DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

Machine Learning for Identifying Fraud in Credit Card Transactions

Mohammad.A.Abbas[1], Bilal Ghazal[2] , Ahmad Ghandour[3]
[1,3] Computer and Commuication Department, Faculty of Engineering, Islamic University of Lebanon
[2]Faculty of Sciences, Lebanese University
[1]ma72161@net.iul.edu.lb., [2]bilal.ghazal@ul.edu.lb, [3]Ahmad.ghandour@iul.edu.lb

Abstract— *Online payment methods for e-commerce and many websites in various fields have increased significantly. Therefore, credit card frauds are easy targets, and their rate is on the rise which poses a major problem for online payments. The basic concept is to examine consumers' purchasing histories to extrapolate their typical behavior patterns, classify cardholders into different groups, and then attempt to detect credit card fraud. Credit card fraud detection based on a machine learning model uses a combination of supervised and unsupervised learning techniques such as Random Forest, Decision Tree, Logistic Regression, and Extreme Gradient Boost. We used the synthetic minority oversampling (SMOTE) technique to balance the dataset. The model is trained on a large set of data related to credit card transactions and uses features such as transaction amount, transaction location, and time of day to identify patterns and anomalies in the data that indicate fraudulent activity. Our goal is to build a model based on machine learning technology that detects and analyzes online shopping fraud Detecting fraud in credit card systems is crucial to protecting consumers from financial losses and maintaining the integrity of the financial ecosystem after collecting Creditcard.csv data. With the help of several algorithms, including Random Forest (RF) algorithm accuracy reached 99%, Logistic Regression (LR) algorithm accuracy reached 97%, and Decision Tree (DT) algorithm accuracy reached 99% Researchers provide a comprehensive method for identifying fraud in credit card transactions Precision Recall F1-score. The proposed system includes four main steps: pre-processing, classification using the algorithm, and checking whether the transaction is fraudulent or not.*

Index Terms— *Detecting fraud, machine learning, Decision Tree, Random Forest, Logistic Regression.*

I. INTRODUCTION

A credit card is a plastic card with personal details like a photo or signature on it., the person's name, and a very important unique card number. Charges for purchases and services are charged to the customer's account which will be debited regularly. today , In addition to ATMs, readers for transactions in online stores and banks, and swiping machines for mobile payments are also being used to read card information.. Card security is determined by the physical security of the card and the confidentiality of the credit card number [1]. A credit card usually is intended for cardholder consumers. For example, it provides the consumer with the possibility to pay later at a certain time by transferring it to the next bill [2]. Card data is read by Automated Teller Machines (ATMs), Point-Of-Sale (POS) readers, and other devices. Every day, credit cards are used for online and offline shopping for the purchase of goods or services. They provide cashless shopping for online and offline shopping with a buy now and pays later feature. With such prevalent use of credit cards, credit card

Received 12/August/2023; Accepted 15/November/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

fraud is relatively on the rise. Before getting a credit card, it is necessary to know every element of your card and its function.

The financial services sector has shown considerable interest in studying methods for detecting credit card theft. Financial organizations lose millions of dollars each year due to fraud. and con artists are continuously on the lookout for new ways to steal money[3]. Credit card fraud detection is challenging. The fraudulent transactions in each transaction dataset can be isolated using feature selection, feature classification, and clustering methods. There are a lot of elements that contribute to the probability of a fraudulent transaction, such as the total amount spent, the buying habits of the consumers, and the outcomes of previous investigations into similar frauds. Because of this, the following components are essential for any method of fraud detection [4].

1- It must effectively identify fraudulent financial dealings.
2- It needs to spot the scam while it's happening.
3- The system must ensure that legitimate purchases are not mistakenly flagged as fraudulent.

Card Issuer's Name: On the face of your credit card, the name of the issuing bank or financial institution will be printed in the uppermost corner. Credit card applications are made at this bank. Name of the Credit Card: Each financial institution issues its own unique set of credit cards, therefore it's reasonable now to give each one its own name. Therefore, the names of all bank-issued credit cards are unique. Credit cards have two names: the issuer and the card. To clarify, XYZ Bank issues the Ultimate card depicted above. This card's full title then becomes XYZ Bank Ultimate Credit CardCard Network The financial organizations that operate behind the scenes to enable all credit card payments are known as credit card networks. The networks that issue credit cards determine which stores accept their cards. They provide a link between the card company and stores that accept credit cards. The five most popular credit card companies are listed below (Visa, MasterCard, American Express, Diners Club, and Rupa). Credit Card Number: This number, which is typically 16 digits long but is only 15 digits for American Express cards, is used to identify both the card's issuing bank and its associated card network. When making a purchase with a credit card, the issuing bank, the card network, and other relevant information are requested via the credit card number.

II. FRAUD DETECTION

Nowadays online transactions are growing as new payment technology. In 2018, Losses from credit card theft in London were expected to be $844.8 million. To mitigate these losses, fraud must be prevented or detected. So there are many algorithms used to detect fraud, especially the artificial neural network giving better performance [5]. Fraud companies and fraudsters must adapt to detect system innovations. Fraud detection and prediction demand precise techniques. A trustworthy system should detect new fraud before it is deployed [6]. Credit card fraud detection, which analyses all card transactions and looks for fraudulent activity among thousands of genuine transactions, is one of the hardest challenges. Fraud detection infrastructure requires knowledge of fraud frequency. The main challenge is creating a reliable, accurate procedure [7], [8].

III. PROBLEM WITH CARD FRAUD DETECTION

It has become imperative to develop effective systems that can detect fraudulent transactions in real time. The main The goal is to create a machine learning model that can detect fraudulent credit card transactions in a huge dataset of credit card transactions. The model must be able to identify patterns and anomalies in the data that indicate fraudulent activity and provide a high level of accuracy

Received 12/August/2023; Accepted 15/November/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

in predicting whether a transaction is fraudulent. The model must also be able to handle imbalanced datasets, where the number of fraudulent transactions is much lower than the number of legitimate transactions. The goal of the project is to provide a reliable The development of a fraud detection system would help reduce credit card losses and improve the confidence of consumers in the security of their transactions.

IV. LITERATURE REVIEW

To spot credit card fraud [9], a model employing backpropagation and ANN has been described. Information such as customer names, transaction IDs, and purchase timestamps are stored. A random sample 20% of the data was used for testing and validation, with the remaining 80% used for training data purposes. Real-time data fraud detection using the proposed technology was evaluated and the model was able to boost the accuracy to 99.96%.

have utilized a wide variety Logistic regression, naive Bayes, and other machine learning techniques approaches multilayer perceptron, and radial basis function networks (Varmedjaet al., 2019). Different techniques for identifying potentially fraudulent credit card purchases are tested using data collected from a European credit cardholder dataset hosted in the Kaggle repository. Oversampling and feature selection methods were utilized. Random Forest (RF) was shown to be the most precise method after being compared to other popular options. Due to the sensitive nature of personal information, doing research and developing algorithms utilizing real-world data is challenging.

Using Naive Bayesian (NB), K-Nearest Neighbors (KNN), and logistic regression (LR) classifiers [10] examine widespread credit card theft. The collection includes the details of 284,807 separate transactions completed by European cardholders. Each approach was used to process both raw data and sanitized data. The results indicated that the NB classifier was the most successful with hiher accuracy of (97.92%), the KNN classifier was the second-most effective (97.69%), and the LR classifier was the least effective (54.8%).

New method developed by [11] for enhancing the SVM algorithm's card recognition capabilities. The success of the SVM technique was heavily dependent on the input parameters and the training data. The Least Square LS-SVM and ensemble technique were developed and employed to foretell which cardholders would be late with their monthly payment. The technique was applied to the University of California Irvine UCI repository's financial datasets in Taiwan.

The authors [12] An unmoderated random forest algorithm is proposed to reduce the number of fraudulent transactions. Additionally, the algorithm is used to analyze credit card fraud detection. Furthermore, Bayesian networks create coordinated aperiodic graphs, which are further used in conditional probability distributions to create aperiodic graphs. The results show that the proposed algorithm based on stochastic structure outperforms its counterparts. The Table I shows a comparison between literature review in this paper.

Received 12/August/2023; Accepted 15/November/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

TABLE I. LITERATURE REVIEW

Title & Authors	Years	Algorithms used	Data set	Accuracy
"Survey of various techniques used for credit card fraud detection" (Agarwal, Iqbal and Mitra)	2020	ANN using back propagation	They divided the data set into 80% training and 20% test data	ANN 99.90
Credit card fraud detection using a new hybrid ML architecture (Malik et al)	2022	LR, RF, DT, XGBOOST, SVM, Adaboost and LGBM	They proposed seven different hybrid machine learning techniques LR, RF, DT, XGB, SVM, and NB, where the result was not good, (Adaboost and LGBM) were used, and they showed an accuracy of 97%.	Adaboost and LGBM 97%
"Credit card fraud detection ML methods" (Varmedja *et al.)*	2019	LR NB RF	The unbalanced European data set was used, and a comparison was made between the three models.	RF 95.50% best accuracy
"Credit card fraud detection using ML techniques: a comparative analysis" (Awoyemi, Adetunmbi and Oluwadare,)	2017	KNN NB LR	Use an unbalanced data set with the three algorithms	KNN 97.92%
"Classification of credit card default clients using ls-svm ensemble" (Lawi and Aziz,)	2018	SVM LR	An unbalanced data set was used applying the Least Square LS-SVM technique using	70.90%
"Credit card fraud detection using ML" (Safa and Ganga,)	2019	NB KNN LR	Three algorithms are used on an unbalanced data set,	LR 97.69%

V. MACHINE LEARNING (ML)

Machine learning (ML) is a branch of computer science that focuses on creating models and algorithms that let computers learn from data and predict or evaluate situations without having to be explicitly programmed. Machine learning necessitates the gathering and processing of data, selection of an appropriate method, and implementation of that algorithm in order to develop a model that can be used to make predictions or classification. The *Fig. 1* shown the type of Machine learning

Classification is a fundamental task in ML that involves predicting the class label of a given input sample based on a set of labeled training examples. The goal of classification is to learn a decision boundary that separates the classes in the input feature space. Classification has many practical applications in fields such as medicine, finance, marketing, and social sciences. Some examples of classification tasks include disease diagnosis, credit risk assessment, customer churn prediction, and sentiment analysis.

ML concepts include different techniques of supervised learning and unsupervised learning. These techniques involve the use of mathematical and statistical methods to extract patterns and ideas from large and complex data sets. Some of the basic concepts in ML include data preprocessing, feature engineering, model selection and evaluation, regularization, validation, and hyperparameter tuning

Received 12/August/2023; Accepted 15/November/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

Supervised learning is a machine learning technique that uses labeled datasets to train algorithms to reliably categorize input or predict results. The objective is to build a model that can predict the output properly based on the input data. When information is collected, it is either utilized for training or testing. By examining the training data, a function may be developed and used to test data for prediction and classification[13].

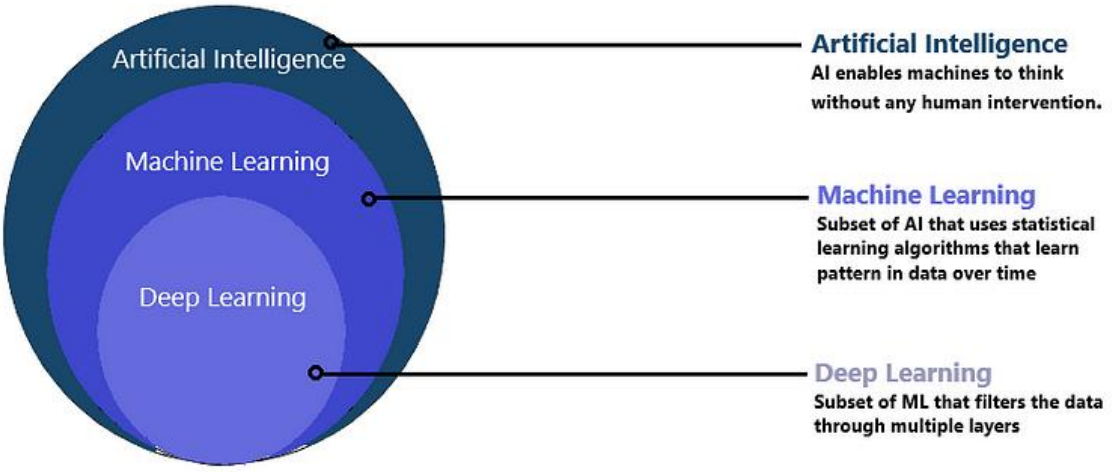

FIG. 1. MACHINE LEARNING IS CLASSIFIED.

A. Decision tree (dt)

A decision tree is a tree-based model that is used for both regression and classification applications. It divides the data into subsets according on the values of the input features and iteratively builds a tree to generate predictions based on the feature values. Decision trees are a popular choice for many machine learning problems since it's straightforward and simple to understand[14].As shown in *Fig. 2.* The confusion matrix of using the DT is shown in *Fig. 3.*

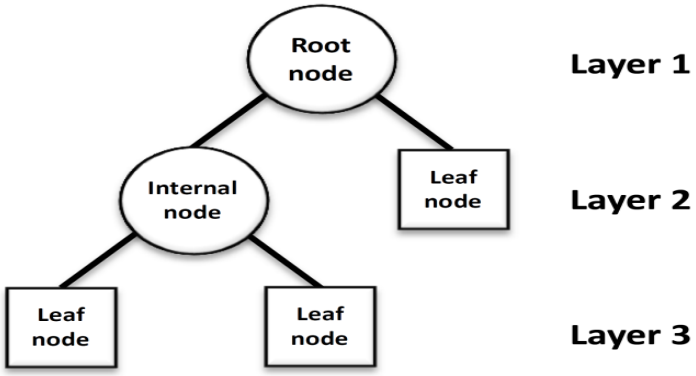

FIG. 2. DECISION TREE STRUCTURE.

Received 12/August/2023; Accepted 15/November/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

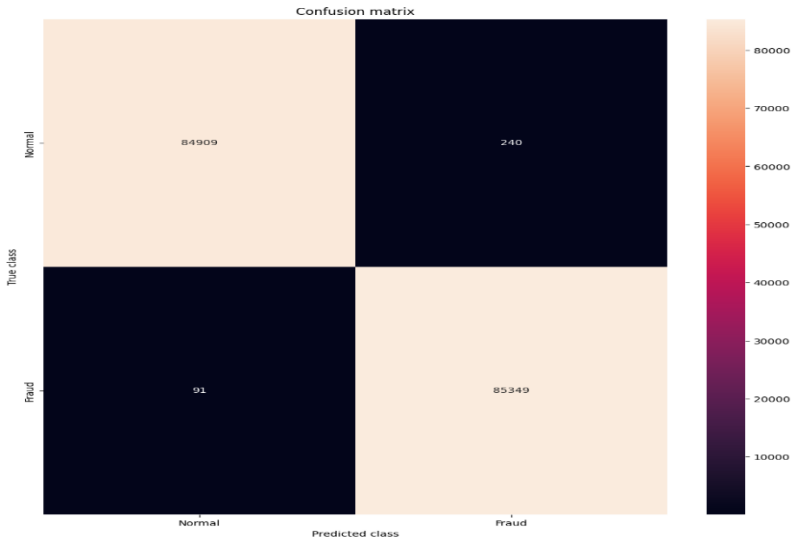

FIG. 3. CONFUSION MATRIX OF USING THE DT ALGORITHM.

B. Logistic regression (LR)

The model of logistic regression is trained by calculating the coefficients Minimizes the difference between predicted probabilities and actual binary response values in the training data. This is usually done by using maximum likelihood estimation or gradient descent optimization. The structure of logistic regression shown in *Fig. 4.* and The confusion matrix of using the (LR) is shown in *Fig. 5.*

Logistic regression has many practical applications in sectors such as medical, finance, and social sciences. It is commonly employed in predictive modeling, risk analysis, and decision-making[15].

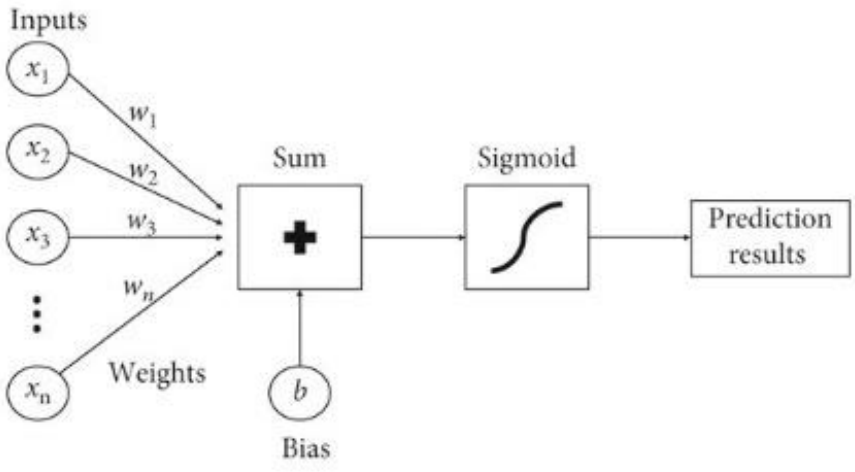

FIG. 4. LOGISTIC REGRESSION STRUCTURE.

Received 12/August/2023; Accepted 15/November/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

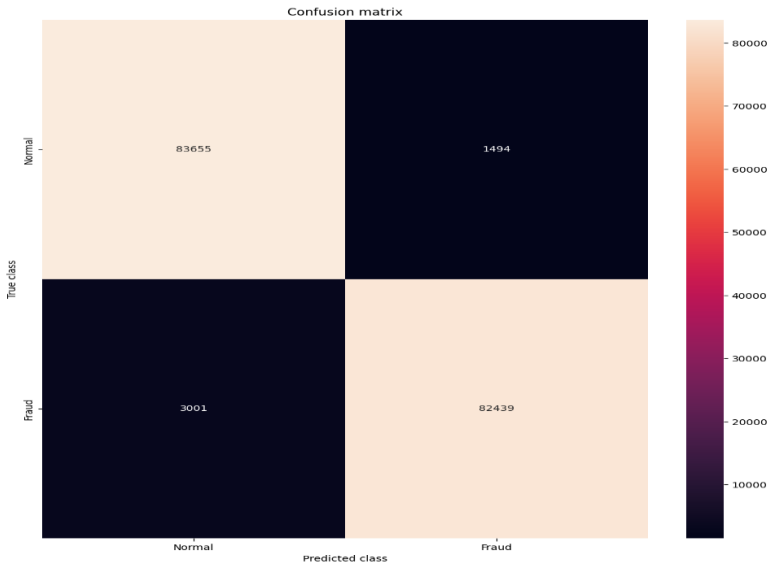

FIG. 5. CONFUSION MATRIX OF USING THE LR ALGORITHM.

C. Random Forest

For many machine learning applications, such as image classification, text classification, and regression analysis, random forest is growing as a popular approach. It has been shown to perform well on various benchmark datasets and can do many jobs at the state-of-the-art performance, the implementation of random feature selection contributes to reducing the correlation between the individual trees, which enhances the performance of the ensemble[16]. The structure of Random Forest shown in *Fig. 6.* and The confusion matrix of using the (RF) is shown in *Fig. 7.*

This method has provided us with high precision in identifying credit card fraud done both in person and online.

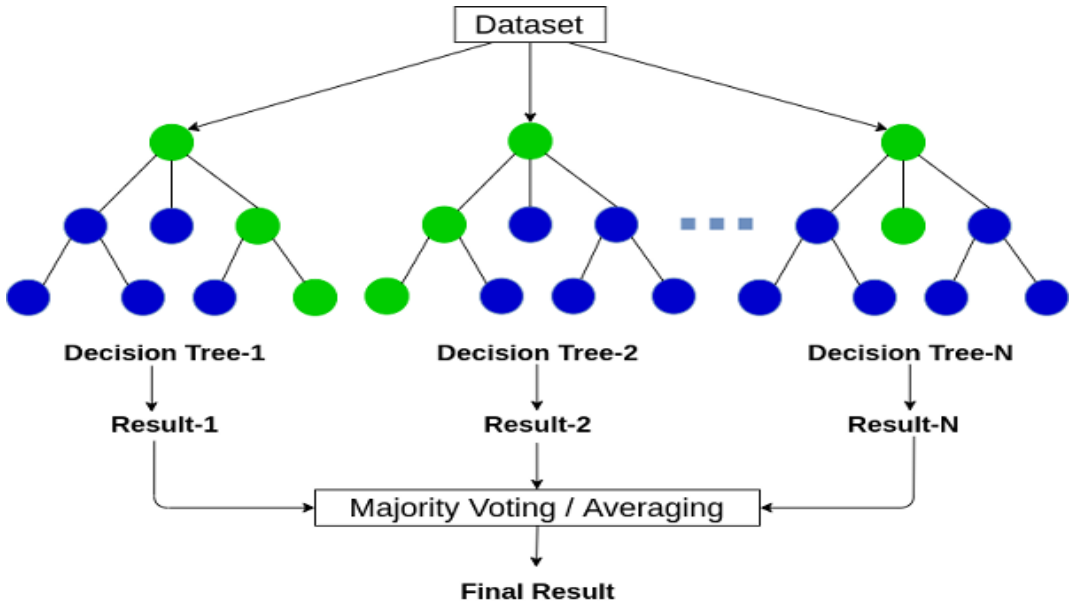

FIG. 6. RANDOM FOREST STRUCTURE.

Received 12/August/2023; Accepted 15/November/2023

ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

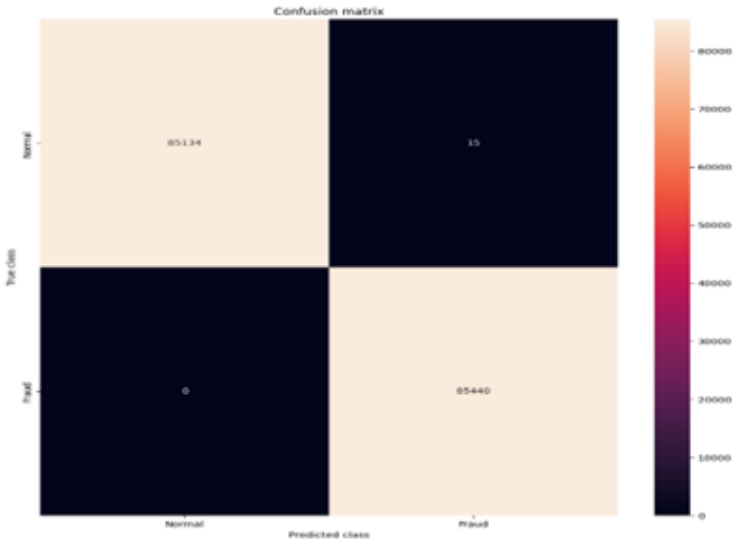

FIG. 7. CONFUSION MATRIX OF USING RF ALGORITHM.

VI. PROPOSED SYSTEM

The proposed model, as shown in *Fig. 8*, consists of four main steps: data collection, preprocessing, machine learning-based classification using algorithms, and dataset division into 70% training data and 30% testing data. The training data is utilized to train three ensemble classifiers, including Random Forest (RF), Logistic Regression (LR), and Decision Tree (DT), while the testing data is used to evaluate the models. Furthermore, the classification results are compared using various evaluation metrics, and the final step involves examining whether a transaction is fraudulent, as explained in detail below. Additionally, to overcome the class imbalance in the credit card dataset, the researcher employed the Synthetic Minority Over-sampling Technique (SMOTE). Class imbalance occurs when The number of examples from the minority group (fraudulent transactions) is much smaller than the number of instances from the majority group (legal transactions). This imbalance can result in biased models with poor performance in detecting the minority class. SMOTE generates synthetic samples of the minority class, thereby increasing its representation in the dataset and facilitating a more balanced approach.

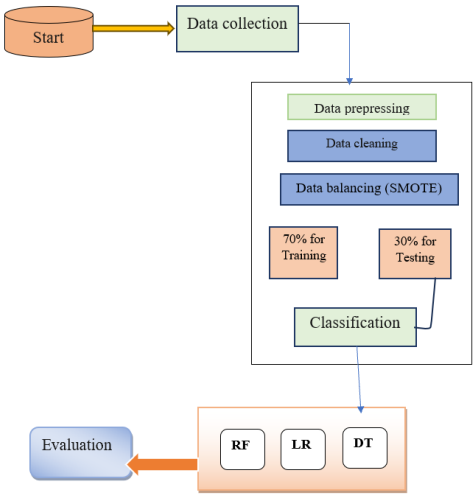

FIG. 8. THE PROPOSED SYSTEM.

Received 12/August/2023; Accepted 15/November/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

A. Collection Data

During the data collection phase, anonymized credit card transactions that are classified as fraudulent or genuine, and relevant credit card transaction data are obtained, including attributes such as transaction amount, merchant information, time, and user details. From the Kaggle website https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud/data Credit Card Fraud Detection, download the dataset. Data collection is very unbalanced. Contains only numeric input variables from the Payment Card Assurance (PCA) conversion.

B. Data Preprocessing

The next step is preprocessing, which includes handling missing values and cleaning up the obtained data. This stage makes that the data is in a format that will work for the next analysis and classification.

C. Data Classification

Throughout this stage, three common Machine Learning algorithms are used: Random Forest, Logistic Regression, and Decision Tree. These algorithms were chosen because of their ability to detect patterns and anomalies in large datasets. Each algorithm is trained on preprocessed data to discover the underlying patterns associated with Transactions that are both legal and illegal.

D. Determine the Validity of a Given Transaction

The trained models are used to categorize a particular transaction as fraudulent or non-fraudulent. The transaction attributes are examined as part of this classification process, and they are compared with the patterns that have been found through the training phase. The suggested system evaluates the transaction in detail, indicating whether it has the potential to be fraudulent or genuine.

PERFORMANCE EVALUATION MEASURES

Measures of evaluation are crucial in machine learning. They help identify which model or system is best for a given task by comparing the performance of various models or systems. Accuracy, precision, recall, and F1-score are a few of the standard evaluation metrics used in machine learning. As can be seen in the calculation of the confusion matrix, which is defined as a matrix in which the test results were dispersed by splitting them into two classes, Table II.

TABLE II. TYPE SIZES AND APPEARANCE

	Positive	**Negative**
Positive	True positive	True negative
Negative	False Positive	False Negative

A- Accuracy: Accuracy measures the degree to which a model's predictions are on average by calculating the ratio of accurately predicted cases to all instances. This equation is used to determine it. (1)[17] :

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad (1)$$

B- Precision: This indicator counts the proportion of positive instances that were predicted correctly out of all positive instances. It focuses on how accurate the predictions are. It is calculated using an equation. (2)[17]:

Received 12/August/2023; Accepted 15/November/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

$$\text{Precision}=\frac{TP}{TP+FP} \qquad (2)$$

C- Recall (Rc): The proportion of accurately expected positive cases out of all actual positive cases is known as the sensitivity or true positive rate. Focuses on reporting positive examples. in the equation (3)[17]:

$$\text{Recall}=\frac{TP}{TP+FN} \qquad (3)$$

D- F- Measure (F1): is a harmonic mean of recall and precision. It provides a balanced evaluation of a model's performance by combining precision and recall into a single metric. The results are as shown in the equation. (4)[17]:

$$F1=2*\frac{\text{Pre}*\text{Rc}}{\text{Pre}+\text{Rc}} \qquad (4)$$

VII. EXPERIMENT AND RESULTS

This section shows the result of data analyzed using a correlation matrix is a table that displays the correlation coefficients between two variables. Correlation coefficients range between and -1 to 1, with -1 indicating strong negative correlation, 1 indicating strong positive correlation, and 0 indicating no correlation. As shown in *Fig. 9*.

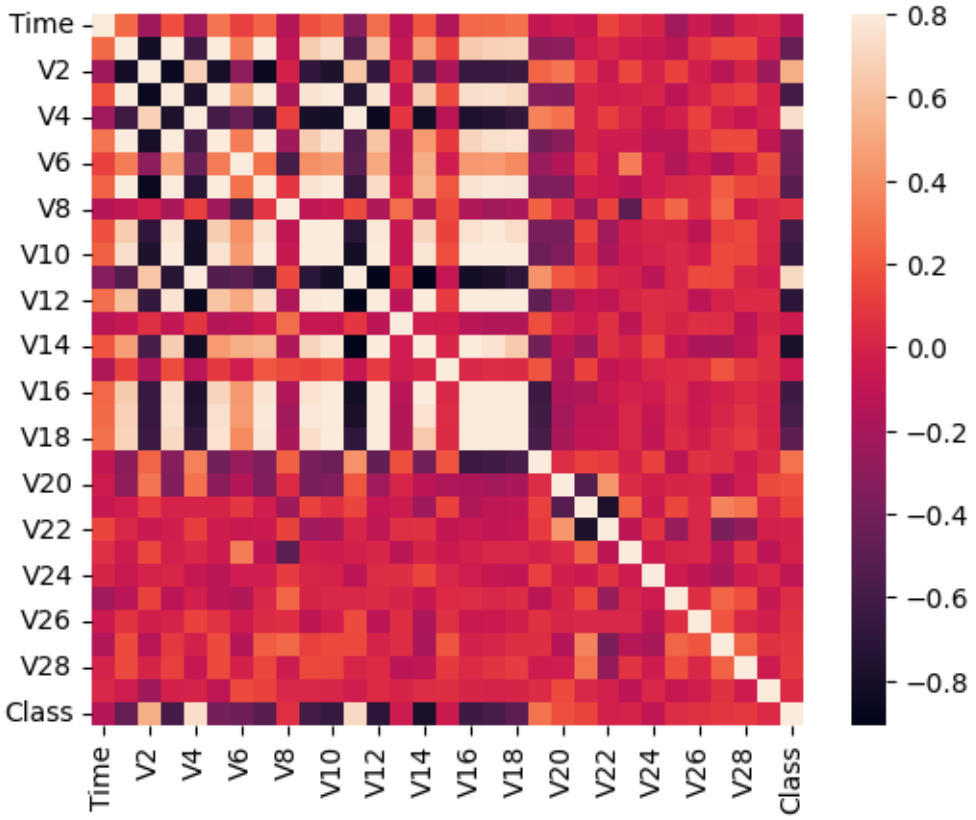

FIG. 9. CORRELATION MATRIX FOR THE PROPOSED SYSTEM FEATURES.

Received 12/August/2023; Accepted 15/November/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

By examining the correlation matrix, we can gain insights into the relationships between different features in the dataset. These correlations can help identify potential features that are strongly related to fraudulent transactions or variables that are highly correlated with each other.

The second stage consists of cleaning the data from missing values and handling outliers. Next, the dataset is split into 70% for training purpose, while the remaining 30% are reserved for training and testing using Oversampling of Minorities Synthetically (SMOTE) technique. After applying SMOTE in the proposed system, we observe a significant change in the distribution of the dataset, particularly in the fraud and valid data categories. The results of the distribution of fraud before (SMOTE) are shown in *Fig. 10* and the distribution of valid data after (SMOTE) is shown in *Fig. 11*.

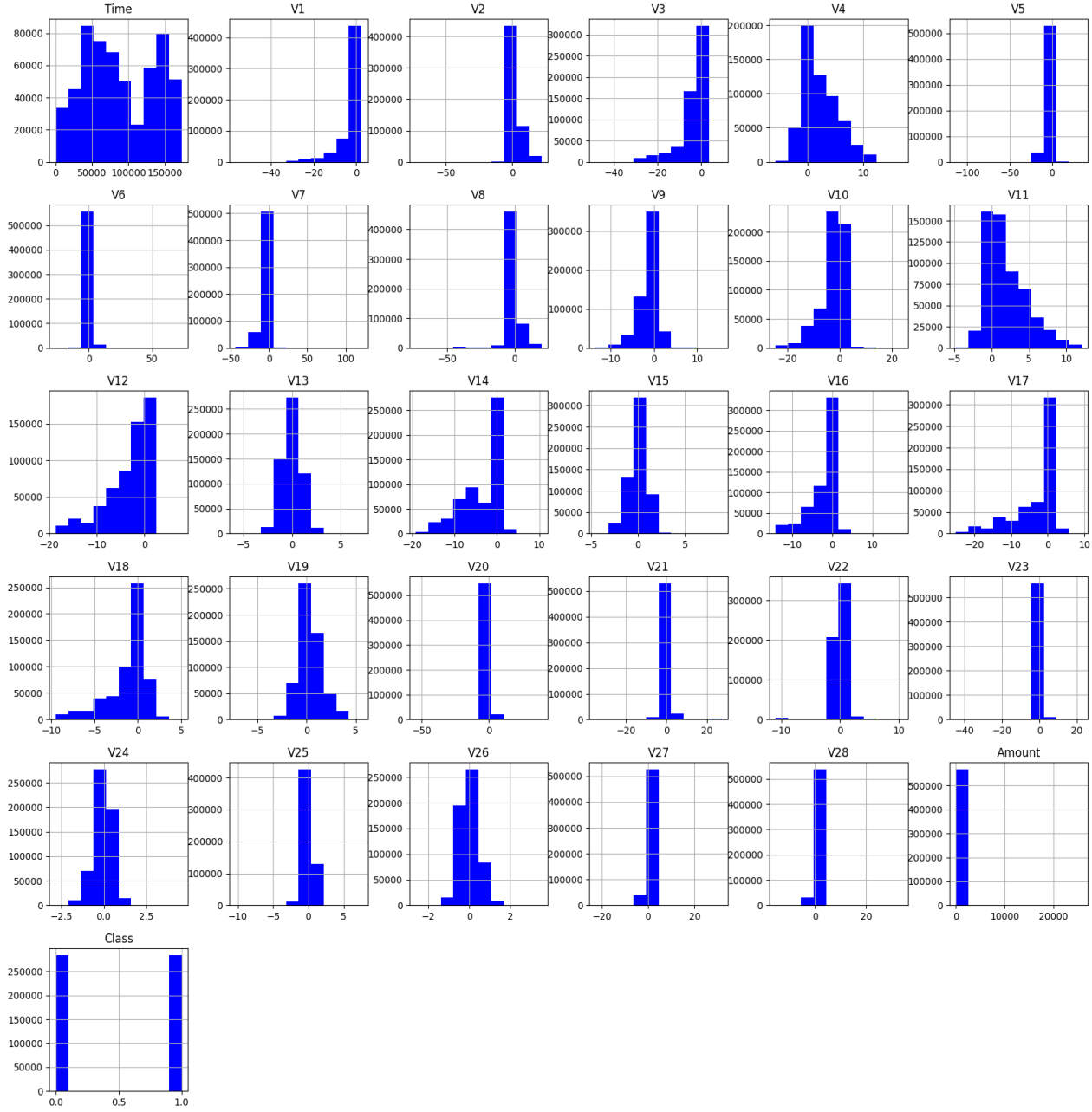

FIG. 10. THE DATA DISTRIBUTION BEFORE SMOTE.

Received 12/August/2023; Accepted 15/November/2023

ISSN (Print) 1811-9212 *ISSN (Online) 2617-3352*

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

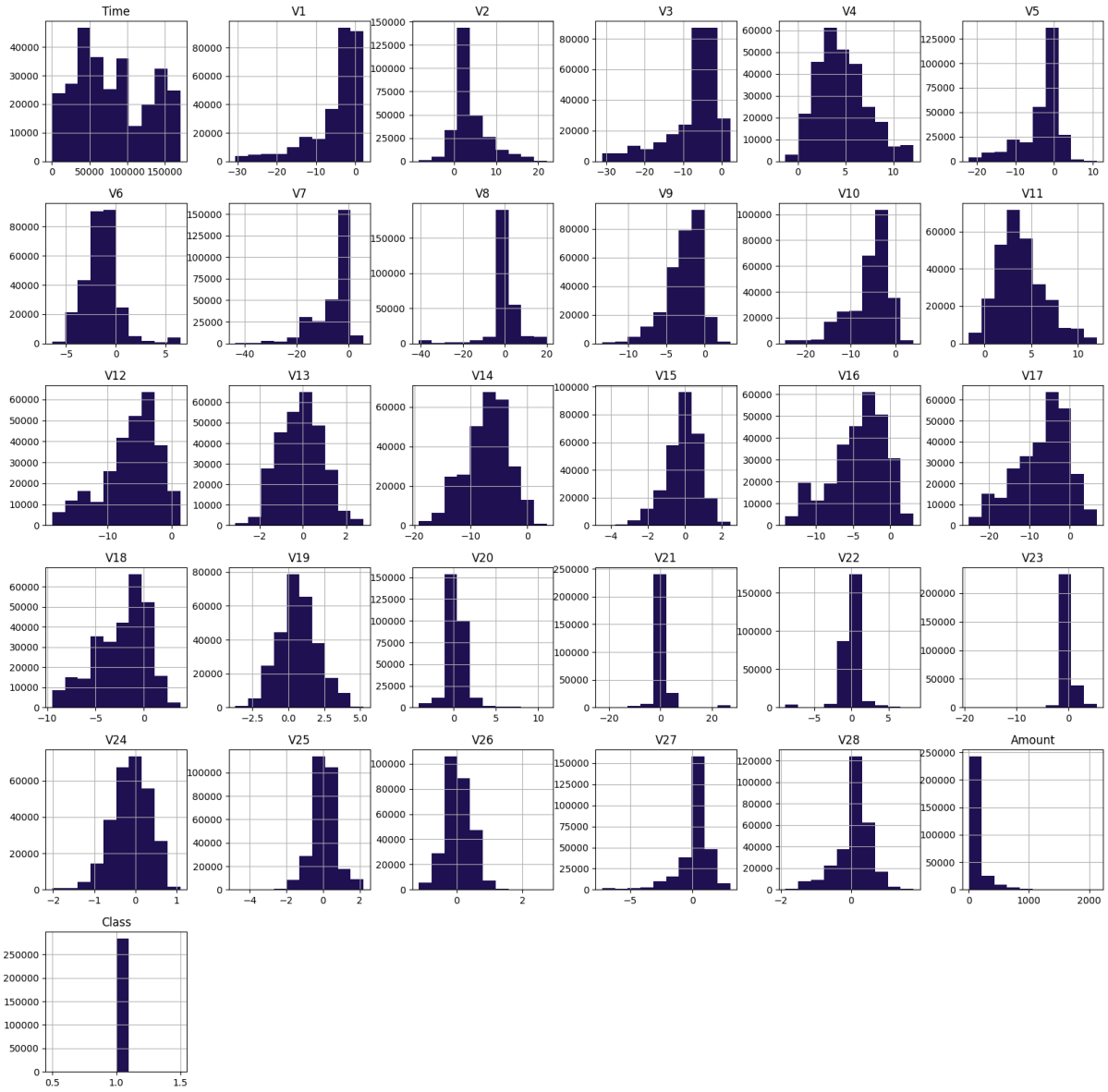

FIG. 11. DATA DISTRIBUTION AFTER SMOTE **ERROR! REFERENCE SOURCE NOT FOUND.**.

From the comparison in Table III. We can observe that all four algorithms achieve high accuracy in classifying credit card transactions. The Random Forest, and Decision Tree algorithms exhibit similar performance, with accuracy, precision, recall, and F1-scores of approximately 0.99. This suggests that these algorithms are well-suited for detecting fraudulent transactions, as they achieve a high level of correctness and effectively minimize false positives and false negatives. Logistic Regression, although slightly lower in performance compared to the other algorithms, still demonstrates a satisfactory level of accuracy, precision, recall, and F1-score, with values around 0.97.

TABLE III. THE EVOLUTION RESULTS FOR ALGORITHMS OF PROPOSED SYSTEM

Classification algorithm	Accuracy	Precision	Recall	F1-score
Random forest	0.99	0.99	1.0	0.99
Decision Tree	0.99	0.99	0.99	0.99
Logistic Regression	0.97	0.98	0.96	0.97

Received 12/August/2023; Accepted 15/November/2023
 ISSN (Print) 1811-9212 *ISSN (Online) 2617-3352*

DOI: https://doi.org/10.33103/uot.ijccce.24.1.6

VIII. CONCLUSIONS

The goal of our research, "Designing and Building a Machine Learning Model for Detecting Credit Card Fraud," was to provide ways for identifying fraud and lowering losses. Many supervised learning algorithms are utilized, including Random Forest, which has an accuracy of 99.99%, Decision Tree, which has an accuracy of 99.98%, and Logistic Regression, which has an accuracy of 97.34%. All of these methods are compared using the same initial dataset. Due of the extreme imbalance in our dataset, sampling and oversampling techniques have been used. Finally, the Random Forest algorithm was the best and most accurate in detecting online fraud, therefore it would be suitable for prediction.

REFERENCES

[1] G. Sandhya, M. Abishek, S. Gunal Kumar, and R. S. Jisenthira Kumar, "Credit Card Fraud Detection using Machine Learning Algorithms," *Lect. Notes Networks Syst.*, vol. 516, no. November, pp. 313–320, 2023, doi: 10.1007/978-981-19-5221-0_30.

[2] P. K. Sadineni, "Detection of fraudulent transactions in credit card using machine learning algorithms," in *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*, IEEE, 2020, pp. 659–660.

[3] A. O. Adewumi and A. A. Akinyelu, "A survey of machine-learning and nature-inspired based credit card fraud detection techniques," *Int. J. Syst. Assur. Eng. Manag.*, vol. 8, no. 2, pp. 937–953, 2017.

[4] P. K. Sadineni, "Detection of fraudulent transactions in credit card using machine learning Algorithms," *Proc. 4th Int. Conf. IoT Soc. Mobile, Anal. Cloud, ISMAC 2020*, no. October, pp. 659–660, 2020, doi: 10.1109/I-SMAC49090.2020.9243545.

[5] G. Sandhya, M. Abishek, S. Gunal Kumar, and R. S. Jisenthira Kumar, "Credit Card Fraud Detection using Machine Learning Algorithms," *Lect. Notes Networks Syst.*, vol. 516, no. 07, pp. 313–320, 2023, doi: 10.1007/978-981-19-5221-0_30.

[6] Y. Jain, N. Tiwari, S. Dubey, and S. Jain, "A comparative analysis of various credit card fraud detection techniques," *Int. J. Recent Technol. Eng.*, vol. 7, no. 5, pp. 402–407, 2019.

[7] U. Fiore, A. De Santis, F. Perla, P. Zanetti, and F. Palmieri, "Using generative adversarial networks for improving classification effectiveness in credit card fraud detection," *Inf. Sci. (Ny).*, vol. 479, pp. 448–455, 2019.

[8] S. M. K. Suha M. Najem, "An Efficient Feature Engineering Method for Fraud Detection in E-commerce," *Iraqi Journal of Computer, Communication, Control and System Engineering*. pp. 40–52, 2021. doi: 10.33103/uot.ijccce.21.3.4.

[9] A. Agarwal, M. Iqbal, and B. Mitra, "Survey of various techniques used for credit card fraud detection," *Int. J. Res. Appl. Sci. Eng. Technol*, vol. 8, no. 7, pp. 1642–1646, 2020.

[10] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," in *2017 international conference on computing networking and informatics (ICCNI)*, IEEE, 2017, pp. 1–9.

[11] A. Lawi and F. Aziz, "Classification of credit card default clients using LS-SVM ensemble," in *2018 Third International Conference on Informatics and Computing (ICIC)*, IEEE, 2018, pp. 1–4.

[12] C. Liu, Y. Chan, S. H. Alam Kazmi, and H. Fu, "Financial Fraud Detection Model: Based on Random Forest," *Int. J. Econ. Financ.*, vol. 7, no. 7, 2015, doi: 10.5539/ijef.v7n7p178.

[13] R. Vijay Bhasker Vangara, S. Prasad Vangara, and V. R. Kailashnath Thirupathur, "A Survey on Natural Language Processing in context with Machine Learning," *Int. J. Anal. Exp. modal Anal.*, 2020.

[14] R. I. Ahmed, R. M. Mohsin, and R. F. Ghani, "Cardiovascular Patients Monitoring Using Internet of Things And Decision Tree," *Iraqi J. Comput.*, vol. 22, no. 4, p. 3, 2022, doi: 10.33103/uot.ijccce.22.4.12.

[15] H. M. Fadhil, M. N. Abdullah, and M. I. Younis, "A Framework for Predicting Airfare Prices Using Machine Learning," *Iraqi J. Comput.*, vol. 22, no. 3, 2022, doi: 10.33103/uot.ijccce.22.3.8.

[16] H. John and S. Naaz, "Credit Card Fraud Detection using Local Outlier Factor and Isolation Forest," *Int. J. Comput. Sci. Eng.*, vol. 7, no. 4, pp. 1060–1064, 2019, doi: 10.26438/ijcse/v7i4.10601064.

[17] D. N. Mhawi and S. H. Hashim, "Proposed Hybrid Ensemble Learning Algorithms for an Efficient Intrusion Detection System," *Iraqi J. Comput.*, vol. 22, no. 2, 2022, doi: 10.33103/uot.ijccce.22.2.7.

Received 12/August/2023; Accepted 15/November/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352