# Analysis and Comparison of Fraud Detection on Credit Card Transactions Using Machine Learning Algorithms

AHMAD UMAR BARMO[1], AHMAD HARUNA[2], YUSUF UMAR WALI[3], KONIKA ABID[4]
[1, 2, 3] *Dept. of Computer Science & Engineering Sharda University, Greater Noida, India*
[4] *Assistant Professor, Dept. of Computer Science & Engineering Sharda University, Greater Noida, India*

***Abstract-*** ***Financial organizations and customers are both very concerned about fraud using credit cards as the use of digital payment methods keeps growing. Strong and effective credit card fraud detection systems are essential given the prevalence of complex fraud schemes and the rising number of online transactions. According to transaction statistics, there are more instances of credit card fraud each year. Researchers are thus actively looking on cutting-edge techniques to identify and shut down these fraudulent enterprises. In an effort to safeguard financial institutions and protect consumers from possible losses, they are focused on utilizing cutting- Techniques for better credit card fraud detection and prevention. The purpose of this study is to focus on machine learning approaches. Logistic regression, K-nearest neighbor, and Naive Bayes with stacked model are the techniques used. The algorithms' output is based on f1 score, recall, accuracy, and precision. As well, the ROC curve is plotted. The evaluation of three credit card fraud detection models revealed distinct strengths and weaknesses. While the Naive Bayes model exhibited the highest overall accuracy (99.7%) and F1 score (37.5%), prioritizing precision and recall led to nuanced considerations. Notably, the introduction of an ensemble learning model, stacking Logistic Regression, K-Nearest Neighbor, and Naive Bayes, significantly boosted overall accuracy to an impressive 98.58%, showcasing the potential for enhanced performance through model combination.***

***Indexed Terms-*** ***Logistic Regression, Naïve Bayes, K-Nearest Neighbor, Fraud Detection, Stacked Generalization***

## I. INTRODUCTION

Credit card fraud is a pervasive issue that is constantly evolving and poses major risks to businesses, customers, and financial institutions all over the world. Credit card fraud, put simply, is when someone uses someone else's credit card without the cardholder's or the card issuer's knowledge for personal expenses. As digital payment methods become more prevalent in our daily lives, the likelihood of fraud is increasing exponentially. Fraudsters constantly devise new strategies to exploit flaws in payment systems, which causes enormous financial losses and undermines consumer confidence. Nowadays, almost everyone uses a credit card, and as a result, fraud is rapidly rising. Transactions may now be made quickly and easily thanks to technology. With only their cell phones or the internet, people may conduct hundreds of transactions. Online credit card purchases are commonplace today. The game is finished if someone accidentally obtains card information since fraud may be readily done. Several safeguards must be in place to prevent the disclosure of credit card information in order to prevent credit card fraud. Credit card information may be compromised in a number of ways, including phishing, card data theft, stolen cards, and others. Using data mining techniques is a crucial means of identifying credit theft. Credit card fraud detection is the activity of classifying fraudulent transactions into two types: legitimate transactions and fraudulent transactions[1]. Most credit card fraud detection systems incorporate artificial intelligence, meta learning, and pattern matching[2].

## II. RELATED WORK

In the realm of credit card fraud detection, various studies have delved into the efficacy of machine learning algorithms, each offering valuable insights. Andhavarapu Bhanusri et al.

[3] compared Naïve Bayes, Logistic Regression, and Random Forest with boosting, finding the latter to

outperform the others in accurately discerning fraudulent transactions. Dejan Varmedja et al. [4] addressed imbalanced datasets using Random Forest, Naïve Bayes, and multilayer perception, employing the Synthetic Minority Over-sampling Technique (SMOTE) for oversampling. They highlighted Random Forest as the most effective and emphasized feature selection and dataset balancing. Pratyush Sharma et al. [5] cautioned against relying solely on accuracy due to dataset imbalance, favoring the Random Forest classifier. M. Ummul Safa et al.

[6] evaluated three methods, with Logistic Regression achieving the highest accuracy. John O. Awoyemi et al. [1] utilized a hybrid approach and identified the K-nearest neighbor algorithm as the best performer. Yogesh Kumar et al. [7] proposed a study employing Random Forest, Logistic Regression, and Support Vector Machine, with Random Forest exhibiting the highest accuracy. Pradheepan Raghavan et al. [8] explored six machine learning algorithms, favoring SVM for larger datasets and convolutional neural networks for shorter datasets, while acknowledging challenges in dynamic settings. Collectively, these studies contribute to the evolving landscape of credit card fraud detection by assessing algorithmic performance and addressing dataset intricacies.

## III. PROPOSED WORK

During the data gathering phase, highly skewed credit card fraud datasets are collected from trusted sources such as financial institutions or credit card firms. Subsequently, during data pre-processing, the collected data undergoes relevant feature removal and handling of missing values, with an additional step to balance class distribution through sampling techniques like oversampling or undersampling. Naive Bayes, k-nearest neighbor, and logistic regression were chosen for model selection since they are widely used in credit card fraud detection literature, setting the stage for a meaningful comparison. Following this, the selected models are trained using preprocessed data and cross-validation techniques to enhance generalization and prevent overfitting. The evaluation phase involves assessing model performance based on metrics like accuracy, time length, and balanced classification rate. Finally, the study closes with a detailed comparison of the three models, resulting in the determination of the most efficient machine learning method.

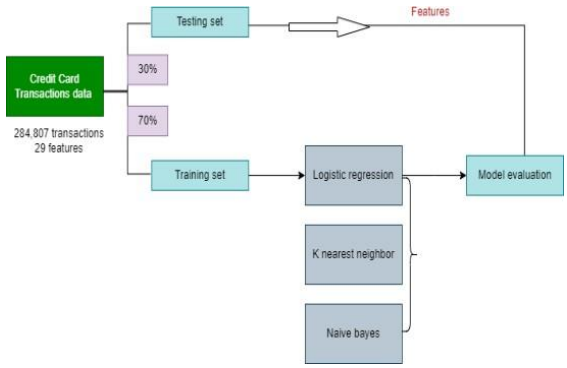


Fig 1Process flow

Stacking is a way to ensemble multiple classifications or regression model. In this paper we combined all the three models with the stacked model to improve the accuracy and efficiency of the models.

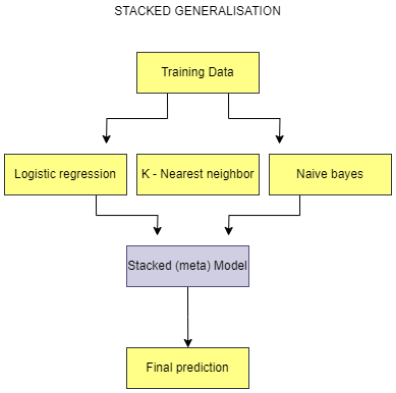


Fig 2 Architecture diagram

### A. *Dataset*

The transaction records data used in this particular research was sourced from Kaggle[9]. It houses records of transactions that was done by European card bearers. The transactions recorded were for two days[10], out of which only 492 out of 284,807 were frauds. This means the data is very skewed (frauds are just 0.172% of all transactions). The data has numbers that came from a math technique called PCA. But we don't know what the numbers mean or where they came from, because that is secret. While the "Time" and"Amount" features were not altered and continue to have their original meanings, the features V1 through V28 reflect principal components

obtained from PCA. The feature "Amount" literally means the amount of the transaction, "Time" means the number of seconds between each transaction and initial transaction of the dataset. The target variable is that of the 'Class' feature, which takes value 1 if the transaction is fraudulent and 0 otherwise.

B. *Logistic regression*

Logistic regression[5][11] is a binary classification technique. It is a sort of generalized linear model (GLM) that is frequently used in statistics and machine learning predicting the chance of an occurrence based on one or more predictor factors. When a target variable can have one of two potential values, commonly expressed as 0 and 1 (or "negative" and "positive," "no" and "yes," etc.), logistic regression is frequently used to solve binary classification issues. The cost function employed in logistic regression, also known as the "sigmoid function," is what we use to convert predicted values into probabilities. When we pass the inputs through a prediction function, the classifier will provide us with a set of probability-based outputs or groups and provide a probability score between 0 and 1[7]. For the purpose of logistic regression, a log-odds or logit function is computed by linearly combining predictor variables (features).

C. *K-nearest neighbour*

K-Nearest Neighbors[11] is a basic and straightforward machine learning technique that may be used for classification and regression problems[12] Being non- parametric and instance-based, it doesn't make any substantial assumptions about the distribution of the underlying data and instead derives predictions from the training set. K-NN uses a distance metric (such as the Euclidean distance, Manhattan distance, and so on) to determine how similar or distant two data points are to one another. The distance measure chosen is determined on the data and task at hand. K-NN assigns a class label to a new data point by majority voting among its K nearest neighbors while performing classification tasks. In other words, it calculates the number of neighbors who belong to each class and chooses the class with the highest number as the anticipated class.

D. *Naïve bayes*

An ML well-known probabilistic classification approachissues is the Naive Bayes method. Using the likelihood of anearlier occurrence, the Bayes theorem determines the likelihood of a subsequent event. Naive Bayes is very beneficial for text categorization and spam filtering, although it may be used for a variety of classification issues. Naïve Bayes' "naive" presumption is that, given the class label, all features are conditionally independent. This suggests that one trait's presence or importance is unrelated to another's. Although this presumption might not always be accurate in data from the real world, it makes computations simpler and performs very well in practice.

E. *Stacked generalization*

Stacked Generalization, often known as stacking, is an approach to ensemble learning that uses a meta-model (or higher-level model) to integrate the predictions of numerous base models (learners). By integrating the capabilities of several base models, stacking aims to improve predictive performance by teaching a meta-model how to integratepredictions in the best possible way. The concept behind stacked generalization is that it can capture the strengths of several base models and learn how to properly weigh their predictions. When compared to employing individual base models alone, this frequently leads in better prediction performance. Stacking is a flexible approach that may be used to perform a variety of machine learning tasks such as classification and regression. Stacking is a strong technique in machine learning ensemble techniques that is used to improve the accuracy and robustness in predictions.

## IV. EVALUATION CRITERIA

Metrics including accuracy, precision recall, and f1-score are evaluated in this research in order to compare various algorithms. This experiment's evaluation is based on four fundamental measures. The acronyms for these terms are True positive (TP), True negative (TN), False positive (FP), and False negative (FN).

True positive (TP): True positives are situations in which the model accurately predicts a transaction as fraudulent and thetransaction is truly fraudulent.

True Negative (TN): True Negative depicts the circumstances in which the model properly predicts a transaction as valid and it is, in fact, a legitimate transaction. False Positive (FP): False Positive shows the situations in which the model mistakenly predicts a transaction as fraudulent when it is really valid.

False Negative (FN): FN denotes instances in which the model mistakenly predicts a transaction as genuine when it is really fraudulent. The Receiver Operating Characteristicscurve is formed by graphing the TPR versus the FPR. This can be done at various thresholds. A ROC curve shows how well a model can tell apart two classes. The x-axis is the false positive rate (FPR), which is how often the model wrongly says "yes". The y-axis is the true positive rate (TPR), which is how often the model correctly says "yes". The area under the curve (AUC) is the space below the ROC curve. It measures how good the model is overall[13].

Confusion Matrix: A confusion matrix tells us more about how well a model works. It shows us which categories themodel gets right and which ones it gets wrongly predicted aswell as the kind of errors being made[14]. It provides atransparent breakdown of the model's predictions and actual results, enabling in assessing the model's capability to reliably distinguish between legitimate and fraudulent transactions.

Accuracy: Accuracy is a measure of how many transactions, both fraudulent and genuine, were properly categorized by the model.

Precision: Precision is the percentage of projected fraudulenttransactions that were really fraudulent.

Recall: Recall, also known as Sensitivity or True Positive Rate (TPR), is the fraction of real fraudulent transactions successfully recognized by the model.

F1 score: The F1 Score is a number that combines how accurate and how complete a model is.

## V. RESULT & DISCUSSION

This research tried four methods—logistic regression, k- nearest neighbor, naive bayes, and stacking model—to predict the data. 70% of the data are used for training and the remaining 30% are used for testing in the evaluation of these models. Different criteria for model comparison have been used to determine which algorithm is the most effective at spotting fraudulent transactions. The most commonly used measurements for machine learning algorithms are accuracy,precision, recall, and f1 score. According to these parameters,the performance of these models was assessed.

```
Logistic Regression Model:
Accuracy: 0.9821518439193381
Precision: 0.07574832009773977
Recall: 0.9117647058823529
F1 Score: 0.1398759165256627
```

Fig 3 Logistic regression

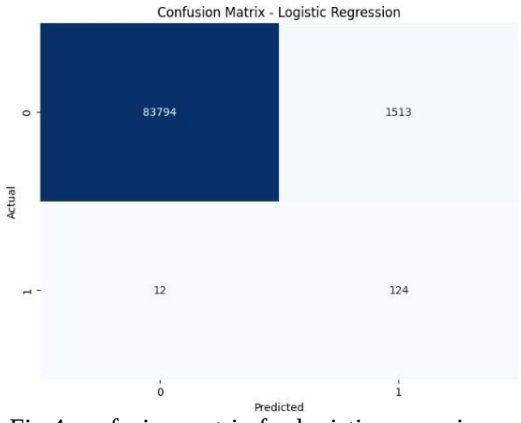


Fig 4 confusion matrix for logistic regression

The confusion matrix of for *Fig 4* shows that there are 83794true positives and 1513 false positives, the true negatives are12 and the false negatives are 124.

```
K-Nearest Neighbor (KNN) Model:
Accuracy: 0.9404281216717578
Precision: 0.013550667714061273
Recall: 0.5073529411764706
F1 Score: 0.02639632746748278
```

Fig 5 K-Nearest Neighbor

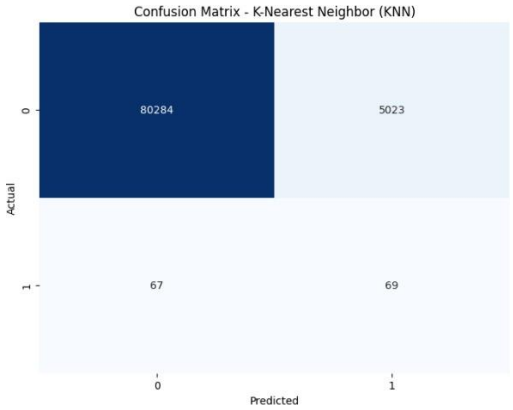


Fig 6 confusion matrix for K- Nearest neighbor

The confusion matrix of *Fig 6* shows that there are 80284 truepositives, 5023 false positives, there are 67 false negatives and 69 true negatives.

```
Naive Bayes Model:
Accuracy: 0.9966059244174479
Precision: 0.2652439024390244
Recall: 0.6397058823529411
F1 Score: 0.375
```

Fig 7 naive bayes

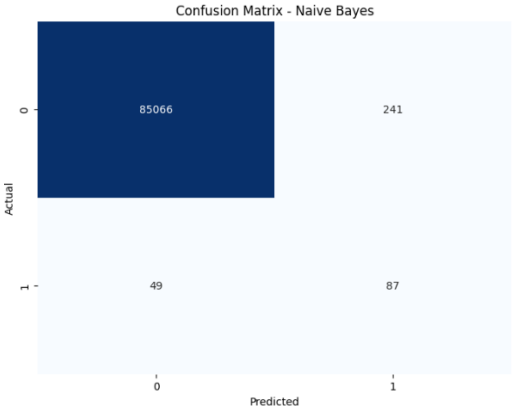


Fig 8 confusion matrix for naive bayes

The confusion matrix for *Fig 8* shows that there are 85066 true positives, 241 false positives, 49 false negatives and 87 true negatives.

```
Stacked Model:
Accuracy: 0.9858034010978078
Precision: 0.09173616376042457
Recall: 0.8897058823529411
F1 Score: 0.1663230240549828
```

Fig 9 Stacked Model

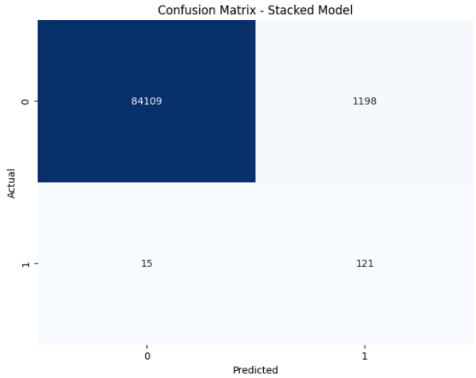


Fig 10 confusion matrix for stacked model.

The matrix for *Fig 10* shows that there are 84109 true positives, 1198 false positives, there are 15 false negatives and 87 true negatives.

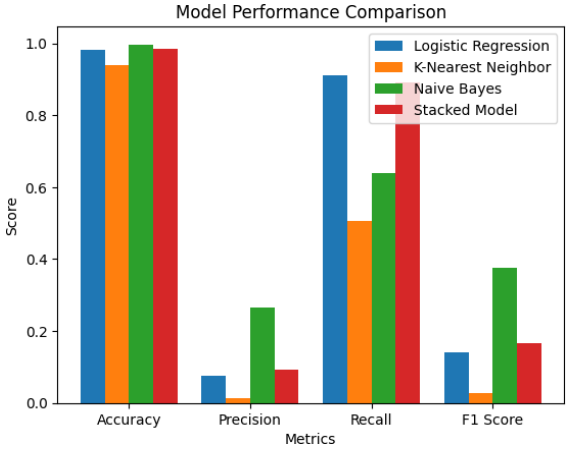


Fig 11 Comparison of the algorithms

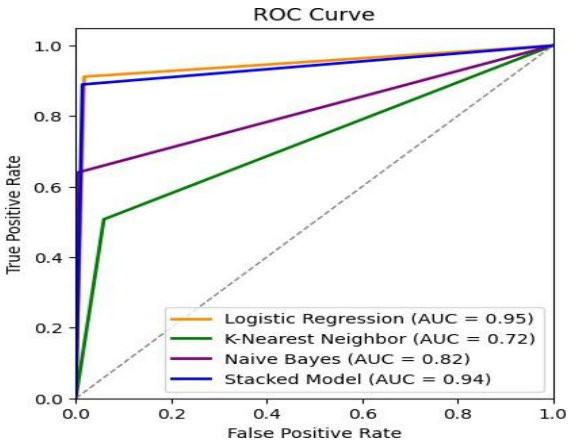


Fig 12 ROC curve for all the models

Table 1 Model Comparison results

| Metrics | Classifiers | | | |
|---|---|---|---|---|
| | *Logistic Regression* | *K-Nearest Neighbor* | *Naïve bayes* | *Stacked model* |
| Accuracy | 0.9821 | 0.9404 | 0.9966 | 0.9858 |
| Precision | 0.0757 | 0.0135 | 0.2652 | 0.0917 |
| Recall | 0.9117 | 0.5073 | 0.6397 | 0.8897 |
| F1 score | 0.1398 | 0.0263 | 0.375 | 0.1663 |
| Roc–Auc | 0.95 | 0.72 | 0.82 | 0.94 |

## CONCLUSION

On a test dataset, the three models' `effectiveness at detectingcredit card fraud was assessed. The accuracy of the Logistic Regression model was 98.2%, correctly recognizing 91.2% of the actual fraudulent transactions (recall), while its precision was just 7.6%. The K-Nearest Neighbor (KNN) model has an extremely low precision and recall of 1.4% and 50.7%, but it nevertheless managed to attain an accuracy of 94.0%. The Naive Bayes model has a precision of 26.5%, a recall of 63.9%, and the greatest accuracy of 99.7%. The Naive Bayes model had the highest F1 score, which balances recall and precision, at 37.5%. The confusion matrices showthat the Naive Bayes model performed best in terms of correctly predicting both fraudulent and non-fraudulent transactions, while the KNN model struggled with false positives and negatives. Overall, the Naive Bayes model showed the most promising performance for detecting creditcard fraud among the three models.

When selecting the right model, it is crucial to take the application's unique requirements into account. The Naive Bayes model would be preferred if reducing false positives (misclassifying legitimate transactions as fraudulent) is a top concern. On the other hand, despite its greater false positive rate, the Logistic Regression model might be a preferable option if properly identifying real fraudulent transactions is essential.

An ensemble learning model (Stacked generalization) was introduced later on to this research paper. This meta model was achieved by combining all of the threemodels (Logistic regression, KNN & Naïve bayes), it achieved a whooping percentage of 98.58% accuracy for all the three models leading to an increase in performance to all of the other metrics. Future work is anticipated to involve further hyperparameter adjustment and fine-tuning in order to perhaps increase robustness and performance overall. Additionally, including interpretability strategies like SHAP or LIME can reveal the models' thought processes, boosting openness and confidence in the system for detecting credit card fraud[6].

## REFERENCES


[1] S. Misra, V. O. Matthews, A. Adewumi, O. S. Covenant University (Ota, IEEE Nigeria Section, and Institute of Electrical and Electronics Engineers, *Proceedings of the IEEE International Conference on Computing, Networking and Informatics (ICCNI 2017) : 29-31 October, 2017, Covenant University, Canaanland, Ota, Ogun State, Nigeria*.

[2] M. Ansari, H. Malik, S. Jadhav, and Z. Khan, “Credit Card Fraud Detection.” [Online]. Available: www.ijert.org

[3] A. Bhanusri *et al.*, “Credit card fraud detection using Machine learning algorithms,” 2020. [Online]. Available:www.questjournals.org

[4] H. Anand, R. Gautam, R. Chaudhry, and R. Chaudary, “EasyChair Preprint Credit Card Fraud Detection Using Machine Learning Credit Card Fraud Detection using Machine Learning,” 2021.

[5] P. Sharma, S. Banerjee, D. Tiwari, and J. C. Patni, “Machine learning model for credit card fraud detection-A comparative analysis,” *International Arab Journal of Information Technology*, vol. 18, no. 6, pp. 789–796, Nov. 2021, doi: 10.34028/iajit/18/6/6.

[6] “M. U. Safa, ‘Credit Card Fraud Detection Using Machine Learning,’ in International Journal of Research in Engineering, Science and Management, IJRESM, Nov. 2019, p. 3.”.

[7] Y. kumar, S. Saini, R. Payal, and A. Professor,

"Comparative Analysis for Fraud Detection Using Logistic Regression, Random Forest and Support Vector Machine," *International Journal of Research and Analytical Reviews*, vol. 7, no. 4, 2020, [Online]. Available: www.ijrar.org

[8] P. Raghavan and N. El Gayar, "Fraud Detection using Machine Learning and Deep Learning," in *Proceedings of 2019 International Conference on Computational Intelligence and Knowledge Economy, ICCIKE 2019*, Institute of Electrical and Electronics Engineers Inc., Dec. 2019, pp. 334–339. doi: 10.1109/ICCIKE47802.2019.9004231.

[9] Univerzitet u Istočnom Sarajevu. Faculty of Electrical Engineering, IEEE Industry Applications Society, Institute of Electrical and Electronics Engineers. Bosnia and Herzegovina Section, Institute of Electrical and Electronics Engineers. Serbia and Montenegro Section, and Institute of Electrical and Electronics Engineers, *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH) : proceedings : March 20-21, 2019, Jahorina, East Sarajevo, Republic of Srpska, Bosnia andHerzegovina*.

[10] D. Tanouz, R. R. Subramanian, D. Eswar, G. V. P. Reddy, A. R. Kumar, and C. H. V. N. M. Praneeth, "Credit card fraud detection using machine learning," in *Proceedings - 5th International Conference on Intelligent Computing and Control Systems, ICICCS 2021*, Institute of Electrical and Electronics Engineers Inc., May 2021, pp. 967–972. doi: 10.1109/ICICCS51141.2021.9432308.

[11] P. Tiwari, S. Mehta, N. Sakhuja, J. Kumar, and A. K. Singh, "Credit Card Fraud Detection using Machine Learning: A Study," Aug. 2021, [Online]. Available: http://arxiv.org/abs/2108.10005

[12] N. S. Alfaiz and S. M. Fati, "Enhanced Credit Card Fraud Detection Model Using Machine Learning," *Electronics (Switzerland)*, vol. 11, no. 4, Feb. 2022, doi: 10.3390/electronics11040662.

[13] Vaigai College of Engineering and Institute of Electrical and Electronics Engineers, *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020) : 13-15 May, 2020*.

[14] A. Mohari, J. Dowerah, K. Das, F. Koucher, and D. Jyoti Bora, "A COMPARATIVE STUDY ON CLASSIFICATION ALGORITHMS FOR CREDIT CARD FRAUD DETECTION." [Online]. Available: www.irjmets.com