms the state-of-the-art machine learning and deep learning techniques. The authors have also carried out tests where they employed deep learning algorithms and balanced the data in an effort to lower the false negative rate. The following techniques were used: XG Boost, Decision Tree, Random Forest, Support Vector Machine, Logistic Regression, and Extreme Learning Method. The analysis of the study project shows the improved results obtained, with optimized f1-score, accuracy, precision, and AUC curve values of 99.9%, 85.71%, 93%, and 98%, respectively.

Merits- By using this method, the imbalance ratio of the datasets is reduced, leading to improved classification outcomes.
Demerits- This approach fails when the dataset is not balanced. Subsequent studies in this field could investigate the use of other state-of-the-art deep learning methods to improve the performance of the model proposed in this work.

### *F. SMOTE sampling along with Logistic Regression [6]*

In [6], The writers have analyzed the various references on fraud translation and compared the five strategies. The authors tested a number of sample techniques and determined which one produced the best results in terms of accuracy. They recommended using SMOTE sampling in conjunction with logistic regression to detect credit card fraud. The authors made use of K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Naive Bayes, ROSE, Unbalanced Dataset, under

sampling, oversampling, Receiver Operator Characteristics (ROC), ROSE, SMOTE, and Logistic Classifier. Logistic regression outperforms the other five techniques with a precision of 99.99% and an accuracy of 97.04%. Merits- It is recommended to use SMOTE sampling in conjunction with logistic regression to detect credit card fraud.

Demerits-Since credit card thieves are able to get past current security measures; credit card fraud is a relevant and serious problem in both offline and online transaction systems.

*G. Predictive classification model by hybridizing [7]*

In [7], because it offers a higher prediction performance, an ensemble of five different algorithms makes up the predictive categorization model that the authors have proposed. They used a variety of classification methods, such as K-Nearest Neighbor (K-NN), Multi-layer Perceptron (MLP), Extreme Learning Machine (ELM), Random Forest (RF), and Bagging classifier. It is found that the proposed classification model has a significantly increased predicted accuracy percentage of 83.83%.

Merits-The suggested model, there is a decrease in fraud detection error and an increase in fraud prediction rate.

Demerits- The availability of real-time datasets affects the fraud detection limitation.

*H. BiLSTM- Max Pooling-BiGRU-Max Pooling [8]*

In [8], SMOTE and random oversampling are unpromising techniques used by writers to create machine learning classifiers. Random under sampling is also used. The BiLSTM-MaxPooling-BiGRU model, which is based on bidirectional LSTM and GRU, was developed by the authors. To obtain precision, they employed the following algorithms. Decision trees, logistic regression, Ada boosting, random forests, and naïve bases. The Author model's results show that SMOTE 90%, Random over sampling 91.37%, and Random under sampling 90%.

Merits-Because SMOTE and random oversampling don't seem to be very accurate, we thought of utilizing a deep learning model to outperform machine learning techniques.

Demerits- Since the authors used other models to develop their model, it is difficult to understand and utilize.

*I. The Isolation Forest with the help of H2O.ai [9]*

In [9], Developing a highly efficient, autonomous fraud detection classifier that can identify fraudulent credit card transactions is the aim of this research project. Researchers have put out a number of ideas and methods for identifying fraud, as well as the use of different algorithms to identify fraudulent tendencies. In this research, we investigate the machine learning method known as Isolation Forest, which trains the system using H2O.ai. When using the isolation forest to detect fraud in credit card transactions, AUCPR discovered that the prediction accuracy of this proposed classification model, which lowers the fraud detection error, was 98.72%.

Merits-In a fraud detection model, the authors shows that their strategy is 98.72% efficient, which is much better than existing fraud detection techniques.

Demerits-The lack of a balanced dataset and its scarcity are the only factors impeding the efficacy of the fraud detection system.

*J. Gradient Boosting Tree (GBT)model [10]*

In [10], The authors' work states that the Gradient Boosting Tree (GBT) model of real-time credit card fraud detection on streaming Card-Not-Present transactions (CNP) searches a variety of card transaction attributes. Numerical, manually constructed numerical, category, and textual features are combined to create a feature vector that will be used as a training example. Two primary topics of this research are automated training dataset creation technique based on sliding windows and character-level word embedding. The name of the merchant can be utilized as a distinctive feature to identify fraudulent activity, and character-level word embedding is required to map the name to a vector of real numbers. Three metrics encoded features, aggregated and encoded feature, embedding, and aggregated and encoded feature—were employed in this experiment. The metrics False-Positive Rate (FPR), recall, precision, and Area under Curve (AUC) are used to assess experiments. Recall is enhanced by 0.029% and AUC for fraud detection is increased by 0.028% and 0.029%, respectively, when the training set is slid.

Merits-In a fraud detection model, the authors show that their strategy is 98.72% efficient, which is much better than existing fraud detection techniques.

Demerits-We are able to obtain each merchant name by employing the character level word embedding technique. We have created an extract feature vector for the automatic production of the transaction dataset using the GBT model.

*K. Random Under-sampling (RUS) [11]*

In [11], The behavior of scams and valid transactions is continually changing, according to authors. Furthermore, the credit

card data is highly skewed, which makes it difficult to predict fraudulent transactions. In this study, three different dataset proportions were used, and skewed datasets were treated with the random under-sampling technique. In this work, three machine learning techniques are utilized: K-Nearest Neighbor, Naïve Bayes, and Logistic Regression. Using the random under sampling method (RUS), the performance of these algorithms is tracked and analyzed to determine how well they distinguish and categorize fraudulent and non-fraudulent transactions from the credit card dataset and to determine whether or not their performance has improved. Python is used for the analysis, and F-measurement, area under curve, precision, sensitivity, specificity, and accuracy are used to determine how well the algorithms perform. In comparison to Naïve Bayes (NB) and K-Nearest Neighbor (KNN), Logistic Regression (LR) had the best performance for all data proportions based on these metrics.

Merits-There are various resampling techniques, which can be applied a skewed or imbalanced data to produce better results.

Demerits-Using random under sampling is that some information could be lost.

### *L. Online Boosting with EFDT [12]*

In [12], The performance of single classifiers can be improved, according to the authors, by using classifier ensembles in data mining or data stream mining. In order to attain high success in prediction with almost no growing memory and time costs, this work suggests an Online Boosting (OLBoost) strategy that first employs the Extremely Fast Decision Tree (EFDT) as a base (weak) learner, then assembles them into a single strong online learner.

Merits- To lower memory use without compromising prediction accuracy.

Demerits- Online tutoring that goes beyond the UCSD-FICO credit card dataset, utilizing other base learners.

### *M. D-AMWSPLAdaboost [13]*

In [13], As to the authors, the self-paced learning chosen in this paper is adaptive hybrid weighted self-paced learning, which modifies the base learner's election strategy in the Adaboost algorithm and improves the objective function of the algorithm. This paper's choice of a self-adaptive threshold finding approach can effectively reduce the impact of human experience on model training. D-AMWSPL Adaboost outperformed other algorithms in terms of convergence speed, increasing the AUC value and F1score in the credit card fraud dataset by 1.41% and 0.47%, respectively, over the classical Adaboost approach. D-AMWSPLAdaboost outperformed the Adaboost algorithm on other datasets, with the largest improvement of 17.01% in terms of F1 score and 13.79% in terms of AUC value. As a result, the studies demonstrate that the suggested D-AMWSPLAdaboost method may successfully be used to identify credit card fraud and improve the resilience of the conventional Adaboost technique.

Merits- Compared to other SPL algorithms, the adaptive threshold iterative parameter algorithm of the SPL algorithm utilized in this research makes parameter selection easier.

Demerits- In order to safeguard consumers' personal information, the authors desensitized the qualities from V1 to V28 in advance, so they do not have unique names. Nevertheless, they conducted further tests using a publicly accessible dataset of credit card transactions.

### *N. The AdaBoost Algorithm [14]*

In [14], One of the most popular, successful, and effective methods for detecting fraud in credit card transactions is machine learning, which employs regression and classification algorithms. the comparison of AdaBoost, KNN, Random Forest, and all three algorithms We discover that: KNN's accuracy is lower than Random Forest's, which is roughly equal to Ada Boost's with a tiny variation. Based on the outcome analysis, we can say that the AdaBoost Algorithm is doing as well as it can to help us achieve our goal of detecting frauds. We claim that while these algorithms can help reduce fraud behavior and transactions to some degree, they cannot totally eliminate or prevent it.

Merits- The authors conclude that the Adaboost algorithm is the most effective in achieving our primary goal, which is the identification of credit card fraud.

Demerits-They discovered that we cannot rely solely on machine learning algorithms to provide the highest accuracy and performance because, in some cases, the results are not accurate or adequate. This is also true for Random Forest, KNN, and Ada Boost.

### *O. Generative Adversarial Networks (GANs) [15]*

In [15], One of the most popular data generation approaches is generative adversarial networks, or GANs, because of its use in huge data contexts. The aim of this work is to present a summary of data augmentation using many GAN variants for credit card fraud detection. GANs employ two different neural networks: the Generator

and Discriminator (D) neural networks. Almost closely matching the real data, the G's task is to input a random noise vector into false data. The D's dual purpose is to gather real samples and to act as a teacher, evaluating the output's effectiveness and

confirming the authenticity of the data.

Merits- The discrete form may be changed into a continuous one using this approach. Trip advisor and Yelp datasets were employed in this study because they are more trustworthy than datasets with human labeling.

Demerits- Even with the significant advancements in GAN approaches, these models are still not fully capable of handling credit card fraud. As a result, future work in the financial industry should concentrate more on developing strong extra GANs for current models.

## IV. PERFORMANCE METRICS AND ANALYSIS OF VARIOUS ALGORITHMS

TABLE I

ACCURACY OF VARIOUS ALGORITHMS

| Sr.No | Methodology Used | Accuracy |
|---|---|---|
| 1 | Federated Learning and ANN [1] | NA |
| 2 | Clustering and similarity-based selection (SBS) [2]<br>ANN, LR, KNN, NB | 98.9<br>98.6<br>96.6<br>98.4 |
| 3 | Decision Tree model [3]<br>Decision TreeKNN<br>Logistic RegressionRandom Forest<br>NB | 79.21<br>81.19<br>69.31<br>78.22<br>85.15 |
| 4 | Optimized Light Gradient Boosting Machine [4]<br>Light GBM Cataboost Proposed Approach | 90.62<br>87.86<br>92.88 |
| 5 | Convolutional Neural Networks (CNN) [5]Convolutional neural networks (CNN) Baseline (BL)<br>RF SVM KNN<br>DT(Decisiontree) Logistic Regression | 96.34<br>99.72<br>99.92<br>99.93<br>99.91<br>99.93<br>99.91 |
| 6 | SMOTE sampling along with Logistic Regression [6]<br>Naive Bayes Random Forest<br>K-Nearest Neighbours Support Vector MachinesLogistic regression | 96.93<br>98.71<br>96.89<br>98.14<br>97.04 |
| 7 | Predictive classification model by hybridizing [7]<br>Extreme Learning Machine (ELM)Multi-layer Perceptron (MLP)<br>Bagging classifier<br>KNN<br>Random Forest Proposed Model | 78.75<br>80.38<br>80.87<br>81.43<br>81.92<br>83.83 |
| 8 | BiLSTM- MaxPooling- BiGRU-MaxPooling [8]<br>Naïve base Voting Random Forest<br>logistic regressionAda boosting Decision Tree | 71.0<br>73.1<br>74.9<br>80<br>77.8<br>69.15 |
| 9 | The Isolation Forest with the help of H2O.ai [9]<br>AUCPR (Area Under Precision-Recall curve) | 98.72 |
| 10 | Gradient Boosting Tree (GBT) model [10]<br>Encoded Agg. + Encoded Emb. + Agg. + Encoded | 0.960<br>0.964<br>0.968 |

| | | |
|---|---|---|
| 11 | Random Under-sampling (RUS) [11]<br>Logistic regression Naı¨veBayes<br>K-nearest neighbours | 0.918<br>0.829<br>0.633 |
| 12 | Online Boosting with EFDT [12] | NA |
| 13 | D-AMWSPLAdaboost [13] | NA |
| 14 | The AdaBoost Algorithm [14]<br>AdaBoost Algorithm | 0.999 |
| 15 | Generative Adversarial Networks (GANs) [15] | NA |

## V. OPEN RESEARCH ISSUES

This research paper provides a comprehensive review of supervised machine learning techniques employed in the detection of credit card fraud. Despite significant advancements in this field, several open research issues persist, necessitating further investigation to enhance the efficacy and robustness of fraud detection systems. This paper identifies and discusses key research gaps and suggests potential avenues for future exploration in this domain. The following some issues regarding open challenges

i. Imbalanced Data Handling-Future research should focus on developing more effective techniques for handling class imbalance, such as advanced sampling methods, cost-sensitive learning algorithms, and synthetic data generation approaches.

ii. Feature Selection and Engineering-Future research should explore novel feature selection methods tailored to fraud detection scenarios, including automatic feature extraction techniques and domain-specific feature engineering strategies.

iii. Model Interpretability and Explainability-Despite the high predictive performance of complex models, such as deep neural networks, their lack of interpretability limits their practical utility in real-world applications. Future research should focus on developing interpretable models and post-hoc explanation techniques that enable stakeholders to understand the rationale behind fraud predictions.

iv. Real-time Detection and Scalability-The timely detection of fraudulent transactions is paramount for minimizing financial losses and mitigating risks. However, many existing fraud detection systems suffer from latency issues and scalability limitations, particularly when processing large volumes of transactional data in real-time. Future research should focus on developing efficient algorithms and scalable architectures for real-time fraud detection, leveraging techniques such as stream processing, distributed computing, and parallelization.

## VI. CONCLUSION

In conclusion, this investigative review has provided a comprehensive examination of supervised machine learning techniques in the context of detecting credit card fraud. Through an analysis of various algorithms, methodologies, and datasets, it is evident that supervised machine learning presents a powerful arsenal for fraud detection in the financial sector. The reviewed techniques, including logistic regression, decision trees, random forests, support vector machines, and neural networks, have demonstrated varying degrees of effectiveness in identifying fraudulent transactions. This research paper contributes to the broader understanding of supervised machine learning techniques for detecting credit card fraud and provides valuable insights for researchers, practitioners, and policymakers striving to combat financial crimes in an increasingly digital landscape. As the field continues to evolve, interdisciplinary collaboration and innovation will be crucial in developing robust, adaptive, and ethically sound solutions to mitigate the risks associated with fraudulent activities in the financial domain.

## REFERENCES

[1] Bin Sulaiman, R., Schetinin, V. & Sant, P. Review of Machine Learning Approach on Credit Card Fraud Detection. Hum-Cent Intell Syst 2, 55–68 (2022), Springer.

[2] Hadeel Ahmad, BassamKasasbeh, BalqeesAldabaybah, EnasRawashdeh “Class balancing framework for credit card fraud detection based on clustering and similarity-based selection (SBS)”, Int. j. inf. tecnol. 15, 325–333 (2023). https://doi.org/10.1007/s41870-022-00987-w, Springer.

[3] SamidhaKhatri, AishwaryaArora, ArunPrakashAgrawal “Supervised Machine Learning Algorithms for Credit Card Fraud Detection: A Comparison “ 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2020, pp. 680-683,

[4] Altyeb Altaher Taha, Sharaf Jameel Malebary “An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine” in IEEE Access, vol. 8, pp. 25579-25587, 2020.

[5] Muhammad Ramzan ,Fawaz Khaled Alarfaj , Iqra Malik, Hikmat Ullah Khan , Naif Almusallam, And Muzamil Ahmed “Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms” in IEEE Access, vol. 10, pp. 39700-39715, 2022.

[6] J. V. V. Sriram Sasank, G. Ram sahith, K.Abhinav, Meena Belwal ” Credit Card Fraud Detection Using Various Classification and Sampling Techniques “2019 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2019, pp. 1713-1718.

[7] Debachudamani Prusti, Santanu Kumar Rath "Fraudulent Transaction Detection in Credit Card by Applying Ensemble Machine Learning techniques" 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-6.
[8] Hassan Najadat, Ola Altiti , Ayah Abu Aqouleh , MutazYounes "Credit Card Fraud Detection Based on Machine and Deep Learning "2020 11th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 2020, pp. 204-208.
[9] Meenu, Swati Gupta, Sanjay Patel, Surender Kumar, Goldi Chauhan "Anomaly Detection in Credit Card Transactions using Machine Learning "International Journal of Innovative Research in Computer Science & Technology (IJIRCST) ISSN:2347-5552, Volume-8, Issue-3, May 2020.
[10] Ali Ye, silkanat(B), Barı, s Bayram, Bilge K¨oro˘glu, and Se, cil Arslan, "An Adaptive Approach on Credit Card Fraud Detection Using Transaction Aggregation and Word Embeddings" © IFIP International Federation for Information Processing 2020, Published by Springer Nature Switzerland AG 2020, I. Maglogiannis et al. (Eds.): AIAI 2020, IFIP AICT 583, pp. 3–14, 2020.
[11] Fayaz Itoo, Meenakshi, Satwinder Singh, "Comparison and analysis of logistic regression, Naı¨ve Bayes and KNN machine learning algorithms for credit card fraud detection" © Bharati Vidyapeeth's Institute of Computer Applications and Management 2020.
[12] Aye Aye Khine, Hint Wint Khin, "Credit Card Fraud Detection Using Online Boosting with Extremely Fast Decision Tree" 2020 IEEE Conference on Computer Applications (ICCA), Yangon, Myanmar, 2020, pp. 1-4.
[13] Wangning ,Siliangchen ,Songyilei ,And Xiongbinlia "AMWSPLAdaboost Credit Card Fraud Detection Method Based on Enhanced Base Classifier Diversity" @ IEEEAccess, DigitalObject Identifier 10.1109/ ACCESS .2023.3290957.
[14] Rakhi Arora, Nitin Dixit, Gaurav Dubey "A Review on Fraud Detection of Credit Cards Through Machine Learning Algorithms", Journal of University of Shanghai for Science and Technology, ISSN: 1007-6735, Volume 25, Issue 01, January – 2023.
[15] Emilija Strelcenia, Simant Prakoonwit "A Survey on GAN Techniques for Data Augmentation to Address the Imbalanced Data Issues in Credit Card Fraud Detection", Mach.Learn.Knowl.Extr.2023,5,304–329.