Efficiency-Driven Machine Learning Models for Credit Card Fraud Detection: A Comparative Study and Optimization Approach

Sumit Haldar, Maisha Chowdhury, Nazifa Bushra, Farah Binta Haque,
Ehsanur Rahman Rhythm and Annajiat Alim Rasel
Department of Computer Science and Engineering (CSE)
School of Data and Sciences (SDS)
Brac University
66 Mohakhali, Dhaka - 1212, Bangladesh
{sumit.haldar, nazifa.bushra, maisha.shabnam.chowdhury, farah.binta.hauqe,
ehsanur.rahman.rhythm}@g.bracu.ac.bd, annajiat@gmail.com

***Abstract*—Since credit card fraud is a serious danger to both consumers and financial institutions, reliable and effective fraud detection systems must be developed. This paper offers detailed research on efficiency-driven machine learning models for credit card fraud detection, with the goal of exploring and optimizing their performance. To assess the efficacy of several deep learning architectures in identifying fraudulent transactions, we compare different classification models and their hybrid equivalents. Our study uses a broad dataset that includes both legitimate and fraudulent credit card transactions, which makes it easier to evaluate the model's performance realistically. Our goal is to maximize the effectiveness of these models by resolving issues with uneven class distribution, interpretability, and computing capacity. We describe the findings of our comparative analysis and optimization strategy, which highlight the advantages and disadvantages of several machine learning models in situations involving credit card fraud detection. The significance of effective and precise fraud detection systems in the quickly changing field of financial security is emphasized by discussing the consequences of our findings for practical implementations.**

***Index Terms*—credit card, fraud, transaction, optimization, security, classification**

I. Introduction

The evolution of digital transactions brought convenience to human beings worldwide and has revolutionized the financial landscape into a seamlessly connected global network.The central point of this digital network are the credit cards which is enabling the transactions at a surprisingly fast and efficient pace. But this widespread popularity of digital transactions has come with a cost , that is the escalation of credit card fraudery. Traditional methods for recognising fraud encounter unparalleled hurdles in keeping up with the growing risks posed by these increasingly sophisticated fraudulent operations.

Credit card theft can manifest itself in a variety of ways, and con artists are always refining their strategies to exploit gaps in their defenses. Here are a few common examples of credit card fraud:

1) Unauthorized Transactions: Using credit card information obtained through hacking or data breaches, fraudsters use this data to conduct transactions with the card holder's knowledge.

2) Account Takeover: Usually by phishing or security breaches, fraudsters are able to gain access to the genuine cardholder's account. Once in control, they might fabricate data, change account information, or engage in criminal activity.

3) Lost or Stolen Cards: Unless a cardholder reports the loss, unauthorized individuals may make fraudulent purchases using an expired or stolen credit card. Scams of this type are common and easy to pull off.

4) Online Purchase Fraud: Credit card details that have been stolen are used by criminals to make purchases online. Because there aren't any physical cards available, it is challenging to recognize this type, which calls for robust online security measures.

5) Fake Credit Cards: Criminals create counterfeit credit cards by employing sophisticated forging methods or data that has been stolen. These fake cards are then used to make unauthorized purchases.

Credit card fraud is such a worldwide problem which affects all the countries of the world including the Republic of Korea. According to the FTC, we have come to a survey that about 1579 data breaches are there involving data points of 179 (million), which also involves fraudulent use of credit cards. Unfortunately the fraudulent use of credit card is the most common type of attack. Moreover, From a newest report of 2023 on credit card fraud transaction, the ratio on US credit card holders experiencing fraud at anytime in their life rose to 65 percent from 58 survey from 2022.

The percentage of US credit and credit card holders who have experienced fraud at some time in their life rose to 65% in 2022 from the 58% recorded in 2021, according to the 2023 Credit Card Fraud Report.

Fraud prevention is crucial in the world of digital transactions, and two primary techniques are used: one is anomaly detection and the other is misuse detection. Misuse detection is used to differentiate which is normal and possibly fraudulent credit card transactions by ML models based on identified patterns, much like a cunning investigator. It's like when a watchful security dog knows the difference between known and unknown activity. However, anomaly detection functions as a watchful defender, identifying the typical characteristics of transactions to build a unique profile and sounding an alarm when deviations take place. It is similar to a security system that is tuned in to normal patterns and prepared to warn users to any unusual events.

Our study attempts to open up new avenues for credit card fraud detection in response to the crucial intersection of financial vulnerability and technology. We provided a detailed analysis of sophisticated machine learning architectures that uses anomaly detection classification. Such as XGBClassifier, Logistic Regression, LSBM and KNN models along with (IF) also known as Isolation Forest. we have also used Local Outlier Factor. Rethinking the fundamentals of fraud detection in credit card transaction in terms of effectiveness, flexibility, and practicality goes beyond just fortifying security mechanisms.In the face of growing fraud complication in financial environments, our work provides direction where innovation and necessity converge. Come along on this voyage into the future of credit card fraud detection, where cybersecurity imperatives and innovation meet as we work towards an even more adaptable and safeguarded financial landscape.

II. Literature Review

In the context of e commerce and electronic payment systems, authors Sun et al. [1] introduced a engine to detect the frauds in credit card transaction utilizing Machine Learning classifier models for example DT or decision tree, NB or naive bayes, Random Forest (RF), Logistic Regression (LR) and also a NN model Artificial Neural Network (ANN).It underscores the necessity for robust ML-based solutions for the challenges faced in the evolving nature of fraudulent transactions and the skewed distribution in fraud datasets. To mitigate the challenge, the paper introduced a notable strategy which optimizes the selection of features incorporating the RF method in its fitness function which is a part of feature selection algorithm based on Genetic Algorithm (GA). Using a dataset created from European credit cardholders, an accuracy of 99.98% for GA-RF is attained which demonstrates the superiority of the GA-based feature selection technique. Still due to the highly confidential nature of credit card transaction datasets, experiments can't be fully replicated because ML models use anonymized attributes. The authors of the article, Xia et al. [2], talked about how the widespread use of internet technology has made credit card fraud a growing problem. It draws attention to the shortcomings of conventional machine learning techniques in identifying unfamiliar attack patterns and presents UAAD-FDNet, a novel framework that makes use of autoencoders, feature attention, and generative adversarial networks (GANs) in an Attentional Anomaly Detection Network which is also Unsupervised. Based on tests using the IEEE-CIS Fraud Detection Dataset and the Kaggle Credit Card Fraud Detection Dataset, the suggested strategy asserts better performance in fraud detection. highlights how technical developments have contributed to the surge in sophisticated credit card fraud, which includes techniques like payment fraud and credit card cashing. Fraud prevention and detection are critical study subjects because of the substantial economic consequences of the problems presented by these developing approaches. The significance of machine learning in fraud detection is highlighted, as it surpasses the constraints of conventional techniques. The authors [2] presented a brand-new unsupervised attentional anomaly detection paradigm and conducted a thorough analysis of the state of fraud detection in cradit card transactions. The suggested UAAD-FDNet uses a generator and discriminator to function according to GAN principles. To improve model training, the generator uses a attention mechanism on features which is channel wise and self-supervised learning; for supervision, a mixed weighted loss function is recommended. The suggested method's superiority is demonstrated by experimental findings on pertinent datasets, which frame detecting fraudulent in credit card transaction as the anomaly detection and provide unique features in UAAD-FDN et. Chung et al. [3] discussed in the context of the growing commerce landscape and IoT devices, how the issue of fraudulent credit card transactions has emerged. By combining three ML models: LR,KNN and LDA full abbreviation will be linear regression, K nearest and linear discriminant analysis respectively. A significant progress has been represented in credit card fraud detection and mostly to improve the sensitivity and accuracy, the incorporation of operators and conditional statements made the method unique from others.The substantial financial loss in the global surge due to internet fraud, notably in the US, due to the alarming rise of credit card fraud incidents (65% of credit card holders falling victim to fraud in 2022, up from 58% in the previous year) the urgency of the matter is highlighted.The study's credit card fraud real-world datasets are subjected to the integrated technique, showcasing recall scores of 1.0000, 0.9701, 1.0000, and 0.9362 for each of four dataset.An automated machine learning framework, PyCaret was used to validate the methodology and show that it outperforms other individual models. But alongside this, the omission of techniques like regularization and sampling methods to address skewed dataset created uncertainties in benchmarking against state-of-the-art fraud detection and thus for developing strategies and enhanced precision ,future research is suggested.

The paper of Alfaiz et al. [4] talks about how the COVID-19 epidemic has made people more reliant on internet services, which has increased credit card theft. Using a dataset of European cardholders, the study investigates 66

machine learning models for credit card fraud detection. Nine algorithms are tested in the first assessment stage, and the top three algorithms are tested in the second stage using 19 resampling strategies. The AllKNN-CatBoost model performs better than previous models with a 97.94% AUC, 95.91% Recall, and 87.40% F1-Score. The research addresses the difficulties associated with imbalanced datasets in fraud detection and suggests a sophisticated method that combines the most efficient machine learning algorithms with practical resampling strategies. The suggested method, which assesses resampling methods and machine learning algorithms for credit card fraud detection, is broken down into two phases. A total of 66 models are obtained by evaluating nine methods and 19 resampling strategies. The best model is found to be AllKNN-CatBoost, which outperforms the others in terms of F1-Score, AUC, and recall. In order to highlight the significance of correcting credit card fraud detection with imbalanced class, the research highlights how uncommon balanced datasets are in real-world circumstances. One month is allotted to the comprehensive evaluation procedure, which illustrates how comprehensive the methodology is. In order to prevent more sophisticated fraud, the paper's conclusion underlines the necessity for proactive solutions that make use of AI and machine learning. All things considered, the work presents a novel strategy to address class disparity in credit card fraud detection, exhibiting encouraging outcomes and providing directions for further investigation.In their paper, the authors Pandi et al. [5] discuss the growing issue of credit card fraud, which is a result of both increased card use and technology improvements. It promotes the use of LSTM-RNN. As an innovative approach to an efficient fraud detection system, it also uses attention mechanism along with it. When compared to alternative classifiers, the suggested model exhibits strong performance and high accuracy. The need of protecting digital transactions is emphasized, but the difficulties in spotting fraudulent activity are acknowledged. Although automated systems are capable of identifying unusual behavior, the analysis that follows needs human interaction, which presents a financial hurdle. The study emphasizes how well machine learning—particularly LSTM-RNN—performs in pattern detection.The topic of fraud detection in real-time is covered, using human rules and machine learning algorithms to quickly identify transactions that seem suspect. The main objective is to identify suspect credit card transactions by using machine learning algorithms (Naive Bayes, SVM, ANN, and LSTM) for pattern identification. The creation of a novel fraud detection system based on LSTM-RNN, effective behavior prediction of illicit service charge behavior, and competitive performance in comparison to current techniques are among the contributions. The study concludes by highlighting the need of adjusting to big data difficulties and presenting an LSTM-based model that is supported by a variety of datasets and performance measures for the identification of credit card fraud.In the study, Almarshad et al. [6] explores the growing issue of credit card fraudulants in this era of digital transactions and emphasizes the demand for sophisticated fraud detection techniques. For the traditional methods, the complexity of fraudulent activities is a matter of struggle to cope up, which is why innovation in the industry is essential. So for this reason, the authors came up with a unique method for detecting fraud using Generative Adversarial Networks (GANs) which eventually solve the imbalance in datasets connected to credit cards and also overcome the difficulties presented by developing fraud approaches. The study emphasizes the model's efficacy against a number of measures, highlighting its resilience, decreased false positives, and increased efficiency while highlighting the significance of developing a substantial dataset for fraud identification. The development of a new GANs-based identification method, dataset generation, resilience, a drop in false positives, a reduction in processing needs, and an increase in accuracy are among the contributions. The study does, however, admit its shortcomings, including the lack of available datasets, privacy issues, and the need for more investigation. Along with this, the study makes clear that more investigation is required to handle new fraud strategies and enhance the suggested methodology because privacy issues arise whenever there is transaction of data using a credit card.

Firdous et al. [7], in their paper have provided a thorough comparison between the different classifier models. Such as LR, RF and support vector machine. These models are compared on the account of recall precision and accuracy. For their research they have used UCI machine learning repository . Their data contained users payment history, bill amount, amount paid by the user and also the payment static from april to september. To Visualize the data they have used MATPLOTLIB and seaborn libraries. They have used the sklearn library to scale and preprocess the data. Moreover this paper also shows different works of researchers and the overview of their existing methods and approaches. The research work briefs about all the algorithms and their respective performance on the dataset. Provided a confusion matrix upon their results. To achieve the goal to evaluate the most accurate and effective model for detecting fraudulent on Credit Card transaction , this paper also offers their own experimental investigation of classification algorithms. From their analysis they have concluded that both super vector machine and random forest have a good accuracy compared to logistic regression. Even by tuning the threshold slightly, best recall, precision and accuracy can be achieved. The writers Kumar et al. [8] tried to build a model which predicts fraud and non fraud transactions with efficiency using machine learning algorithms. The prediction will be on account of time and the amount of the transaction using machine classifier models. Moreover they have used linear algebra, in constructing more complex ML models. The authors started by briefing about the need for such an efficient solution for Credit card fraud detection. Moreover they have described each algorithm's working process and their performance upon the DATASET. This research paper worked with a dataset containing the transaction form europe

card owners where 492 out of 2,84,807 are fraud transactions. The Dataset was then converted into PCA transformation thus it contains numeric values. To predict the result they have used statistical analysis to visualize the data. Furthermore, they have used a confusion matrix of all the ML models such as RF, NB, ANN, logistic regression for better comparison. Upon their thorough analysis, they have concluded that ANN has more accuracy of 98 percent from the rest of the ML models. The authors of this paper, Jadhav et al. [9] worked on modeling the dataset for credit card fraud detection. They tried to achieve detection of fraudsters and the number of them. They have used machine learning algorithms such as Artificial neural networks,Gradient Boosting Algorithms. Logistic regression and DTM. On analyzing logistic regression, and more of the ML models. The authors conclude in their study that the gradient boosting model has more accuracy than any of the ML models and also it helps in teaching how the credit card detection model can be improved.

III. Dataset

Dataset contains 2,84,807 transactions of Europe card owners. From here only 492 transactions are fraud. Thus, the dataset is not balanced. It has fewer fraud cases compared to the huge number of transactions. Also, the dataset is in PCA transformation only containing the numeric values. Here only time and amount are not converted into PCA value. All the other datas from Volume1(V1) to Volume28(V28) are converted to PCA values. For privacy concerns, many values are already given as PCA values. As for the feature value in this dataset, we have two feature class's value type. Containing 1 denotes fraud and 0 for normal transactions.

Fig. 1. Transection Dataset.

IV. Methodology

We are going to follow a few steps to understand the problem and the data. First, visualization and statistical analysis upon the data will be performed, to see how much of the data is imbalanced. Later on, to balance the data, oversampling and scaling will be used. standardization and normalization will also be used to scale the data. Furthermore, to run the ML models and visualize the dataset, Numpy, Matplotlib and seaborn libraries will be used respectively.

In our visionary approach to revolutionize credit card fraud detection, we strategically incorporate six avant-garde models: XGBClassifier, Logistic Regression, LSBM and KNN models along with Isolation Forest and Local Outlier Factor. Each model is carefully selected based on its innovative features, which raises the bar in our quest for a flexible and effective fraud detection system.

A. XGBoost

Extreme Gradient Boosting, or XGBoost, is a strong collaborative learning method renowned for its effectiveness and output. It functions by merging the predictions of several weak models, usually decision trees, to provide a final forecast that is reliable and accurate. To improve the model's accuracy, the boosting formula consists of repeatedly changing the weights of examples that are incorrectly identified.

XGBoost involves a weighted sum of decision trees. The prediction ($\hat{y}$) for a given instance is calculated as:

$$\hat{y} = \sum_{i=1}^{N} f_i(x) \tag{1}$$

where $f_i(x)$ represents the prediction of the individual decision trees.

We opt for XGBoost because of its capacity to manage intricate relationships in data, which makes it a good fit for detecting credit card fraud in situations with complex patterns. Additionally, it was picked due to its outstanding performance in a number of machine learning contests and practical uses. Because of its capability in managing intricate links within the data, it is especially well-suited for credit card fraud detection, where fraudulent patterns can be complex and dynamic. By integrating many decision trees, XGBoost's ensemble approach improves the predicted accuracy and resilience of the model. Furthermore, the way it handles outliers and missing data is flexible enough to meet the difficulties that are frequently present in datasets used for fraud detection.

B. Logistic Regression

An essential algorithm that is widely used as binary value classifier is logistic regression. In spite of its name, its purpose is to forecast the likelihood that an instance will belong to a particular class. It is a linearly combined input value that is converted into a range of 0 and 1, which represents the probability, via the function called logistic (sigmoid). Because of its ease of use and interpretability, logistic regression is a fundamental baseline model.

In logistic regression, the probability (p) of an instance belonging to the positive class is modeled using the logistic function:

$$p = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n)}} \tag{2}$$

where $\beta_0, \beta_1, \ldots, \beta_n$ are the coefficients and $x_1, x_2, \ldots, x_n$ are the

We include it in our study to establish a benchmark against more complex models and to provide insights into linear relationships within the data. Logistic Regression is a fundamental yet powerful algorithm, especially in binary classification tasks like fraud detection. It's selected for its simplicity, interpretability, and efficiency. Logistic Regression provides a baseline understanding of linear relationships within the data, helping us establish a benchmark against more complex models. The transparency of its coefficients also aids in interpreting the influence of individual features on the likelihood of fraud.

C. LightGBM

Another gradient boosting framework with an emphasis on efficiency and speed is called LightGBM. It uses 'Gradient-based One-Side Sampling' as an approach to effectively train on big datasets. The key to the model's success is its capacity to deal with categorical characteristics and scale effectively to large numbers of features and cases.

The general formula for boosting in LightGBM involves a weighted sum of decision trees similar to XGBoost. The specific details depend on the chosen objective function and parameters.

LightGBM is chosen for its balance between accuracy and computational efficiency, essential considerations for credit card fraud detection where large datasets and real-time processing are common. Because of its effectiveness, scalability, and capacity for handling huge datasets, it is included. LightGBM is a gradient boosting framework that performs well when trained on large datasets and keeps a high level of predicted accuracy. Because of its novel Gradient-based One-Side Sampling method, which improves computational performance, it is a good fit for applications like credit card fraud detection where real-time processing is essential. Given the variety of financial data, the model's ability to handle category characteristics is very useful.

D. K-Nearest Neighbors (KNeighborsClassifier)

K-Nearest Neighbours is a simple and basic method. It sorts instances according to the majority class of their k-nearest neighbors. To calculate proximity, the formula includes measuring distances between occurrences. The prediction in K-Nearest Neighbors is based on a majority vote of the k-nearest neighbors. For binary classification:

$$\hat{y} = \text{argmax} \left(\sum_{i=1}^{k} I(y_i = 1) \right) \tag{3}$$

where $I(y_i = 1)$ is an indicator function equal to 1 if $y_i = 1$ (positive class) and 0 otherwise.

K-Nearest Neighbours is included for its simplicity and efficacy, particularly in circumstances where instances with identical attributes tend to belong to the same class. It offers a counterpoint to more complicated models, providing insights into the function of proximity in fraud detection. This technique sheds light on the function of proximity in fraud detection. In contrast to more sophisticated models, K-Nearest Neighbours provides a plain perspective on the effect of neighboring examples on categorization outcomes. Because of its simplicity, it is also computationally efficient, allowing for rapid examination of small patterns within the data.

E. Isolation Forest

The idea behind the Isolation Forest model is to build random decision trees in order to isolate anomalies. Its unique method of separating outliers from less partitioned cases makes it extremely effective in outlier detection. It stands for an anomaly detection paradigm of accuracy and quickness. Designed to function as a group of cyber investigators using intuitively sharp random decision trees, the model quickly finds irregularities in a clever forest of choices. Its prowess extends beyond mere outlier detection; it operates with the finesse reminiscent of an algorithmic Sherlock Holmes.

The formulation of the anomaly score, s(x,n) bears a mark of mathematical elegance, symbolizing not just efficiency but an orchestrated demonstration of the algorithm's intrinsic ability to unravel the anomaly's seclusion within the data forest. The formula is-

$$s(x, n) = 2^{-\frac{E(h(x))}{c(n)}} \tag{4}$$

Where, E(h(x)) is the average path length for a given data point x. c(n) is a normalization term derived from the average path length of an unsuccessful search in a binary search tree with n nodes.

We incorporated Isolation Forest in our study to leverage its intrinsic efficiency and rapid anomaly isolation capabilities. The dynamic and continually evolving nature of credit card fraud patterns necessitates a model with the capacity to swiftly identify irregularities. The utilization of Isolation Forest aligns seamlessly with our objective of optimizing computational efficiency without compromising accuracy.

F. Local Outlier Factor (LOF)

LOF uses fluctuations in data point densities to award anomaly scores, working on the basis of the local density deviation principle. Because of this, LOF is especially good at identifying minute variations that point to fraudulent activity inside the complex world of credit card transactions.Beyond mere anomaly detection, LOF demonstrates a profound understanding of the intricacies inherent in local patterns, revealing anomalies with the cultural fluency akin to that of a data whisperer.The reachability distance, local reachability

density, and ultimate LOF score formulae, which regulate LOF, are similar to the brushstrokes of an expert painter.

Formulas:

$$rd(a,b) = \max\{k\text{-distance}(b), d(a,b)\} \tag{5}$$

$$lrd(a) = \frac{1}{\sum_{b \in N_k(a)} rd(a,b)} \tag{6}$$

$$LOF(a) = \frac{\sum_{b \in N_k(a)} \frac{lrd(b)}{lrd(a)}}{|N_k(a)|} \tag{7}$$

LOF's adaptability to local variations in data density enhances our fraud detection capabilities. By incorporating LOF, we aim to explore its effectiveness in capturing nuanced fraud patterns that might go unnoticed by models with a global perspective. This sensitivity aligns with our pursuit of a comprehensive and adaptive fraud detection system.

V. Result and Analysis

To address the crucial problem of credit card fraud detection, our study used six anomaly detection models: XGBClassifier, Logistic Regression, Local Outlier Factor (LOF), LightGBM, KNN and Isolation Forest. The accuracy of each model in differentiating between legitimate and fraudulent credit card transactions was subjected to a thorough evaluation.

A. Analysis on XGBClassifier

The XGBClassifier model proved to be a formidable tool for detecting credit card fraud because of its ability to handle intricate patterns found in transaction data. XGBClassifier showed its effectiveness in capturing complex relationships within the feature space with an astounding accuracy of 99.97%.

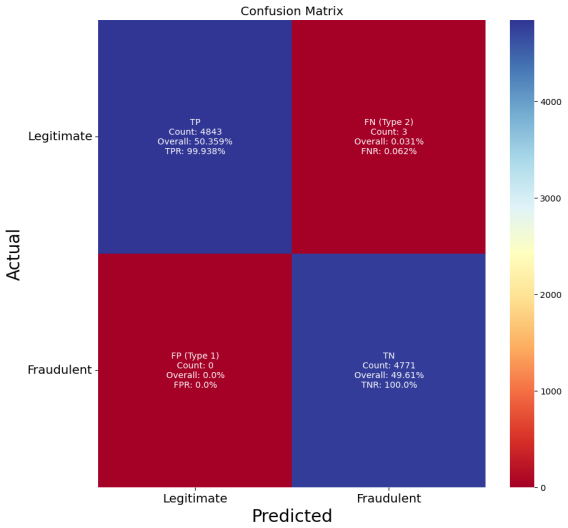

Fig. 2. Confusion Matrix for XGBClassifier

The capacity of the model to learn from and adjust to non-linear dependencies in the data was made possible by its ensemble nature, which made use of gradient boosting. Moreover, XGBClassifier's regularization methods prevented overfitting and guaranteed a balanced precision-recall performance. The model is particularly well-suited for tasks requiring strong pattern recognition and high accuracy, as demonstrated by its perfect precision, recall, and F1-scores, which demonstrate its mastery of both classes.

B. Analysis on Logistic Regression

In terms of interpretability and simplicity, Logistic Regression proved to be a reliable option, even though its accuracy was marginally lower 97.47% than that of the tree-based models. With a 96% accuracy rate, it demonstrated its remarkable ability to identify regular patterns in the classification of non-fraudulent transactions (class 0). For fraudulent transactions, however, the model's precision slightly decreased (class 1). Because it balances interpretability and precision, Logistic Regression is a useful option, especially when clear decision-making is essential. When model interpretability is a top concern in an application, the model's simplicity and ease of understanding make it a desirable choice.

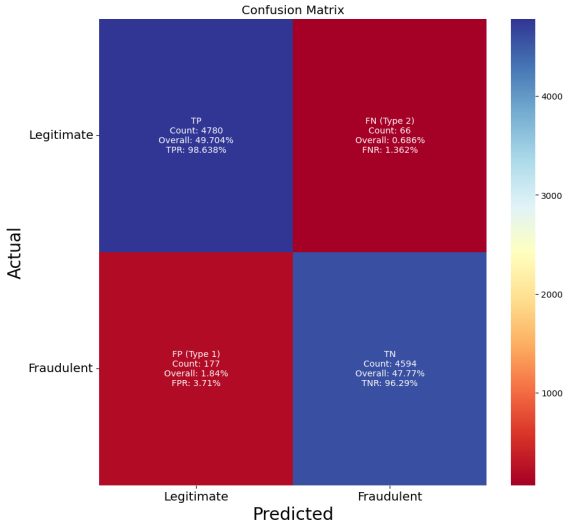

Fig. 3. Confusion Matrix for Logistic Regression

C. Analysis on LightGBM

LightGBM showed remarkable performance (99.97%) along with efficiency when processing large-scale datasets. The preprocessing procedures were simplified by the model's histogram-based tree construction method and its capacity to handle categorical features without requiring one-hot encoding. Perfect precision, recall, and F1-scores for both classes demonstrate LightGBM's high accuracy, which highlights its applicability for tasks involving credit card fraud detection. It is especially well-suited for applications where computational resources are an issue because of its speed and efficiency. The

model's performance demonstrates why using a histogram-based gradient boosting strategy is advantageous for tasks requiring accuracy and efficiency.

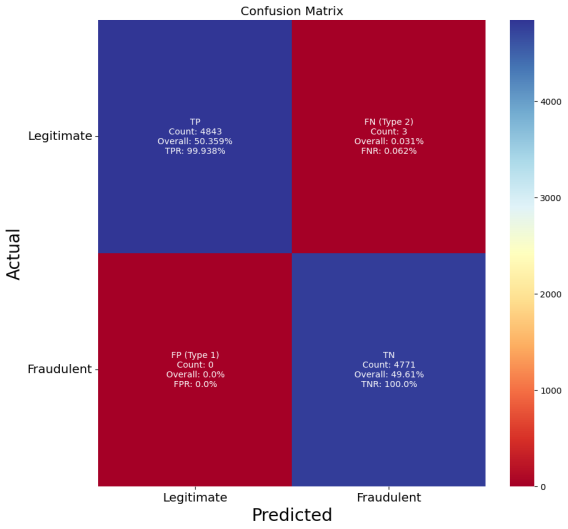

Fig. 4. Confusion Matrix for LGBM

D. Analysis on KNN

With an accuracy of 99.96%, KNN proved to be a strong rival, demonstrating its ability to identify patterns using similarity metrics. The model is an attractive option for tasks where patterns may be non-linear and irregular because of its ease of use and flexibility to different data distributions. It is imperative to acknowledge the possible computational expense linked with KNN, especially when dealing with larger datasets.

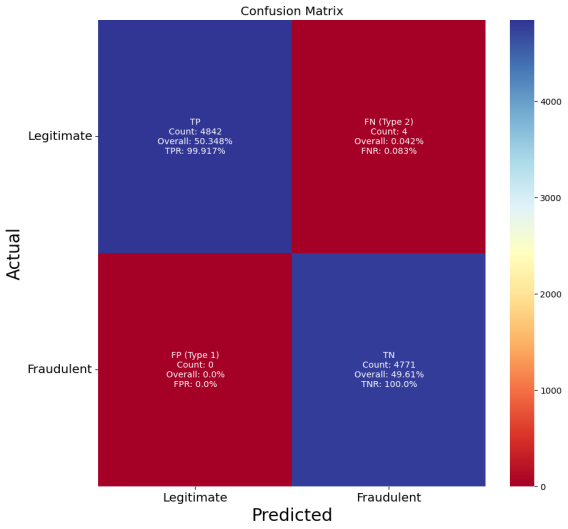

Fig. 5. Confusion Matrix for KNN

The model's efficacy in detecting fraudulent transactions based on instance similarity is highlighted by its perfect precision, recall, and F1-scores for both classes. When interpretability and simplicity are important, KNN's non-parametric nature makes it a competitive alternative in the detection of credit card fraud.

E. Analysis on Isolation forest and LOF

In the beginning, the credit card fraud detection task produced results with almost 99% accuracy from both the Local Outlier Factor (LOF) and Isolation Forest models. On the other hand, it was clear from a closer look at their classification reports that they performed poorly in class 1 (fraudulent transaction prediction). Due to the dataset's extreme imbalance, which included a small percentage of fraudulent transactions, the models tended to prioritize accuracy by categorizing most transactions as non-fraudulent (class 0).

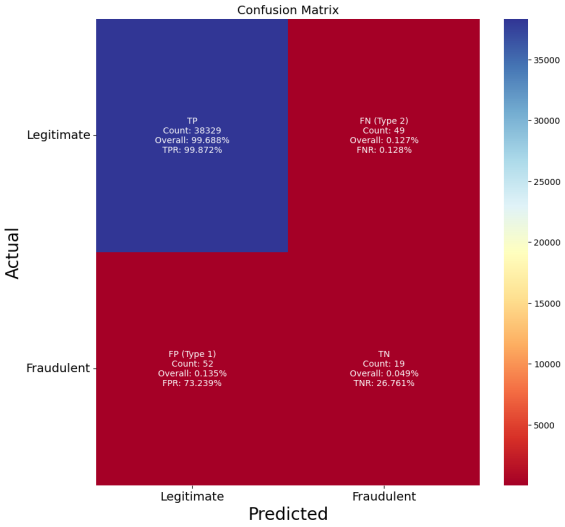

Fig. 6. Confusion Matrix for LOF

The previously given results are the consequence of an attempt to rectify this imbalance by balancing the dataset. Although 50.11% and 50.15% accuracy, respectively, were obtained for LOF and Isolation Forest, the classification reports show ongoing difficulties in accurately detecting and forecasting fraudulent activity. These models are still unable to identify the subtle patterns linked to credit card fraud, with precision, recall, and F1-scores for class 1 remaining low or close to zero. This demonstrates the complexity of fraud detection tasks and emphasizes the need for additional research and different modeling strategies to improve performance when data is unbalanced and exactly that is what we did and described above.

Performance Table						
	LOF	KNN	LightGBM	LR	XGB	IF
Accuracy	50.10%	99.95%	99.96%	97.47%	99.96%	50.15%
Precison	0.5	1.00	1.00	0.96	1.00	0.5
Recall	0.99	1.00	1.00	0.99	1.00	1.00
F1 score	0.67	1.00	1.00	0.98	1.00	0.67

TABLE I
COMPARISON OF 5 DIFFERENT MODELS

When it comes to computational efficiency, LightGBM and XGBClassifier showed exceptional speed, which makes them the best choices for large-scale applications. Even though it required less computing power, logistic regression demonstrated a compromise between accuracy and readability. KNN performed exceptionally well at capturing patterns due to its simplicity and flexibility, but it may have issues with scalability. Logistic regression was the clearest choice in terms of interpretability, offering insights into the process of making decisions. On the other hand, despite being ensemble models, XGBClassifier and LightGBM could be viewed as "black-box" systems due to their high accuracy and efficiency.

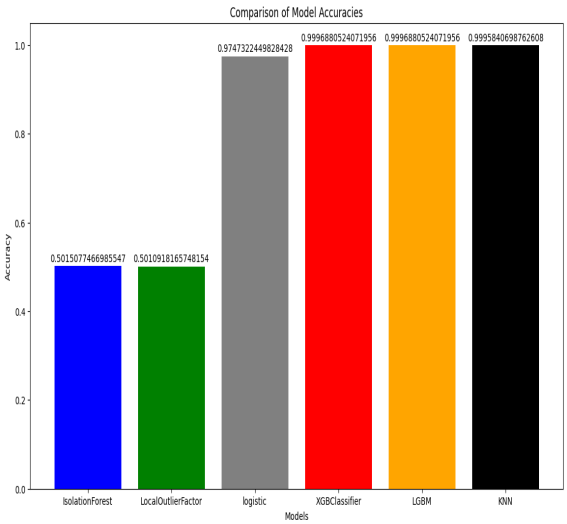

Fig. 7. Accuracy Comparison

The specific requirements of the credit card fraud detection task should be carefully considered before selecting one of these models. Transparency, computational efficiency, or attaining maximum predictive accuracy are the three main priorities, and each model offers unique benefits tailored to a particular application.

VI. Challenges

The highly unbalanced nature of the credit card transaction dataset was one of the main difficulties our research team faced. The anomaly detection methods faced intrinsic challenges due to the notable disparity between the quantity of fraudulent and non-fraudulent cases. Specifically, the models showed a tendency to perform exceptionally well on the majority class (transactions that are not fraudulent), frequently attaining near-perfect accuracy, precision, and recall for these cases. On the other hand, the unequal distribution resulted in less than ideal identification of the minority class (fraudulent transactions). Due to the small number of positive examples, the models had difficulty generalizing patterns linked to fraudulent activities, which led to lower precision, recall, and F1-score values for the minority class. This imbalance made it difficult to develop a thorough and balanced credit card fraud detection system, so it was important to carefully consider the best methods for resolving the imbalance.

VII. Conclusion

In conclusion, this study evaluated machine learning models for credit card fraud detection in a methodical manner. Even with the dataset balancing, Local Outlier Factor (LOF) and Isolation Forest still had issues, whereas XGBClassifier, LightGBM, and KNN showed strong performance. The difficulties LOF and Isolation Forest encountered demonstrate how complex fraud detection jobs can be. The results highlight how crucial careful model selection and preprocessing are to correcting imbalances.

Based on our research, we prioritized XGBoost to fight credit card theft due to its amazing accuracy rate of 99.97%. The optimization work focused on developing ensemble methodologies and enhancing methods for feature design in order to better overall model performance. XGBoost is a good choice for identifying fraud activities due to its capacity to handle complex interactions within the data. When it came to accuracy, XGBoost performed better than each of the models evaluated for this study. LightGBM came in close second with a match accuracy of 99.97%, demonstrating its efficacy on large-scale datasets. K-Nearest Neighbors (KNN) likewise performed well, achieving 99.96% accuracy. Even though it was slightly lower, at 97.47%, the logistic regression model provided a reliable and comprehensible baseline. In this case, accuracy—the percentage of correctly classified occurrences relative to all instances—is the efficiency metric. When it comes to credit card fraud detection, accuracy is a crucial performance metric since it demonstrates how well the model can discriminate between legitimate and fraudulent transactions. But, if fraudulent transactions are rare compared to legitimate ones, other metrics like precision, recall, and F1-score are just as important to include for a comprehensive assessment of unbalanced datasets.

VIII. Future Work

For future research, exploring advanced ensemble methods, refining feature engineering techniques, and deploying deep learning architectures like autoencoders could enhance credit card fraud detection. Investigating temporal aspects, transaction sequences, and user behavior patterns may provide valuable insights. Additionally, focusing on model interpretability and explainability methods is crucial for building trust and facilitating real-world adoption. Developing transparent and interpretable models ensures stakeholders can comprehend decision-making processes, fostering confidence in the reliability of credit card fraud detection systems.
Future studies should also investigate novel approaches to improve performance when dealing with unbalanced and complex datasets, supporting continuous efforts to fortify

financial systems against fraud.

REFERENCES

[1] E. Ileberi, Y. Sun, and Z. Wang, "A machine learning based credit card fraud detection using the ga algorithm for feature selection," *Journal of Big Data*, vol. 9, 02 2022.

[2] S. Jiang, R. Dong, J. Wang, and M. Xia, "Credit card fraud detection based on unsupervised attentional anomaly detection network," *Systems*, vol. 11, p. 305, 06 2023.

[3] J. Chung and K. Lee, "Credit card fraud detection: An improved strategy for high recall using knn, lda, and linear regression," *Sensors*, vol. 23, p. 7788, 09 2023.

[4] N. Alfaiz and S. Fati, "Enhanced credit card fraud detection model using machine learning," *Electronics*, vol. 11, p. 662, 02 2022.

[5] J. Roseline, G. Naidu, V. Pandi, S. Rajasree, and D. Mageswari, "Autonomous credit card fraud detection using machine learning approach," *Computers and Electrical Engineering*, vol. 102, p. 108132, 09 2022.

[6] F. A. Almarshad, G. A. Gashgari, and A. I. A. Alzahrani, "Generative adversarial networks-based novel approach for fraud detection for the european cardholders 2013 dataset," *IEEE Access*, vol. 11, pp. 107348–107368, 2023.

[7] I. Journal, "Credit card fraud detection using machine learning," Jan 2023.

[8] V. S, "Credit card fraud detection using machine learning algorithms," *International Journal of Engineering Research and*, vol. V9, 08 2020.

[9] I. Journal, "Credit card fraud detection using machine learning," Sep 2021.