IJNRD.ORG ISSN : 2456-4184

INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG

An International Open Access, Peer-reviewed, Refereed Journal

CREDIT_CARD FRAUD DETECTION USING ML ALGORITHMS

[1]Mr. Bharath M , [2]Sindhu Shree H R, [3]Pavan , [4]Sneha and [5]Rahul H
[1]Asst Prof, Dept of CSE, Sri Venkateshwara College of engineering, Bangalore,
[2,3,4,5]UG scholar, Dept of CSE, Sri Venkateshwara College of engineering, Bangalore

Abstract– This paper examines the increasing problem on creditcard(CD) fraud as evolving within the rapidly developing world of electronic commerce. With creditcards becoming more widespread as a payment method, the no of false transactions are also increasing at an alarming rate across the world and leading to significant monetary loss. In the face of such challenge, a no of sophisticated approaches rooted in artificial intelligence, It gives an overview of the techniques, stressing the necessity for effective fraud detection systems to protect damage from credit card issued by banks. Machinelearning(ML) with approaches that includes DT, LR, RF, Genetic Algorithms and ensemble techniques is then considered to demonstrate the efficacy of finding out genuine transactions from all fraudulent ones. Real-world datasets are utilized in the research to show how competitive this work against existing systems. This last section concludes on the ever evolving fraud detection ideas and a broad understanding of different counter–measures across domains with cases in creditcard fraud, telecommunication fraud and computer intrusion.

Keywords: random forest(RF), decision tree(DT), oversampling, under sampling, logistic regression(LR), credit card(creditcard), fraud detection, classification, training set, test set.

I. INTRODUCTION

Credit card technology was first developed by Bank of America in 1958. By 2020, there were 1.4 million allegations of identity theft, with 393,207 of those cases employing Credit Conversion Factor (CCF). CCF emerged as the second most prevalent category of purported identity theft. [2]. The Credit Conversion Factor (CCF) is a measure used in banking and finance to find the credit risk connected to various financial instrument types. It is often employed in the calculation of credit exposure for regulatory purposes.

The Annual Fraud Report on United Kingdom Finance for 2022 states that through both authorized and unauthorized criminal activity, over In 2022, 1.2 billion were taken., amounting to an astounding Every minute, 2300 are gone. Remarkably, 18% of Authorized Fraud involving Push Payment (APP) incidents came through phone lines, and 78% of cases started online. [6] .

Fraud using credit cards is among the leading problems in the world. In 2020 the credit card scam cases had increased to 45,120 [1]The last ten years have seen an exponential growth in the Internet. As a result, the growth and utilization of amenities such as electronic bill payment, tap and pay, and e-commerce, among others. Fraudsters have therefore intensified their efforts to target credit card transactions. [3][4]**.** Fraud detection involves observing users activities to estimate, understand, or avoid objectionable behaviour that accommodates fraud, intrusion, and defaulting. There are far more legitimate transactions than there are fraudulent ones. Also, the dealing patterns usually modify their applied mathematical properties throughout your time [1] .

identifying and stopping credit card theft is of utmost importance to maintain trust within the monetary system and protect the interests of consumers and businesses. Traditional fraud detection methods, such as manual inspection and rule-based systems, have proven to be insufficient in dealing with the evolving techniques employed by fraudsters. As a result ,There is an demanding need for more_advanced and automated approaches to address this issue effectively.

Two of the several techniques for securing credit card transactions are tokenization and credit card data encryption. Despite their general effectiveness, these techniques fall short of completely safeguarding transactions with credit cards from fraud. Through the application of dataset from past experiences, or ML, artificial intelligence(AI) may learn from computers and enhance their predictive capabilities without needing special programming. [3][4].

Machinelearning techniques are nesessaary for detecting fraudulent transactions in creditcard fraud detection. These algorithms analyse patterns and characteristics within large datasets to discern between actions that are fraud and those that are lawful. For the purpose of detecting creditcard fraud, ML methods can be categorized into supervised, unsupervised, and semi-supervised learning methods.

Labelled training datasets is used by supervised learning algorithms to understand patterns linked to both fraudulent and non-fraudulent transactions.

The usage of unsupervised learning methods when labelled fraud data is unavailable.

Semi_supervised learning algorithms combine elements of both supervised and un_supervised learning. They utilize a small amount of labelled fraud data along with a large amount of unlabelled data to improve the fraud_detection. In this study, we apply ML techniques such as LR , RF and decision trees. This work makes use of a credit_card fraud dataset that was created by credit cardholders throughout Europe. These datasets frequently contain a diffrent characteristics that could negatively affect the classifiers' performance during training. [3]. So Consequently, to balance the dataset we perform under sampling and oversampling methods . Several criteria are used to assess these machine learning algorithms, including accuracy, recall, and F1 score [5].

The remainder of this essay is structured as follows: In Section 2, It covers a literature on creditcard_fraud, It is supplied with an outline of the models as well. The process and methodology for identifying credit_card fraud are covered in Section 3. The experiments' results are given in Section4, along with a discussion regarding the analysis. The comparative_analysis is ended in Section 5, which further suggests further research.

II. . LITRETURE SURVEY

1.This paper explores detection of creditcard-fraud utilizing machine learning, specifically Logistic_Regression, RF Naïve Bayes Classifier, DecisionTree, and Support Vector Machine(SVM). The dataset from Kaggle comprises 284,807 transactions by European credit card holders, with only 492 fraud transactions (0.172% of total). Bcz of the imbalanced dataset, experiments with different test sizes were conducted.

The Random_Forest model frequently outperforms Naïve Bayes, Logistic Regression, and SVM in the results. Regardless of the training/testing split (80%/20% or 20%/80%), Random Forest consistently provides superior outcomes, making it the preferred choice for detecting creditcard theft[1].

2. This research focuses on credit_card fraud detection, calculating the effectivenes of KNN comparing with other machine-learning methods. Utilizing public data and high-class disparity statistics, the proposed approach achieved impressive results: 99.95% accuracy, 97.2% precision, 85.71% recall, and 90.3% F1-score.

The study used a real-world credit_card transactions dataset[3] from GitHub, containing 331,624 transactions, including 571 fraudulent ones. The dataset was split into 80% training and 20% testing for three different credit card datasets: European, Brazilian, and Chinese [2].

3. This paper proposes an ML-based creditcard fraud_detection engine using Decision_Tree (DT), Logistic Regression (LR), and Artificial Neural Network (ANN) classifiers. Evaluated on a dataset from European cardholders sourced from Kaggle, the proposed engine can predict fraud transactions promptly when integrated into bank systems. The study uses 80% of the data for training and 20% for testing, employing accuracy(acc), precision(pre), and recall_scores for evaluation. Comparing the three algorithms, Neural Networks outperform with an accuracy score of 0.99, precision(pre) of 0.74, and recall of 0.70, demonstrating its effectiveness in predicting fraudulent transactions [3].

4.The study presents a genetic algorithm (GA) based ML based credit card fraud detection engine that selects features. utilizes Decision_Tree(DT),Random_Forest(RF),Logistic_Regression(LR), Artificial Neural Network(ANN), and Naive Bayes(NB) machine learning classifiers. To verify the performance, For this study, they employ a dataset containing credit card transactions done by European cardholders over the course of two days in September 2013. In total, there are 284807 transactions in this dataset, of which 0.172% are fraudulent. The 30 features (V1,..., V28) in the datasetss are Time and Amount. On Google Colab, the experimental procedures were carried out. The Scikit-Learn machine learning framework was employed in this study. RF algorithms achieved the greatest test accuracy (TAC) of 99.94%. On other hand, the most precise findings were obtained with the RF approach.[4].

5.This study uses a dataset of 550,000 records from European-cardholders in 2023 to compare cutting-edge machine-learning algorithms for creditcard fraud_detection. For both non-fraudulent and fraudulent_transactions, perfect accuracy, recall and an F1_score of 1.00 are attained by using LR, RF , extra trees, and LGBM. TN, FN, TP, and FP are examples of assessment metrics that show accurate classification free of errors. While models like Random_Forest and XGBoost, with longer training , offer a trade_off between training time and result accuracy, LightGBM and LR, with training times in seconds, are effective for real-time fraud_detection [5].

6.In order to overcome shortcomings in credit-card fraud_detection, this study suggests a ideal model that combines Support Vector Machine(SVM), K-Nearest Neighbor(KNN), Random_Forest (RF), Bagging, and Boosting classifiers. To address dataset imbalance, the model uses the Synthetic Over-sampling Technique (SMOTE) in conjunction with under-sampling. Of the 2,84,807 transaction in the data_set, which represents two days' worth of transactions by European credit cardholders, 492 incidents include fraud, accounting for just 0.172% of the total transactions. Performance was evaluated using the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) values of the ensemble model, which are displayed in the confusion matrix. The model has improved 100%, based on the results, with the bagging classifier and P_M_2 obtaining the highest accuracy percentages in the testing sample, 94.73%. Together, the RF and SVM classifiers With 94.21% accuracy, the RF and SVM classifiers come in close second, while the Boosting, KNN, P_M_1, and LR classifiers have ACCs of 93.68%, 93.68%, 93.68%, and 93.0%, in that order [6].

7. LR, decision trees(DT), and random forests algorithms were used in identifying credit card fraud. Transactions that are frauds are designated as the "positive class," and genuine transactions are classified as the "negative class." Error rate, accuracy, sensitivity, and specificity are all considered in performance evaluation. Additionally, the report offers information on popular fraud tactics, techniques for identification, and recent advancements in the sector. With a tenth of the dataset, the MLmethod [7] This study uses a dataset of 284,808 transactions from Kaggle, which represents a European bank, to examine the efficacy of shows precision of 28%, which improves to 33% when the complete dataset is used. Nevertheless, the technique reaches above 99.6% accuracy. The notable high degree of accuracy is ascribed to the significant disparity between legitimate and fraudulent transactions.[7]

2.1 RESEARCH GAP

Although there has been a lot of development in creditcard_fraud detection, there are still issues. and areas that have not been completely solved or accomplished. Some of these include **Imbalanced Datasets:** Creditcard fraud is a relatively rare event compared to legitimate_transactions, leading to imbalanced-datasets where fraudulent-transactions are significantly outnumbered by non-fraudulent ones. Imbalanced datasets can hinder machine learning models' performance, as they frequently exhibit bias in favour of the majority class.

False Negative and Positive Results: Achieving an optimal balance between detecting as many fraud cases as possible (low false negatives) while reducing the number(no) of false positives (flagging legitimate transactions as fraud) is a challenging task. High false positives can result in customer inconvenience and unnecessary investigations, while high false negatives can result in missed fraudulent activities.

Cross-Channel Fraud Detection: Fraudulent activities often involve multiple channels, such as online transactions, mobile payments, and point-of-sale transactions. Detecting and linking fraudulent activities across different channels to form a comprehensive view of fraud patterns is a complex task. Developing effective cross-channel fraud detection techniques is an area that requires further exploration.

Privacy Concerns: Fraud detection often requires access to sensitive customer information and transaction data. Balancing the necessity of efficient fraud detection with privacy concerns is a complex task. Striking the right balance between privacy preservation and fraud detection accuracy remains an ongoing challenge.

III. . METHODOLOGY

3.1 MACHINE_LEARNING ALGORITHMS APPLIED TO THE DETECTION OF CREDIT CARD -FRAUD

3.1.1 LOGISTIC REGRESSION

A statistical approach called LR is frequently applied to binary classification problems such as creditcard fraud_detection. It makes predictions about whether or not a creditcard transaction is fraudulent. This program improves security by succinctly and successfully recognizing possible fraud incidents. The process begins with the preparation of a dataset containing historical transactions, featuring labelled outcomes of fraud or non-fraud. Relevant features, such as transaction amount ,result of a PCA transformation, and time, are selected for model training. The logistic function, or sigmoid function, is employed to transform predicted values into probabilities between 0 and 1. The instance is trained by adjusting weights to maximize the likelihood of observed outcomes, typically accomplished through optimization techniques like gradient descent. Upon completion, the model defines a decisive boundary, allowing the categorization of transactions into fraudulent or non-fraudulent classes. To gauge its effectiveness, evaluation metrics like accuracy, precision, recall, and F1-score are utilized on an independent testing set. Models demonstrating proficiency can subsequently be operationalized for real-time predictions, effectively identifying potential instances of credit_card fraud in newly processed transactions.

- **model training :**

Logistic Regression estimates the probability of an event occurring, in this case, the probability of a creditcard_transaction being fraudulent.

The logistic function (sigmoid function) is used to map the predicted values into probabilities between 0 and 1.

$P(Y=1)=1+e-(\beta0+\beta1X1+\beta2X2+...+\beta nXn)1$

Y is the binary outcome (1 for fraudulent, 0 for non-fraudulent),$\beta0$ is the intercept $\beta1,\beta2,...,\beta n$ are the coefficients connected to the features $X1,X2,...,Xn$.

3.1.2 DECISION TREE:

Recursively, the Decision Tree method divides the dataset assembling the features into a structure like a tree, with each node representing a choice depending on a particular feature. This division procedure keeps going till a stopping condition is met, ensuring the tree does not grow excessively complex. The resulting decision rules, found along the paths from the root_node to the leafnodes, take the form of "If-Then" statements and provide transparency into the decision-making process. Decision Trees handle imbalanced datasets naturally, but adjustments can be made to address class imbalance. The model's performance is calculated on a testing set using standard metrics, and if satisfactory. Decision Trees may struggle with capturing difficult relationships, leading to the use of another methods like random forest.

A assortment of data with several rows and columns where last column is referred as a class and the rows or tuples are referred as instance and other columns are called as attributes or features.

3.1.3 RANDOM FOREST

Random Forest, an ensemble learning algorithm, is widely used in creditcard fraud_detection because of its ability to control complex data linkages and reduce overfitting. The algorithm operates by constructing a forest of DT, each trained on a portion of the information and employing variety of features. Through the process of bootstrap sampling, random subsets of the original-dataset are dicipiated using substitute for training each tree, ensuring diversity in the data. Additionally, feature randomization at each node of the Decision Trees introduces variability by considering only a random_subset of features for splitting. During prediction, each tree independently provides its outcome, and the ultimate classification is determined through a majority or weighted vote from all trees in the ensemble. This voting mechanism enhances the model's robustness and accuracy, particularly in handling imbalanced datasets commonly encountered in creditcard fraud detection

3.2 METHODS USED

In this section all the above mentioned algorithm are processed by under_sampling and over_sampling method.

using under_sampling and over_sampling techniques is often crucial to address the class imbalance, where instances of non-fraudulent transactions significantly outnumber those of fraudulent transactions. Under sampling involves reducing the no of instances from the majority class(non-fraudulent) to produce a dataset that is more balanced., while oversampling involves increasing the instances of the minority_class (fraudulent) to accomplish a balanced distribution. the minority class's (fraudulent) instances in order_to reach a balanced distribution.

3.2.1 UNDERSAMPALING:

To overcome class imbalance in a dataset—a situation in which one class contains noticeably less instances than the other—under sampling is a machine learning approach. Under sampling is the tenchinque of randomly deleting instances from the majority class(non-fraudulent transactions) in the context of credit_card fraud-detection, when fraudulent_transactions are often rare relative to valid ones, to establish a more balanced class distribution. This lessens the likelihood that the model will be biased towards of the majority_class and enhances its capacity to recognize and pick up on trends connected to the minority_class (fraudulent transactions).

3.2.2 OVERSAMPALING:

Oversampling is a ML technique used to address class imbalance in datasets where one class is significantly underrepresented in comparison to the other. creditcard fraud_detection, when instances of fraudulent-transactions are frequently infrequent, oversampling is the practice of increasing the no. of instances in the minority class(fraudulent transactions). To create a more equitable distribution of classes is the aim. One popular method to do this is to replicate existing instances of the minority_class.

3.3 PROCEDURE :

Initially, the dataset was gathered from the Kaggle to identify instances of creditcard fraud. Next, to analyse the dataset and find the fraud, we import the necessary library functions. next we read the supplied dataset after importing the necessary library functions. Next, we determine the dataset's shape i.e. the no of rows and columns, and then we gather the dataset's information, including the total number-of-rows, total_number of columns, datatypes for each column, and memory requirement.

Then we preprocess the dataset import a Standard Scaler to normalize the features and ensure consistency in the range of values across the dataset. This normalization is essential for machine_ learning models, as it helps prevent features with larger scales from dominating the learning process.

we conduct a thorough scan of the_dataset to find and subsequently remove any duplicate values. Upon completion of this cleansing process, we proceed to store the Feature Matrix in the variable X, and the Response (Target) in the vector y. This strategic arrangement paves the way for the subsequent division of the- dataset into the Training-Set and Test Set, a pivotal step in the ML workflow. This division allows for the efficient training of ML models on one subset, while simultaneously providing a distinct subset for the rigorous evaluation of model performance. To achieve this, we designate 20% of the_dataset as the test set, reserving the remaining 80% for the training set, ensuring a robust and reliable assessment of model capabilities.

then, we use under_sampling and over_sampling to handel with the unbalanced data collection. Under sampling involves fewer examples from the majority_class(non-fraudulent) to produce a more balanced dataset. Next, we apply different ML_techniques and identify which one produces the highest level of accuracy, f1_score and recall score. But the disadvantages of under sampling is that we lose lot of valuable data so we perform oversampling .

In oversampling we are going to utilize SMOTE(synthetic minority oversampling technique) to do oversampling. SMOTE helps address the imbalanced class distribution by introducing diversity to the minority_class, making the model less prone to overfitting on the limited available examples. It is particularly useful when the underrepresented group is small and additional instances are needed to increase the robustness of the model for machine learning. and Next, we utilize various ML methods and determine which one yields the highest level of accuracy, f1_score and recall score. Then we determine which is the finest and we save the model .

IV. . RESULTS AND DISCUSSIONS

4.1 DATASET SUMMARY

The dataset was obtained via Kaggle. The dataset includes creditcard transactions performed by cardholders throughout Europe in September-2013.

This dataset shows the transactions that took place over a two-day period. Of the 284,807 transactions, 492 were fraudulent. because of the extreme imbalance 0.172% of all transactions in the dataset fall into the positive_class. called as fraud transactions.

Its sole input variables are numbers that come from a PCA transformation. Regretfully, Kaggle is unable to offer the original features and further context for the data owing to privacy concerns. Features V1, V2,... The primary components identified by PCA are V28; "Time" and "Amount" are the only features that haven't undergone PCA transformation. The seconds that pass between every transactions and the dataset's initial transaction are contained in the feature "Time." The transaction's total amount is shown as the feature "Amount," which can be utilized for example-based learning that is cost-sensitive. As the response_variable, the 'Class' feature takes value 1 in the event of frauds and 0 otherwise.

4.2 EXPERMINATAL CONFIGUARATION

The trial procedures were performed using Google Collab. importing the certain library functions such as pandas , NumPy , warnings , Standard Scaling , seaborn .

4.3 EVALATION CRITERIA

To assess the outcomes of the classification algorithms there are various parameter such as Accuracy score, classification report, F1-score etc. Some important definitions are:

- **True positive (TP)**- It is an outcome in which the model accurately predicts the positive class.
- **False positive (FP)**- It occurs when the positive class is predicted wrongly by the model.
- **True negative (TN)**- It is an outcome in which the model accurately predicts the negative class.
- **False negative (FN)**- It is an outcome whereby the model forecasts the negative-class inaccurately.
- **Accuracy**: Accuracy is the proportion of number of correct predictions that means (TN+TP)/(TP+TN+FN+FP).
- **F1 Score**: The Formula One Score is the harmonic mean of precision and recall values for a classification problem.
- **Sensitivity**: It is also known as recall score . It's the ratio of actual positive cases correctly identified. That means TP/(TP+FN)

4.4 EXPERIMENTS WITH ML MODELS

To find credit card fraud, three ML-techniques were used. Data from 80% of the training dataset and 20% of the testing dataset were utilised to asses the algorithms. Accuracy, precision score , f1_score, and recall score are used to analyse the performance .

4.4.1 FOR UNDER SAMPLING :

Accuracy obtained for logistic_regression is **0.9263157894736842**

Precision score obtained for LR is **0.94**

Recall score obtained for LR is **0.921568627450980**

F1_score for LR is **0.9306930693069307**

Accuracy obtained for decision_tree is **0.8789473684210526**

Precision score obtained for decision_tree is **0.8761904761904762**

Recall score obtained for DT is **0.901960784313725**

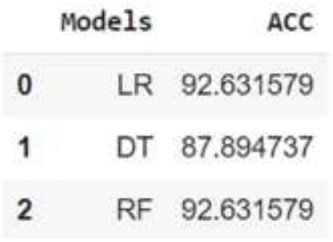
F1_score obtained for decision tree(DT) is **0.888888888888889**

Accuracy obtained for random_forest is **0.9263157894736842**

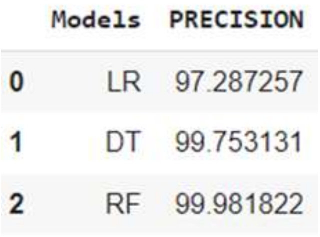
Precision score obtained for RF is **0.958333333333333**

Recall score obtained for RF is **0.9019607843137255**

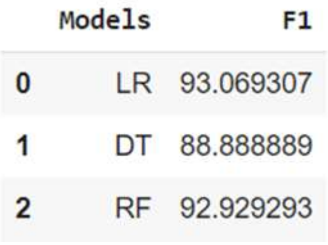
F1_score obtained for random_forest is **0.9292929292929293**



	Models	ACC
0	LR	92.631579
1	DT	87.894737
2	RF	92.631579

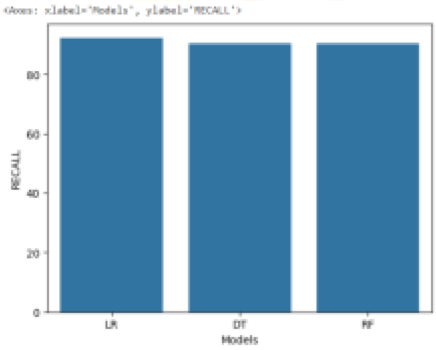


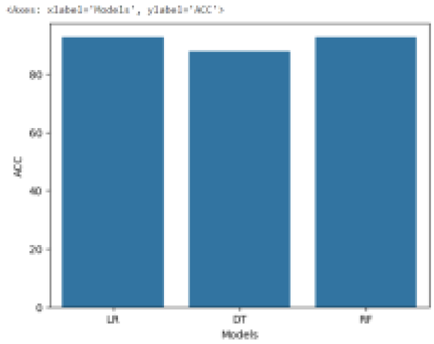
	Models	PRECISION
0	LR	97.287257
1	DT	99.753131
2	RF	99.981822

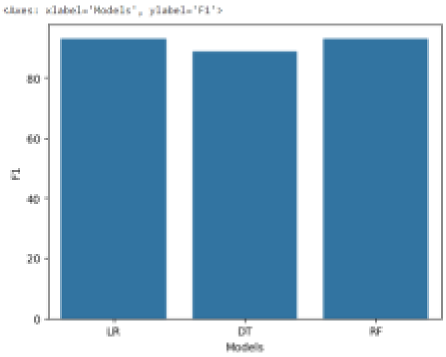


	Models	F1
0	LR	93.069307
1	DT	88.888889
2	RF	92.929293

4.4.1.1 BAR GRAPH







4.4.2 FOR OVERSAMPLING

Accuracy resulted for logistic_regression is **0.9443566263308987**

Precision score obtained for LR is **0.972872566851**

Recall score for LR is **0.914131956438739**

F1_score for logistic_regression is **0.9425879926887566**

Accuracy resulted for decision_tree is **0.9983193429993823**

Precision score obtained for DT is **0.9976958525345622**

Recall score resulted for DT is **0.9993006503951325**

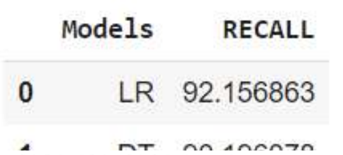
F1_score resulted for DT is **0.998497606652458**

Accuracy resulted for random_forest is **0.999909153675642**

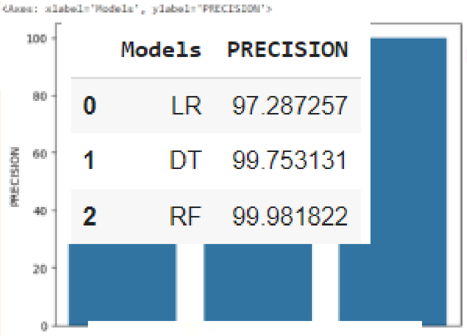
Prec **18224783233**

Reca

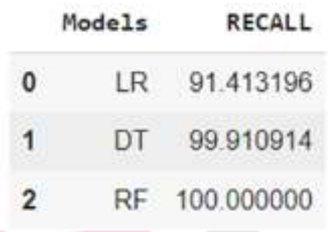
F1_s **1303083**



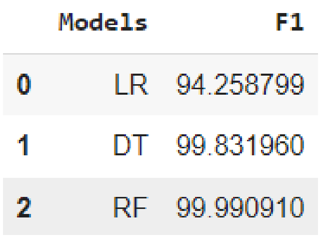
	Models	RECALL
0	LR	92.156863
1	DT	[illegible]



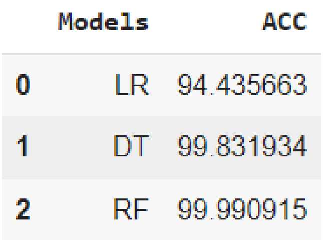
	Models	PRECISION
0	LR	97.287257
1	DT	99.753131
2	RF	99.981822



	Models	RECALL
0	LR	91.413196
1	DT	99.910914
2	RF	100.000000

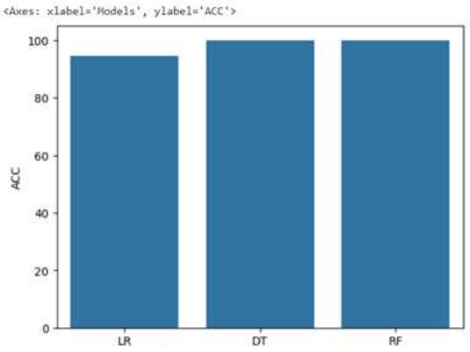


	Models	F1
0	LR	94.258799
1	DT	99.831960
2	RF	99.990910

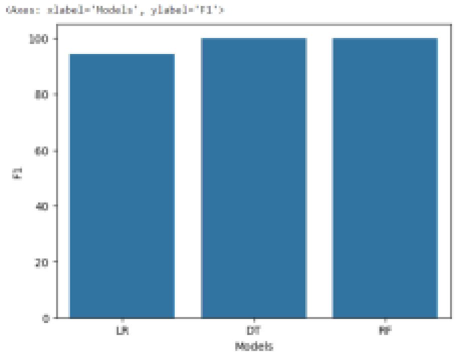


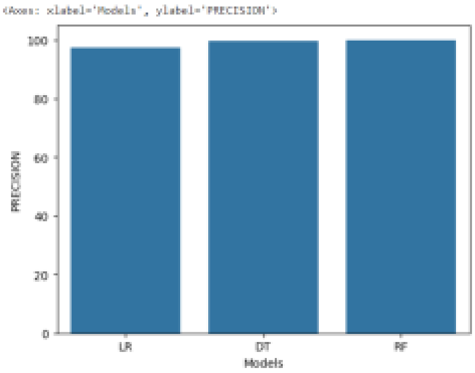
	Models	ACC
0	LR	94.435663
1	DT	99.831934
2	RF	99.990915

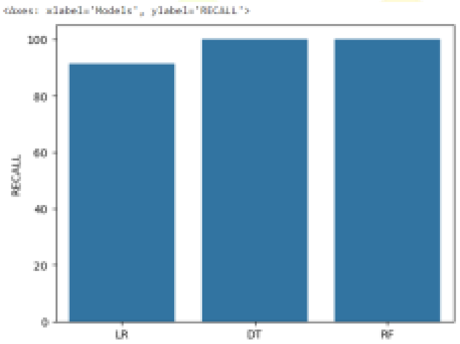
4.4.2.1 BAR GRAPH



5.1







When all three machine-learning models are analyzed, random forest produces results that are more accurate than decision_trees and logistic_regression.

V. . CONCLUSION

The significance of applying techniques for machine learning to improve security measures in financial transactions is highlighted by the study on identifying creditcard theft. The review of several algorithms, such as Random Forest, decision trees, and LR, emphasizes how crucial it is to choose the right model for fraud_ detection. The evaluation's measurements of accuracy, recall, and precision demonstrate how well these models distinguish between authentic and fraudulent transactions.. The results highlight the necessity of an all-encompassing strategy that takes into account variables like dataset imbalance, algorithm efficiency, When selecting a machine_learning model for creditcard fraud_detection . We determined the precision, recall (sensitivity), F1-score, and other performance indicators for each method, which allowed us choose the best one out of the few that were employed. Among the algorithms for machine learning, random forest produced the best results. Having a 100% recall score, 99% f1 score, and 99% accuracy.

5.1 REFERENCES

[1]. Indrani Vejalla , Sai Preethi Battula , Kartheek Kalluri , Hemantha Kumar Kalluri , "Credit_Card_Fraud_Detection_Using Machine_Learning Techniques" , 2023 2nd International_ Conference on Paradigm Shifts in Communications Embedded Systems, Machine Learning(ML) and Signal Processing (PCEMS), 2023 IEEE

[2[. Anjankumar, S. Poonkuntran, and Ananya Singhai , " A Novel Methodology for Credit-Card Fraud_Detection using KNN Dependent Machine Learning Methodology" , Proceedings of the Second International Conference on Applied Artificial Intelligence and Computing (ICAAIC 2023) , 2023 IEEE Xplore

[3]. Prateeksha M.S1, B. Naga Swetha1 and Manjula Patil ," CREDIT_CARD FRAUD_DETECTION USING_MACHINE-LEARNING" , April 2023

[4]. Emmanuel Ileberi1, Yanxia Sun and Zenghui Wang , "A machine learning based credit_card fraud_detection using the GA-algorithm for featureselection", April 2023 International Journal_of Advanced-Research

[5]. Ayesha Aslam and Adil Hussain "A Performance Analysis of Machine_Learning Techniques for Credit_Card_Fraud Detection", January 2024

[6]. Abdul Rehman Khalid 1, Nsikak Owoh 1,* , Omair Uthmani 1, Moses Ashawa 1 , Jude Osamor 1 and John Adejoh 2, "Enhancing Credit_Card_Fraud Detection: An Ensemble Machine Learning Approach", 3 January 2024

[7]. MS Joitson, Preejamol Prasad, Rimil Joseph, Jayakrishnan B," Credit _Card Fraud Detection_ using Machine_Learning", International Journal_ of Engineering _Research & Technology (IJERT) ,ICCIDT - 2023 Conference Proceedings , ISSN: 2278-0181

[8] . Asha RB , Suresh Kumar KR," Credit card fraud detection using Artificial Neural Network", Global transitions proceedings 2021.

[9]. Seera .M, Lim CP, Kumar A, Dhamotharan L, Tan KH. An intelligent payment card fraud detection system. Ann Oper Res 2021;1–23.

[10]. Y. Fang,Y. Zhang and C.Huang, "Credit card fraud-detection based onmachin learning,"*Comput.Mater.Contin.*, vol. 61, no. 1, 2019.

[11]. S. Dhingra, "Comparative analysis of algorithms for credit card fraud-detection using data mining: A

review," *J. Adv. Database Manag. Syst.*, vol. 6, no. 2, pp. 12–17, 2019

[12]. A. Mishra, C. Ghorpade, "Credit Card Fraud- Detection on the Skewed Data Using Various Classification and Ensemble Techniques" 2018 IEEE International Students' Conference on Electronics, Electrical and Computer_Science(SCEECS) pp. 1-5. IEEE

[13]. P. Kumar, F. Iqbal, Creditcard fraud_identification using machine_learning approaches, in: Proceedings of the 1st International_Conference on Innovations in Information and Communication Technology (ICIICT), CHENNAI,

India, 2019, pp. 1–4, doi:10.1109/ICIICT1.2019.874149

[14]. Yashvi Jain, NamrataTiwari, Shripriya Dubey, Sarika Jain," A Comparative _Analysis of Various Credit_Card Fraud_Detection Techniques", International _Journal of Recent_ Technology and

Engineering(IJRTE)_ISSN: 2277-3878, Volume-7 Issue-5S2, January 2019