International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930

# FraudShield: Detection of Fraud in Credit Card based on Machine Learning Techniques with integration of web-based Framework

*Tejas Prakash D*
*20191IST0163*
*Dept. of IST*
*Presidency University*
*Banglore-64 India*
*201910100303@presidencyuniversity.in*

*Grishma A V*
*20191IST0182*
*Dept. of IST*
*Presidency University*
*Banglore-64 India*
*201910102130@presidencyuniversity.in*

*Sneha M*
*20191IST0150*
*Dept. of IST*
*Presidency University*
*Banglore-64 India*
*201910100331@presidencyuniversity.in*

*Shweta A Nair*
*20191IST0157*
*Dept. of IST*
*Presidency University*
*Banglore-64 India*
*201910101983@presidencyuniversity.in*

*Sai Meghana J S*
*20191IST0187*
*Dept. of IST*
*Presidency University*
*Banglore-64 India*
*201910101615@presidencyuniversity.in*

*Dr. Kuppala Saritha*
*Associate Professor*
*School of CSE*
*Presidency University*
*Banglore – 64 India*
*kuppala.saritha@presidencyuniversity.in*

## ABSTRACT

In this paper, a general term for both fraud and theft committed with or while using a credit card for payment is "credit card fraud." In light of the daily increase in scammers, numerous different types of fraud are committed using credit cards. In order to address this issue, numerous methods like logistic regression, decision trees, KNN, and Naive Bayes algorithms are used. Various options are considered for this transaction, and the best one is implemented. Filtering out the above-mentioned tactics is done mostly to achieve the goal of detecting fraud and producing better results. Credit card fraud has increased due to the growing popularity of using credit cards for online purchases. The problem has been addressed through the use of machine learning algorithms to identify fraudulent transactions made with credit cards. In this paper, suggesting a system for detecting credit card fraud that employs machine learning techniques to distinguish between legitimate and fraudulent transactions. In order to examine previous transaction data and spot trends that point to fraudulent activity, the proposed system combines both supervised and unsupervised methods of learning. Additionally combining the developed machine learning methods with a web-based platform to deliver an intuitive user interface for immediate form of fraud detection. The test findings

demonstrate that the suggested framework detects fraudulent transactions with a high degree of accuracy and with a minimal amount of false positives.

**Keywords:** Fraud Detection, Credit Card, Decision Tree, Logistic Regression, KNN, Naïve-Bayes, Streamlit, MLP, Accuracy, Precision, Macro average, Weighted average.

## 1. INTRODUCTION

The consumers and financial organisations are equally impacted by the severe issue of credit card theft. Fraudulent actions can harm the credibility of the financial company as well as result in significant monetary losses for both parties. In order to promptly identify forged transactions and reduce losses, it is necessary to design successful and effective fraud detection systems. Credit card fraud has been successfully identified using machine learning techniques. These methods entail building a model from a dataset of confirmed illegal and fraudulent transactions, then using the model to forecast the likelihood of fraud for new transactions. In comparison to conventional rule-based systems, the application of techniques based on machine learning in the identification of fraud in credit cards has a number of benefits and irregularities in massive datasets that human analysts could miss. In contrast to rule-based systems, which need manual updates to be successful, they are also able to adjust to a new fraud patterns as they appear. In this regard, the project's goal is to look into how well different machine learning approaches work to identify credit card fraud. In order to determine which method is most efficient, the research will examine multiple datasets and use a variety of preprocessing strategies, feature selection techniques, and modelling algorithms.[1,2]

## 2. OBJECTIVES

The objectives that come across in fraud detection of credit cards are: (a) To identify and foresee the outcome of unauthorised credit card activity. (b) Analyse a few effective machine learning algorithms, identify the one with the best accuracy, and suggest a model. (c) Add a machine learning model to the web-based framework for better user interface and user experience. (d) Find pertinent dataset features that can aid in the detection of fraud. Identify relevant attributes that capture the patterns and traits of fraudulent transactions by extracting and engineering them. (e) Create a system that can process incoming transactions made with credit cards in real-time and identify whether they are counterfeit or genuine using the used machine learning models [2].

## 3. METHODOLOGY

### 3.1 Existing methods

In the current system, research on an instance involving detecting credit card fraud, where data normalisation was used before cluster analysis, and with outcomes obtained through the use of clustering and neural networks based on fraud detection, demonstrated that by accumulating qualities, neuronal activity inputs can be minimised. Additionally, normalised data should be used, and MLP training is recommended [3]. This study was built on unsupervised learning.

Finding innovative strategies for identifying fraudulent activity and improving the accuracy of outcomes were the two main purposes of this article. Personal information in the data set used for this study is kept isolated and it is based on the real transactional figures collected by a major European corporation. The algorithm typically has a 50% accuracy rate. Discovering an algorithm and dropping the cost measure were the two main purposes of this paper. The result was 23%, and the chosen algorithm had the lowest risk. [2,3].

**Disadvantages**

1. The gains and losses attributable to fraud detection are adequately represented in this study by a novel collative comparison metric.

2. The suggested cost measure is used to offer a cost-sensitive strategy centred around Bayes minimum risk.

### 3.2 Proposed method

Proposing a model in the system that is being suggested here to identify fraud behaviour in transactions with credit cards. The bulk of essential characteristics which are required to distinguish between legitimate and illegal transactions may be offered by this method. With the development of technology, it becomes more difficult to identify the idea and pattern of faked transactions. The advancement of artificial intelligence (AI), machine learning, and other relevant information technology disciplines has made it possible to automate this process and minimise part of the intense labour that is necessary to detect credit card fraud [4]. Teo identify credit card fraud. To discover which machine learning algorithm is best, comparisons are made between many algorithms, including random forests, decision trees, logistic regression, and Naive Bayes. Determine the best algorithm that credit card merchants can use to identify fraudulent transactions. Finally, integrating the machine learning model with the web based framework using streamlit, it is a web based framework for better user interface and user experience. Then creating menus, inputs fields for prediction, classification reports and display model graph in the web framework [5].

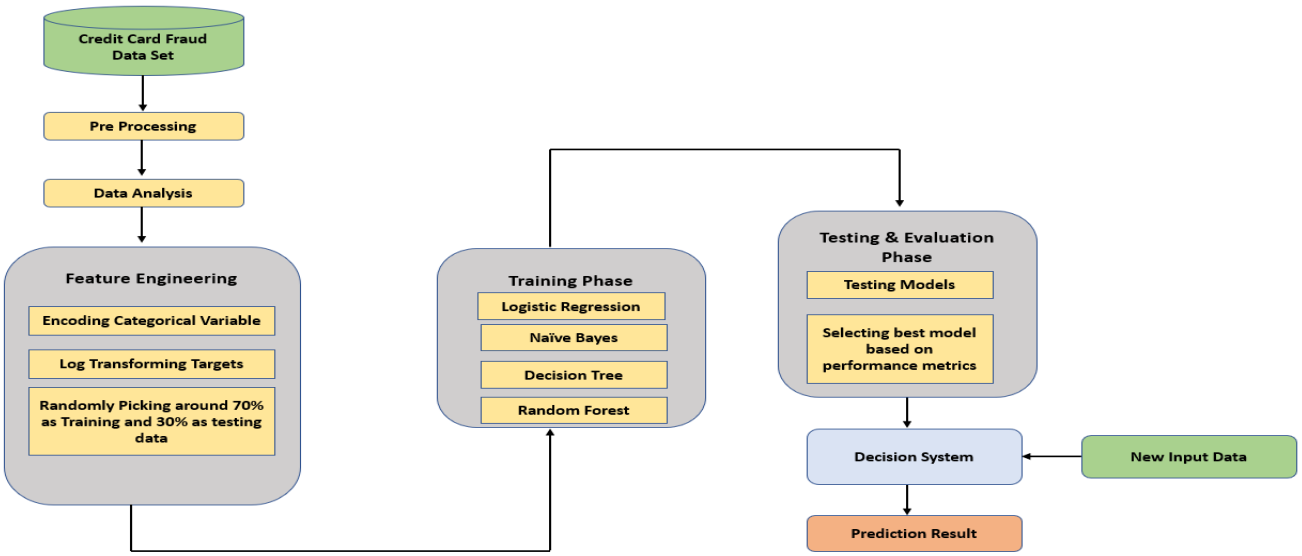


*Fig.1. Architecture*

## 4. MODULES

Data collection is the first stage of the research; the data being gathered consists of a number of actions, some of which are genuine and others of which are fraudulent. The project's data collection phase is the initial stage; this dataset consists of a number of operations, some of which are genuine and others of which are fraudulent. The first stage of the project is data collection; this dataset consists of a number of operations, several of which are genuine and others of which are fraudulent. Credit Card Dataset using the Kaggle website as source, may access a credit card payment information set. This process of the dataset: In this module, selected data is prepared, cleaned up, and sampled. Dataset loading is a variety of library functions can be used to load the dataset. To read CSV function of the python pandas module was used in this case to load a data collection in CSV or Microsoft Excel format. Creating a model for the data that was trained is now used to create the model after the data was divided into test and training samples, each of which was given a 70% and 30% weighting. Determining the module's accuracy using a variety of algorithms, this stage determines the module's correctness. Streamlit Web Framework in the web application, will incorporate the machine algorithm graphs, user input, and accuracy result.

## 5. ALGORITHMS AND FRAMEWORK

### 5.1 Logistic Regression

Logistic regression (LR) is a well-known supervised learning method that is commonly used for classification tasks. It is a statistical technique for studying a collection of data in which any number of independent variables influence the outcome. The purpose of the logistic regression technique is to determine the best model that best describes the connection between the variable that is dependent (or responder) and the variables that are independent (or predictors) [5]. Logistic regression, unlike linear regression, has a binary or categorical response variable. The advantages are: (a) The logistic regression method is a straightforward and basic algorithm. Because it is simple to put into practise and interpret, it is a common solution for many issues related to classification. (b) The method of logistic regression is a quick technique that can easily handle enormous datasets. (c) Logistic regression is a versatile approach that is capable of handling both large and small datasets. The disadvantages are: (a) Logistic regression presupposes that there is a linear connection between the independent factors and the response variable's log-odds. (b) The logistic regression method may not perform effectively if the connection is non-linear. Outliers affect logistic regression because it is sensitive to them. (c) Non-linear correlations between the independent factors and a response variable cannot be handled using logistic regression [6].

### 5.2 Decision Tree

A decision tree is a regression analysis and classification machine learning algorithm. The decision tree paradigm has a tree-like structure, with each internal node indicating an investigation of a characteristic, each branch reflecting the testing result, and symbolising a class identification or a numerical value [6]. The tree can be "learned" by subdividing the source set depending on the outcome of attribute testing. This method is repeated recursively on each derived subset, which is known as recursive partitioning. Because the development of a classifier that uses decision trees requires no domain expertise or parameter setup, it is suitable for exploratory learning and discovery. High-dimensional data can be handled via decision trees. Classifiers have high accuracy in general decision trees. The decision tree inference is a common inductive way of learning classification information [7]. Decision trees categorise instances by moving them through the tree through the root to a leaf node that offers the instance's categorization. Beginning at the lowest point of the tree, an instance is categorised by checking the attribute indicated by this node and then going along the tree branch according to the numerical value of the property.

### 5.3 Naive-Bayes

Naive Bayes is a common classification technique based on Bayes' probability theory. It is a straightforward yet effective algorithm that is commonly used in the classification of text, filtering spam, and recommendation systems. The name "naive" comes from the assumption that each of the characteristics is unrelated to each other. To begin, the algorithm computes the estimated likelihood of each class supplied with a set of parameters. This is accomplished by the use of the Bayes' theorem, which indicates that the likelihood of an expectation (in the current instance, the class) offered by the information (the features) correlates to the likelihood of the data supplied by the hypothesis multiplied by the initial likelihood of the hypothesis [8]. The advantages are: (a) Naive Bayes is a straightforward algorithm that is simple to grasp and apply. It does not necessitate the use of complex iterative algorithms, as many other machine learning techniques do. (b) The naive Bayes principle is a quick approach that can handle big, high-dimensional datasets. (c) To create accurate predictions, Naive Bayes takes only a small amount of training data. (d) Naive Bayes can deal with insignificant features and is unaffected by them [9]. The disadvantages are: (a) Naive Bayes presupposes that the characteristics are independent of one another, which is not necessarily the case in real-world datasets. (b) The naive Bayes technique has limited expressive capacity and may be incapable of capturing complicated feature interactions. (c) Naive Bayes presupposes a predefined distribution of probability for the attributes, which may or may not be appropriate for the dataset. Because it implies a discrete distribution of probability for the features, Naive Bayes is unsuitable for continuous data. (d) Naive Bayes is best suited for categorical information and may struggle with continuum or numerical features [10].

### 5.4 KNN

In the KNN model, statistics uses a non-para-supervised learning method called the k-nearest neighbour technique (k-NN). A class member is the product of the k-NN classification. An object is classified by the unanimous consent of its neighbours, with the object being given the classification that is most common among its k (positive, frequently tiny) nearby objects. When k is equal to 1, the item is simply classified as the object's lone nearest neighbour [11]. With the k-NN classification approach, all processing is put on hold until the function being classified has been evaluated and only a remote model has been constructed. The accuracy of the aforementioned method can be greatly improved by identifying the source data if the features represent a variety of physical measurements or arrive at vastly different scales because the method uses distances for categorization [12]. Applying weights to neighbour contributions to make the near neighbours contribute more to the median than the distant neighbours is an effective method for regression as well as classification. Assigning each neighbour a weight of 1/d, where d represents their distance from one another, is a common way to weigh objects. In both the k-NN classification and the k-NN regression, the neighbours are selected from a group of elements where the class or object value for a property has been established.This is the algorithm's training set; however, no explicit training is necessary [13,14].

### 5.5 Streamlit

For the detection of fraud in credit cards, with the help of the freely available web application framework Streamlit, programmers may use Python to build interactive data-driven apps. With Streamlit, developers can easily create data visualizations, interactive dashboards, and machine learning models that can be deployed as web applications. Streamlit provides a simple and intuitive interface for creating applications, allowing developers to focus on the content and functionality of their applications rather than the technical details of web development. Streamlit provides a number of features to make building web applications easier, including: A simple and intuitive API for creating user interfaces and data visualizations.Automatic reactivity, this enables developers to construct applications that are interactive that are updated in immediate time as the consumer interacts with them. Built-in support for popular data science libraries such as Pandas, Matplotlib, and Plotly. Easy deployment to a variety of cloud platforms, including Heroku and Google Clouds [15]. Overall, Streamlit is a powerful tool for creating interactive data-driven applications with Python, and is well-suited for data scientists and developers who want to quickly prototype and deploy web applications. Streamlit is a versatile web application framework that can be used for a wide variety of applications in data science, machine learning, and beyond. Here are some examples of the uses of Streamlit:

(a) Interactive data exploration: Streamlit makes it easy to create interactive data visualizations and exploration tools, allowing users to explore and analyze data in a more intuitive and engaging way.

(b) Machine learning model development and deployment: Streamlit can be used to develop and deploy machine learning models as web applications, allowing users to interact with and test models in real-time.

(c) Dashboard creation: Streamlit is well-suited for creating interactive dashboards that allow users to explore and analyze data from a variety of sources.
(d) Prototyping and experimentation: Streamlit provides an easy-to-use interface for prototyping and experimenting with new data science ideas and techniques, allowing users to quickly test and iterate on new ideas.
(e) Education and training: Streamlit can be used to create interactive educational tools and tutorials, allowing students and learners [16,17].

# 6. TECHNIQUES

### 6.1 Repeat retailer

The fraud detection of a credit card using a repeat retailer is the technique that utilizes the history of transactions made at a particular retailer to identify potentially fraudulent transactions. The basic idea is that if a cardholder has made several legitimate transactions at a particular retailer in the past, then any future transactions at that retailer are more likely to be legitimate as well. The system maintains a history of transactions made by each cardholder at each retailer. When a new transaction is made, the system checks to see if the cardholder has made any previous transactions at the same retailer. If the cardholder has made previous transactions at the retailer, the system calculates various metrics, such as the average transaction amount, where the time is between transactions, and the location of the transactions [18]. The system compares the metrics of the new transaction to the historical metrics of the cardholder's previous transactions at the retailer. If the metrics of the new transaction are significantly different from the historical metrics, the system flags the transaction as potentially fraudulent and triggers a review process. Repeat retailer is just one of many techniques used in detection of credit card fraud, and is often used in combination with other techniques, such as anomaly detection and machine learning. By leveraging the history of transactions made by each cardholder, repeat retailer can help identify potentially fraudulent transactions and reduce the incidence of credit card fraud. Through graph model which are depicting the analysis part based on the dataset, column of ‘repeat retailer’. Predicting the percent of ‘yes’ is 88.2% and ‘no’ is 11.8%.

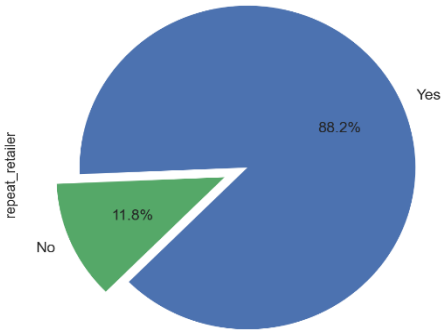


***Fig.2. Pie chart of Retain Retailer***

### 6.2 Used chip

The detection of fraud in credit card using used_chip is a technique that utilizes the information stored on the chip of a credit card to identify potentially fraudulent transactions. The basic idea is that the information stored on the chip can provide additional authentication and validation that can help verify the legitimacy of a transaction.The system reads the information stored on the chip of the credit card, including the card number, expiration date, and other information [19]. The system compares this information to the information provided by the merchant, such as the transaction amount, the merchant name, and the location of the transaction. If the information provided by the merchant matches the information stored on the chip, the system assumes that the transaction is valid and approves the transaction. If the information provided by the merchant does not match the information stored on the chip, the system flags the transaction as potentially fraudulent and triggers a review process. Used_chip is just one of many techniques used in detection of credit card fraud, and is often used in combination with other techniques, such as repeat retailer analysis and machine learning [20]. By utilizing the information stored on the chip of a credit card, used_chip can help verify the legitimacy of a transaction and reduce the incidence of credit card fraud. .Through graph model which are depicting the analysis part based on the dataset, column of 'used chip'. Predicting the percent of 'yes' is 65.0% and 'no' is 35.0%.

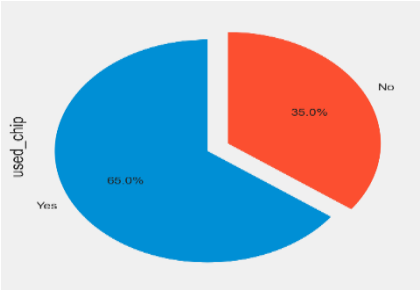


***Fig.3. Pie chart of Used chip***

### 6.3 Used_pin_number

Credit card fraud detection using 'used pin number' is a technique that utilizes the personal identification number (PIN) entered by the cardholder during a transaction to identify potentially fraudulent transactions. The basic idea is that if a transaction is made using a cardholder's stolen credit card, the thief is unlikely to know the correct PIN number, and this can be used to identify potentially fraudulent transactions. The cardholder enters their PIN number during the transaction. The system compares the entered PIN number to the PIN number stored on the credit card's chip. If the entered PIN number matches the stored PIN number, the system assumes that the transaction is legitimate and approves the transaction [21,22]. If the entered PIN number does not match the stored PIN number, the system flags the transaction as potentially fraudulent and triggers a review process. Used_pin_number is just one of many techniques used in detection of

fraud in credit card, and is often used in combination with other techniques, such as anomaly detection and machine learning. By utilizing the PIN number entered by the cardholder, used_pin_number can help verify the legitimacy of a transaction and reduce the incidence of credit card fraud [23,24]. However, it is important to note that this technique is not foolproof and can be compromised in cases where the PIN number has been stolen or the thief has managed to guess the correct PIN number. Through graph model which are depicting the analysis part based on the dataset, column of 'used pin number'. Predicting the percent of 'yes' is 89.9% and 'no' is 10.1%.

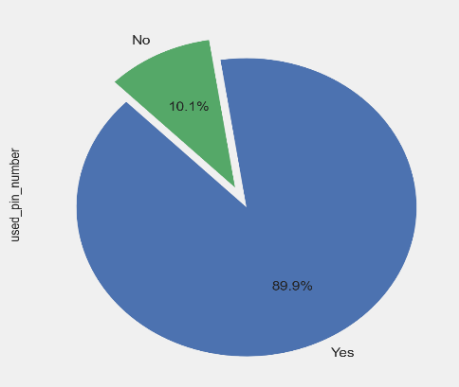


***Fig.4. Pie chart of Used Pin Number***

### 6.4 Online order

The fraud detection of credit card using online_order is a technique that utilizes the information provided during an online order transaction to identify potentially fraudulent transactions. The basic idea is that certain patterns and behaviors can be used to identify potentially fraudulent transactions made online [25]. The system analyzes the information provided during the online order transaction, including the IP address of the device used to make the transaction, the shipping address, and the billing address. The system compares this information to the cardholder's historical information, such as their location, typical shipping and billing addresses, and other relevant information [26]. The system looks for patterns and anomalies in the information provided, such as a shipping address that is significantly different from the cardholder's billing address or an IP address that is located in a different country. If the system detects any suspicious patterns or anomalies, it flags the transaction as potentially fraudulent and triggers a review process [27]. Online_order is just one of many techniques used in detection of fraud, and often used in group with other techniques, where as like in machine learning and anomaly detection. By analyzing the information provided during an online order transaction, online_order can help identify potentially fraudulent transactions and reduce the occurrence of fraud. However, it is important to note that this technique is not foolproof and can be compromised in cases where the thief has access to the cardholder's personal information, such as their shipping and billing addresses [28]. Through graph model which are depicting the analysis part based on the dataset, column of 'online order'.

Predicting the percent of ‘yes’ is 65.1% and ‘no’ is 34.9%.

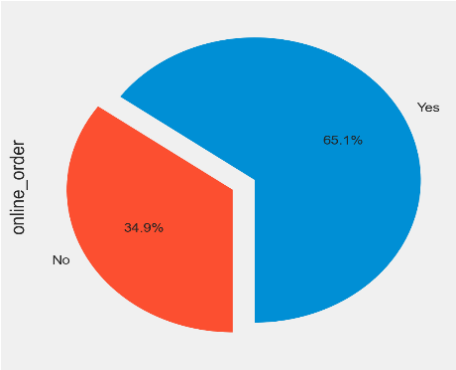


***Fig.5. Pie chart of Online Order***

## 7. RESULT

Based on the precision and accuracy scores offered, it is critical to analyse the fraudulent credit card detection application's particular evaluation standards and goals. If accuracy is important, Naive Bayes has the highest accuracy score. It is worth mentioning, however, that Naive Bayes had the lowest precision score, indicating a higher false-positive rate. Decision Tree obtained the highest precision score if precision is important. This means that it was more accurate in classifying fraudulent transactions. The decision tree, on the other hand, had a somewhat lower accuracy score. When accuracy and precision were combined, logistic regression performed moderately in both measurements, with a good precision score and an accuracy score. Logistic regression produced a fair balance of precision and accuracy. K-Nearest Neighbours (KNN) also performed well, with an accuracy and precision score ofBased on the ratings, logistic regression appears to be the model of greatest choice since it achieves a decent mix of accuracy and precision. However, the best appropriate model is ultimately determined by the application's specific requirements and goals, and other variables like complexity of computation, interpretability, and scalability must also be considered.

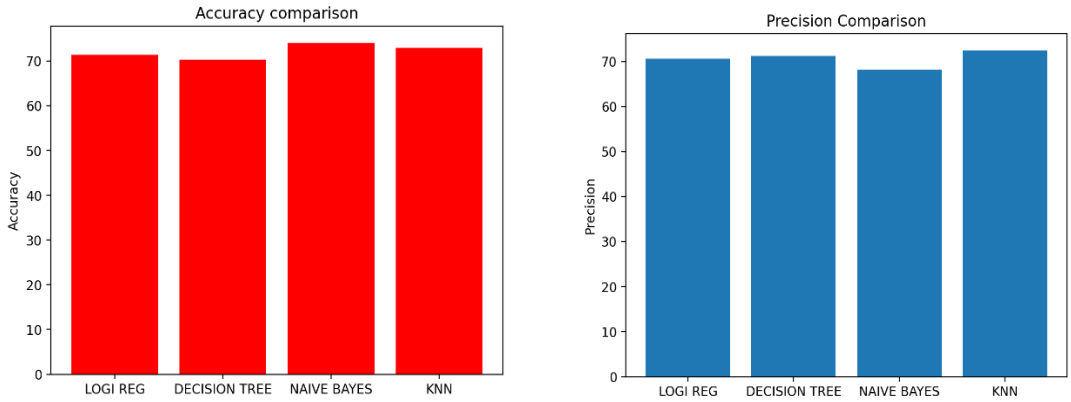


***Fig.6. Comparison graphs of four models based on accuracy and precision***

## 8. CONCLUSION

Finally, credit card theft is a serious worry for both financial institutions and customers. Machine learning algorithms have been shown to be useful in the real-time detection of fraudulent transactions. In this paper, establishing a system for detecting credit card fraud by combining both supervised and unsupervised algorithms to discover patterns that signal fraudulent behaviour. And also combined the learned algorithmic learning model with a web-based structure to create a straightforward user experience for real-time identification of fraud. The experimental results indicated that the suggested framework detected fraudulent transactions with high accuracy while minimising false positives. The proposed methodology can be used by banking organisations in order to enhance their theft identification abilities and avoid financial losses caused by credit card theft. Future studies can concentrate on enhancing the suggested framework's accuracy and investigating the use of alternative machine learning approaches to address this challenge. This credit card fraud detection architecture, which employs streamlining and machine learning, is highly successful in preventing monetary harm caused by credit card fraud. Future research can concentrate on increasing the framework's performance and investigating the use of more sophisticated machine learning methods. The suggested credit card fraud identification framework based on stream-lit and machine learning solves this challenge effectively. To accurately identify fraudulent transactions, the platform includes several machine learning algorithms such as decision tree, XGBoost, random forest and logistic regression.

## 9. REFERENCES

[1] Raj, S. Benson Edwin, and A. Annie Portia. "Analysis on credit card fraud detection methods." In *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)*, pp. 152-156. IEEE, 2011.

[2] Ghosh, Sushmito, and Douglas L. Reilly. "Credit card fraud detection with a neural-network." In *System Sciences, 1994. Proceedings of the Twenty-Seventh Hawaii International Conference on*, vol. 3, pp. 621-630. IEEE, 1994.

[3] Chaudhary, Khyati, Jyoti Yadav, and Bhawna Mallick. "A review of fraud detection techniques: Credit card." *International Journal of Computer Applications* 45, no. 1 (2012): 39-44.

[4] Srivastava, Abhinav, Amlan Kundu, Shamik Sural, and Arun Majumdar. "Credit card fraud detection using hidden Markov model." *IEEE Transactions on dependable and secure computing* 5, no. 1 (2008): 37-48.

[5] Awoyemi, John O., Adebayo O. Adetunmbi, and Samuel A. Oluwadare. "Credit card fraud detection using machine learning techniques: A comparative analysis." In *2017 international conference on computing networking and informatics (ICCNI)*, pp. 1-9. IEEE, 2017.

[6] Sahin, Yusuf, and Ekrem Duman. "Detecting credit card fraud by ANN and logistic regression." In *2011 international symposium on innovations in intelligent systems and applications*, pp. 315-319. IEEE, 2011.

[7] Kiran, Sai, Jyoti Guru, Rishabh Kumar, Naveen Kumar, Deepak Katariya, and Maheshwar Sharma. "Credit card fraud detection using Naïve Bayes model based and KNN classifier." *International Journal of Advance Research, Ideas and Innovations in Technoloy* 4, no. 3 (2018): 44.

[8] Husejinovic, Admel. "Credit card fraud detection using naive Bayesian and c4. 5 decision tree classifiers." *Husejinovic, A.(2020). Credit card fraud detection using naive Bayesian and C* 4 (2020): 1-5.

[9] Saheed, Yakub K., Moshood A. Hambali, Micheal O. Arowolo, and Yinusa A. Olasupo. "Application of GA feature selection on Naive Bayes, random forest and SVM for credit card fraud detection." In *2020 international conference on decision aid sciences and application (DASA)*, pp. 1091-1097. IEEE, 2020.

[10] Varmedja, Dejan, Mirjana Karanovic, Srdjan Sladojevic, Marko Arsenovic, and Andras Anderla. "Credit card fraud detection-machine learning methods." In *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH)*, pp. 1-5. IEEE, 2019.

[11] Yee, Ong Shu, Saravanan Sagadevan, and Nurul Hashimah Ahamed Hassain Malim. "Credit card fraud detection using machine learning as data mining technique." *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)* 10, no. 1-4 (2018): 23-27.

[12] Malini, N., and M. Pushpa. "Analysis on credit card fraud identification techniques based on KNN and outlier detection." In *2017 third international conference on advances in electrical, electronics, information, communication and bio-informatics (AEEICB)*, pp. 255-258. IEEE, 2017.

[13] Ganji, Venkata Ratnam, and Siva Naga Prasad Mannem. "Credit card fraud detection using anti-k nearest neighbor algorithm." *International Journal on Computer Science and Engineering* 4, no. 6 (2012): 1035-1039.

[14] Vengatesan, K., A. Kumar, S. Yuvraj, V. Kumar, and S. Sabnis. "Credit card fraud detection using data analytic techniques." *Advances in Mathematics: Scientific Journal* 9, no. 3 (2020): 1185-1196.

[15] Zareapoor, Masoumeh, K. R. Seeja, and M. Afshar Alam. "Analysis on credit card fraud detection techniques: based on certain design criteria." *International journal of computer applications* 52, no. 3 (2012).

[16] Nancy, A. Maria, G. Senthil Kumar, S. Veena, NA S. Vinoth, and Moinak Bandyopadhyay. "Fraud detection in credit card transaction using hybrid model." In *AIP Conference Proceedings*, vol. 2277, no. 1, p. 130010. AIP Publishing LLC, 2020.

[17] Kaur, Darshan. "Machine Learning Approach for Credit Card Fraud Detection (KNN & Naïve Bayes)." In *Machine Learning Approach for Credit Card Fraud Detection (KNN & Naïve Bayes)(March 30, 2020). Proceedings of the International Conference on Innovative Computing & Communications (ICICC)*. 2020.

[18] Saheed, Yakub Kayode, Usman Ahmad Baba, and Mustafa Ayobami Raji. "Big Data Analytics for Credit Card Fraud Detection Using Supervised Machine Learning Models." In *Big Data Analytics in the Insurance Market*, pp. 31-56. Emerald Publishing Limited, 2022.

[19] Adewumi, Aderemi O., and Andronicus A. Akinyelu. "A survey of machine-learning and nature-inspired based credit card fraud detection techniques." *International Journal of System Assurance Engineering and Management* 8 (2017): 937-953.

[20] Mehbodniya, Abolfazl, Izhar Alam, Sagar Pande, Rahul Neware, Kantilal Pitambar Rane, Mohammad Shabaz, and Mangena Venu Madhavan. "Financial fraud detection in healthcare using machine learning and deep learning techniques." *Security and Communication Networks* 2021 (2021): 1-8.

[21] Handa, Akansha, Yash Dhawan, and Prabhat Semwal. "Hybrid analysis on credit card fraud detection using machine learning techniques." *Handbook of Big Data Analytics and Forensics* (2022): 223-238.

[22] Tiwari, Pooja, Simran Mehta, Nishtha Sakhuja, Ishu Gupta, and Ashutosh Kumar Singh. "Hybrid method in identifying the fraud detection in the credit card." In *Evolutionary Computing and Mobile Sustainable Networks: Proceedings of ICECMSN 2020*, pp. 27-35. Springer Singapore, 2021.

[23] Kazemi, Zahra, and Houman Zarrabi. "Using deep networks for fraud detection in the credit card transactions." In *2017 IEEE 4th International conference on knowledge-based engineering and innovation (KBEI)*, pp. 0630-0633. IEEE, 2017.

[24] Faraji, Zahra. "A Review of Machine Learning Applications for Credit Card Fraud Detection with A Case study." *SEISENSE Journal of Management* 5, no. 1 (2022): 49-59.

[25] Prusti, Debachudamani, and Santanu Kumar Rath. "Web service based credit card fraud detection by applying machine learning techniques." In *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*, pp. 492-497. IEEE, 2019.

[26] Ahammad, Jalal, Nazia Hossain, and Mohammad Shafiul Alam. "Credit card fraud detection using data pre-processing on imbalanced data-Both oversampling and undersampling." In Proceedings of the International Conference on Computing Advancements, pp. 1-4. 2020.

[27] Ata, Oğuz, and Layth Hazim. "Comparative analysis of different distributions dataset by using data mining techniques on credit card fraud detection." Tehnički vjesnik 27, no. 2 (2020): 618-626.

[28] Shirgave, Suresh, Chetan Awati, Rashmi More, and Sonam Patil. "A review on credit card fraud detection using machine learning." International Journal of Scientific & technology research 8, no. 10 (2019): 1217-1220.