WCRG Global Research WORLD COMET RESEARCH GROUP

International Journal of Commerce, Finance and Digital Economy
Vol. 4, No. 1, 2021
Doi: 10.67228/3071642X/IJCFDE-2021PI3C9D
PP. 01-17

Original Article

Fraud Detection in Online Banking Using Machine Learning

Noah Wright[1], Isabella Moore[2]

[1]*Data Engineering Lead, SAP, Germany.*

[2]*HR Director, Siemens, Germany.*

Received: 03-04-2021 Revised: 03-24-2021 Accepted: 03-29-2021 Published: 04-02-2021

ABSTRACT

Machine Learning-based fraud detection systems provide an effective and intelligent solution for securing online banking platforms against evolving cyber threats. By analyzing transaction patterns and detecting anomalies in real time, ML algorithms such as Logistic Regression, Decision Tree, Random Forest, and XGBoost can accurately identify fraudulent activities while reducing false alarms. Experimental results indicate that ensemble models, particularly Random Forest, achieve superior detection performance. The proposed framework enhances fraud prevention, minimizes financial losses, improves customer trust, and offers a scalable and adaptive approach for modern digital banking security.

KEYWORDS

Online Banking, Fraud Detection, Machine Learning, Financial Security, Random Forest, Transaction Analysis, Cybersecurity, Data Mining.

ISSN: 3071-642X
PP: 01-17.

1. INTRODUCTION

1.1. Background of Online Banking Fraud

Online banking has revolutionized banking services delivery by enabling customers to access and operate their banking accounts from anywhere and anytime using any Internet-enabled device. Users experience much greater convenience, speed and ease of access with services like transferring funds, paying bills, managing their accounts, and shopping online. The digital banking story has been accelerating across the globe, and the amount of financial transactions online has risen significantly. But the same speedy digital evolution has also allured the cybercriminals who take advantage of technology weaknesses to increase their profit. Online banking fraud consists of fund transfer fraud, phishing, identity theft, account takeover, credential theft and transaction manipulation. These scams can lead to substantial monetary losses for customers and financial institutions, as well as erode confidence in electronic monetary services. Fraudulent schemes are getting more sophisticated, and it is a challenge to detect them, necessitating intelligent monitoring and advanced security measures. Hence, the need to develop effective fraud detection mechanisms has become a critical need to ensure the security of transactions, protect the assets of customers and maintain trust in online banking services.

1.2. Need for Machine Learning-Based Fraud Detection

The rise of online banking and digital payment methods has created more opportunities for financial fraud than ever before. Traditional fraud detection methods based on fixed rules and manual monitoring are often unable to detect sophisticated and evolving fraud patterns. Machine learning provides an intelligent and automated solution that can sift through huge amounts of transaction data and spot suspicious transactions in real time. The following are some of the reasons why machine learning can be so important in the context of fraud detection systems.

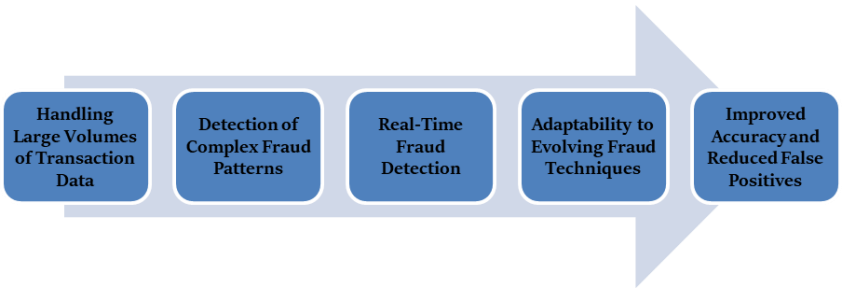

Fig 1 - Need for Machine Learning-Based Fraud Detection

1.2.1. Handling Large Volumes of Transaction Data

In today's banking landscape, millions of transactions can take place daily, which makes it impractical to manually monitor and analyze them with rules. Machine learning algorithms are capable of processing a tremendous amount of data in a short time to detect suspicious patterns that could be a red flag for fraud. It allows financial institutions to continue to have a strong fraud monitoring system as transactions grow.

ISSN: 3071-642X
PP: 01-17.

1.2.2. Detection of Complex Fraud Patterns

Fraudsters are continually finding ways around conventional security. With historical transaction data, machine learning models can capture complex relationships and patterns that may not be obvious to traditional methods. Machine learning systems can detect even minute variations in behavior and suspicious transaction patterns, enhancing the accuracy of fraud detection and the chances of thwarting attacks.

1.2.3. Real-Time Fraud Detection

In an online banking context, fraudulent transactions have to be identified and prevented before any financial harm comes to it. Machine Learning models can analyze transactions as they happen and classify them as legitimate or suspicious transactions at the moment of the transaction. The speed at which this decisions can be made can help banks stop potentially fraudulent transactions, send alerts and prevent fraud before it happens, and help keep money more secure.

1.2.4. Adaptability to Evolving Fraud Techniques

Legacy fraud detection systems have to be manually updated on a regular basis, as fraudsters develop new methods. Machine learning models, on the other hand, can ingest new transaction information to continuously update their model. This flexibility allows financial institutions to effectively address emerging threats and ensure robust protection against evolving cybercrime operations.

1.2.5. Improved Accuracy and Reduced False Positives

The major disadvantage of traditional fraud detection systems is that they tend to generate numerous false alarms because of too strict of a set of rules. Machine learning algorithms enable data-driven decision making and enhance the accuracy of the classification while differentiating between the legitimate and fraudulent transactions. Consequently, they help mitigate false alarms, limit follow-up investigations, improve the customer experience, and improve fraud detection efficiencies.

1.3. Challenges in Online Banking Fraud Detection

Given the growing volume of digital transactions and the sophistication of today's cyber threats, online banking fraud detection is an evolving and complex challenge. A major problem is the highly skewed distribution of transactions, with a tiny fraction of the transactions being fraudulent. This imbalance makes it hard to train machine learning models effectively to recognize fraud patterns, since they will tend to learn the majority class of non-fraudulent transactions. Dynamic fraud patterns are a major challenge as well. As time goes on, cybercriminals are continually changing the ways they attack and how they do it, rendering successful detection strategies less effective over time. This means that fraud detection systems need to be flexible enough to adapt to evolving fraud practices and new threats. The need for real-time detection adds to the fraud detection process. Financial transactions need to be analyzed on the spot and accurate decisions made within seconds, to avoid possible financial losses due to unauthorised financial transactions. If detection is delayed, fraudulent transactions may be executed without any problems, thereby seriously harming

ISSN: 3071-642X
PP: 01-17.

the customers and banks. Also, with today's banking systems, they process high volumes of transactions and sometimes millions of transactions per day. Handling such a vast amount of information in a timely and efficient manner with high detection accuracy demands powerful computational power and scalable analytical models. Minimization of false positives is another major challenge, meaning that the actual and false positive transactions are classified as fraud. Many false alerts may bother customers, cause delays in the banking process, and cause more work on fraud investigators. An effective fraud detection system, therefore, should strike a balance between the ability to detect real fraud cases and minimise unnecessary alerts. The key to meeting these challenges is to leverage advanced machine learning techniques, intelligent data analysis, and adaptive detection mechanisms that can deliver accurate, scalable, and real-time fraud prevention in today's online banking landscape.

2. LITERATURE SURVEY

2.1. Traditional Fraud Detection Techniques

The traditional fraud detection systems were mostly based on rule based and threshold based methods, which were manually designed by domain experts. The systems were based on known fraud patterns, blacklists of known fraud accounts, geographic restrictions, frequency limits, and/or transaction amount limits. The primary benefit of these methods was that they were simple, required minimal computation and were easily implemented. But they did not have the capacity to evolve with fraud tactics, and were less effective with complex and newer fraud schemes. The fraudsters were constantly changing their methods making it difficult to keep up with the changing rules and resulting in a lot of false positives for rule based systems.

2.1.1. Z-Score Analysis

Z-score analysis is a statistical method that breaks down the transaction data to determine if there are any anomalies in the data by calculating how far the transaction is from the average behavior of the customer or data set. Operational transactions involving quantities that are significantly larger or smaller than the average transaction are identified as suspicious. It's a simple approach which is quite efficient to compute and implement and has been used in simple fraud detection tasks. It does assume that data is normally distributed, however, and may not always be able to identify more complicated fraud schemes in a vast data set.

2.1.2. Bayesian Models

Bayesian models are based on probability theory and attempt to determine the probability of a transaction being fraudulent using existing evidence and knowledge. These models continuously will be updated with newer transaction data in order to adaptively make decisions based on fraud probabilities. Bayesian techniques can be very useful in dealing with uncertainty and incomplete information. However, they are very dependent on the quality of their a priori probabilities and assumptions and can become less effective in the face of rapidly changing fraud behaviors.

ISSN: 3071-642X
PP: 01-17.

2.1.3. Regression Analysis

Multiple features of transactions have been widely used in regression analysis (usually logistic regression) to predict the probability of fraud. It looks at how much data variables like transaction size and number are associated with the probability of fraudulent transactions. Regression models are simple and easy to understand, and they are also efficient in terms of computation, which is a welcome attribute in financial institutions. They tend to fall short of dealing with complex, non-linear relationships found in today's transactional data, which diminishes their ability to detect more complex fraud schemes.

2.2. Machine Learning Approaches for Fraud Detection

Machine learning has greatly enhanced fraud detection by allowing systems to learn patterns directly from the historical data of transactions. Machine learning models discover hidden relationships and adjust to new fraud behaviors on their own, whereas traditional rule-based systems require manual intervention, collaboration, and implementation. Unlike traditional rule-based methods, machine learning models can automatically detect hidden relationships and adapt to new fraud behaviors without needing manual intervention, collaboration or implementation. Often, supervised learning algorithms are employed, which are trained with labeled data of both legitimate and fraudulent transactions. These are more accurate for detection, scalable and less reliance on manually written rules. Therefore, machine learning is one of the most popular technologies employed in today's fraud detection solutions.

2.2.1. Logistic Regression

Logistic Regression is a supervised learning algorithm which is often used in binary classification problems, like fraud detection. It attempts to calculate the probability for the transactions to be in the legitimate or fraudulent class with a logistic function. It is a simple, fast and highly interpretable model that can be used to understand the impact of various features on the prediction of fraud. It works well for linear relationships, but may not perform as well in more complex and non-linear transaction patterns.

2.2.2. Decision Trees

Decision Trees classify transactions according to the values of features: they recursively split the data according to the value of the feature and build a tree-like structure of decisions. Model is easy to understand and easy to visualize and thus are suitable for fraud detection applications where explainability is a requirement. Decision Trees can model non-linear relationships and interactions between variables. They, however, tend to overfit when trained on noisy or imbalanced data sets, and their ability to generalize may be affected.

2.2.3. Random Forest

Random Forest is an ensemble learning technique that uses the decision tree technique for making predictions by combining multiple decision trees to get better predictions and reduce the effect of noise. They are trained on a random subset of data and features and the final prediction

ISSN: 3071-642X
PP: 01-17.

made by majority voting. The process minimises overfitting and enhances the model for fraud detection. Random Forest is a very popular method that is widely used for financial data with many dimensions and is very efficient, scalable, and performs well.

2.2.4. Support Vector Machines (SVM):

Support Vector Machines are classification algorithms that can determine the best separating line between fraudulent and legitimate transactions, called a hyperplane. In high-dimensional data, SVMs are effective, and can be used to model complex relationships with kernel functions. They frequently perform well in terms of classification accuracy particularly when the classes are well separated. But, SVM training is costly when dealing with large-scale transaction data sets, which makes it less useful in real-time fraud detection applications.

2.2.5. XGBoost

XGBoost (Extreme Gradient Boosting) is one of the most recent ensemble learning methods that sequentially increase the number of trees in the ensemble and correct the errors of the previous trees. It uses regularization methods to avoid overfitting and enhance model generalisation. XGBoost has gained popularity in fraud detection due to its high accuracy, fast execution speed, and ability to handle missing values and imbalanced datasets. It is always among the top performing algorithms in fraud detection research and industry practice.

2.3. Deep Learning and Hybrid Models

The fact that deep learning can learn complex and hierarchical patterns, automatically from large volumes of transaction data, makes it a useful tool for fraud detection. Unlike traditional machine learning methods that rely heavily on manual feature engineering, deep learning models can extract meaningful representations directly from raw data. These methods can be especially helpful in spotting advanced fraud schemes that change over time. Moreover, hybrid models based on the combination of machine learning and deep learning techniques are more beneficial, as they benefit from the best of both worlds: superior detection accuracy, robustness and adaptability.

2.3.1. Artificial Neural Networks (ANN)

Artificial Neural Networks are models of computation that are based on the structure and operation of the human brain. ANNs are made up of neurons that are interconnected into layers that take in data and learn patterns when they are trained. The applications of ANNs in fraud detection are useful in the ability of the network to find complex correlations between the parameters of transactions, which are not revealed by conventional means. They are great at learning non-linear patterns which helps them detect fraudulent activities. They are, however, typically demand significant data sets and processing power in order to work well.

2.3.2. Convolutional Neural Networks (CNN)

A Convolutional Neural Network (CNN) is a type of deep learning model that is primarily used for image processing but also has widespread applications for fraud detection. CNNs are comprised of convolutional layers which automatically identify features of interest from structured and

ISSN: 3071-642X
PP: 01-17.

unstructured data. For financial systems, CNNs can detect subtle transaction patterns and irregularities that are signs of financial fraud. They automatically extract features, which decreases the requirement for manual feature extraction but may be complicated and time-consuming to train.

2.3.3. Long short-term memory networks (LSTM):

A special class of recurrent neural networks, called Long Short-Term Memory Networks are designed to model temporal dependencies in sequential data. Because financial transactions are time-based, LSTMs can be especially useful in analyzing the history of transactions and detecting any unusual behaviors. They are able to recall long-term patterns and identify minor shifts in customer use that could indicate fraud. This makes LSTMs very well suited for real-time fraud detection systems that are based on sequential transaction analysis.

2.3.4. Hybrid Models

To get better results, hybrid models of fraud detection are used in which combinations of algorithms are applied, such as machine learning classifiers and deep learning architectures. Such models are able to benefit from the advantages of various methods, including the interpretability of traditional machine learning and the pattern recognition ability of deep learning. Utilizing a variety of methodologies, hybrid models can help minimize false positives, increase accuracy of fraud detection and improve system reliability. Thus, they are a great research opportunity in contemporary financial fraud detection.

2.4. Research Gap and Motivation

While significant advances have been achieved in the field of fraud detection through machine learning and deep learning approaches, there are still a number of challenges that have yet to be addressed. The first problem is data imbalance; fraudulent transactions make up a small percentage of the data set, so models are more likely to be biased in favour of genuine transactions. Furthermore, sophisticated algorithms can be resource-intensive, which can pose challenges for deployment in high-volume financial settings. Another key challenge is model interpretability, with complex models like deep neural networks being "black boxes" that are hard for financial institutions to interpret and understand the decision-making process behind. Inspired by these drawbacks, the current research project seeks to establish an effective and scalable machine learning methodology that can deliver high detection accuracy, tackle the challenge of class imbalance, minimize computational load, and enable real-world deployment in the fraud detection systems.

3. METHODOLOGY

3.1. Proposed System Architecture

The suggested fraud detection system intends to precisely recognize fraudulent transactions by means of an organized procedure of processing steps. The system starts with data collection of transactions and then proceeds with data pre-processing and feature engineering before analysis. The processed data set is then used for training machine learning models, and the trained model is used

ISSN: 3071-642X
PP: 01-17.

for the classification of transactions as legitimate or fraudulent. This architecture allows for a high detection accuracy, has a high scalability and is efficient with large datasets.

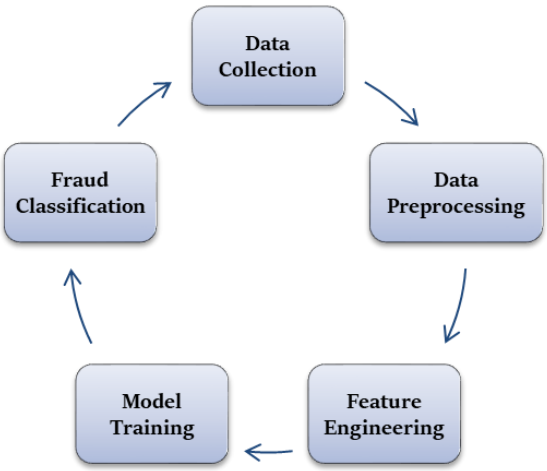

Fig 2 - Proposed System Architecture

3.1.1. Data Collection

The initial phase of the proposed system involves the collection of data, which includes historical records of transactions from trusted sources like banking databases, payment gateways, or fraud detection data sets made available publicly. The data collected usually contains transaction information, like amount, time of the transaction, customer details, merchant details, and fraud labels. The quality of the data is crucial for creating a robust fraud detection model, as the effectiveness of machine learning algorithms heavily relies on the quality and quantity of available data.

3.1.2. Data Preprocessing

Data preprocessing entails cleaning and processing data to be used by machine learning algorithms. This phase involves dealing with missing data, cleaning up duplicate records, inconsistencies, and outliers. Also, categorical variables can be converted to numeric, and numeric features can be normalized and/or standardized to be equal-sized. Good preprocessing ensures good quality data and allows for the model to learn more accurately the meaningful patterns.

3.1.3. Feature Engineering

The job of feature engineering is to make and choose the most relevant features that will facilitate fraud detection. During this phase, new variables can be created based on the attributes of transactions, such as transaction frequency, average spending patterns, or transaction intervals. Feature selection Methods are also used to eliminate redundant or irrelevant features and to find out the most informative features. Correct feature engineering increases the capability of the model to predict and thus the performance of the overall detection.

ISSN: 3071-642X
PP: 01-17.

3.1.4. Model Training

Model Training: This step entails training a machine learning model with the processed data to learn how to identify genuine and fraudulent transactions. Labelled transaction data can be used to train various algorithms, including Logistic Regression, Decision Trees, Random Forest, Support Vector Machines, and XGBoost. In this phase, the model identifies patterns and relationships within the data, tunes its parameters to reduce prediction errors. Proper training is crucial for achieving high accuracy and reliable fraud detection performance.

3.1.5. Fraud Classification

The last step of the proposed system is fraud classification, in which the trained model type is applied to classify the new transactions as fraudulent or not. Each transaction received is passed through and processed by the features of the transaction and sent to the trained model which assigns a class to the transaction. Potentially fraudulent transactions can be flagged up for additional investigation or blocked automatically. It helps organizations detect fraud in real-time and minimizes financial losses for financial institutions while making transactions secure.

3.2. Data Preprocessing

One of the key stages in the proposed fraud detection framework is data preprocessing, which is essential for ensuring that the data fed into the machine learning models is of high quality and relevant. Raw transaction data can include missing values, duplicate records, inconsistent data and outliers that can impact the accuracy of the model. Therefore, to bring the data into a structured and reliable format for analysis, preprocessing is carried out to clean and transform the data. A key pre-processing step is the handling of missing values, which involves the identification of missing information and the application of suitable techniques for dealing with it, such as deletion, mean substitution, median replacement or predictive imputation. The correct handling of missing data prevents the loss of information and helps to avoid bias in the results of machine learning algorithms when using the dataset. One of the key pre-processing tasks is outlier detection. An outlier is a transaction that is markedly different from the rest, either through data entry errors, system malfunctions or unusual customer activity. Some outliers might be valid fraud cases, but the statistical values which are computed from extreme values can be misleading and the performance of the models can be affected. Outliers can be detected and treated using various statistical methods, including the Z-score method, Interquartile Range (IQR), and statistical threshold analysis. Data preprocessing is also performed, such as data normalization which is the process of scaling the numerical attributes in the data to a common range. The different value ranges of transaction features (e.g., transaction amount, account balance, transaction frequency) mean that no particular feature becomes so prominent that it dominates the learning process. Various techniques are used for enhancing model convergence and maximizing prediction accuracy, including Min-Max Scaling and Standardization. Also, duplicate elimination is carried out to eliminate duplicate transaction records that might occur in the data gathering or database incorporation system. Duplicated entries may contain bias, contribute to the computational complexity and result in wrongly trained models. The dataset is now more uniform, and more representative of true transaction behavior, due to the

ISSN: 3071-642X
PP: 01-17.

elimination of duplicate records. In conclusion, data preprocessing plays a vital role in improving data quality, minimizing noise, and preparing the data for optimal feature extraction and machine learning model construction, which leads to better and more reliable fraud detection.

3.3. Machine Learning Algorithms

The proposed fraud detection system uses a variety of machine learning algorithms to detect fraud accurately and efficiently. The algorithms are trained on the patterns and transactions from the past and then determine whether or not new transactions are fraudulent or not. Different algorithms excel at different aspects of complex financial data, resulting in greater accuracy and reliability in the detection process. These algorithms are commonly used in fraud detection-related research and practical applications and have been selected: Logistic Regression, Decision Tree, Random Forest, and XGBoost.

3.3.1. Logistic Regression

Logistic Regression is a supervised machine learning model widely used in binary classification tasks, like fraud detection. It models the relationship between the input features and the class label, predicting the chance a transaction is in the fraudulent class. The algorithm applies a logistic (or sigmoid) function to the output to produce a probability value between 0 and 1. This probability allows for identification of legitimate or fraudulent transactions. Logistic Regression is easy to understand, very efficient and very readable; this enables an analyst to understand how each of the transaction attributes contributes to the prediction of a fraudulent transaction. It is an easy to implement and quick to process model and is a good baseline in fraud detection systems.

3.3.2. Decision Tree

The Decision Tree is a classification algorithm that classifies data with a hierarchical tree structure, using a number of decision rules. The model continues to divide the data again according to the most informative features and branches to final classification results. Certain attributes, including transaction amount, frequency, location and account activity, can help identify suspicious transactions in fraud detection. Decision Trees can be easily understood by analysts and the decision making process can be visualized, making them easy to interpret. They are able to learn non-linearities in transaction data, but can become very complex and prone to overfitting when trained on large numbers of transactions and without proper pruning.

3.3.3. Random Forest

Random Forest is an ensembled learning method on which multiple Decision Trees are used to increase the accuracy of the prediction and decrease overfitting. The algorithm produces many trees based on random samples of the data and features, rather than a single tree. The final classification is a majority voting, with each tree making its own prediction whether a transaction is fraudulent or not. This is a collective decision making process for which there is a greater robustness and accuracy than the individual Decision Trees. Random Forest is a powerful tool for fraud

ISSN: 3071-642X
PP: 01-17.

detection due to its ability to process high-dimensional data, to deal with large data sets and to detect complex patterns in transaction behavior, and to possess a good generalization performance.

3.3.4. XGBoost

XGBoost (Extreme Gradient Boosting) is a highly sophisticated ensemble learning method that sequentially constructs decision trees and tries to rectify the mistakes of the trees already created. This enhancement allows the model to capture complex fraud patterns with high accuracy. The XGBoost algorithm is one of the most powerful classification algorithms, thanks to the regularization methods it incorporates into the model to prevent overfitting and to ensure that the model is able to generalize from the training data. It can effectively process missing data, highly skewed datasets and large volume of transaction data, which are typical problems in fraud detection. Given its predictive power, speed of execution and scalability, XGBoost is often employed in the latest financial fraud systems and has been shown to perform better than the traditional machine learning approaches.

3.4. Performance Evaluation Metrics

The metrics used to evaluate the performance of machine-learning models for fraud detection play a critical role in assessing how well they perform in fraud detection. The problem of detecting fraud is a classification problem and thus, multiple evaluation metrics are used for the evaluation of the model from different angles of approach. These metrics can be used to assess the accuracy of the model in detecting fraud, while also reducing false alerts. Most widely used evaluation criteria are: Accuracy, Precision, Recall, and F1-Score. Together, these measures provide a comprehensive understanding of the model's classification capability and reliability in real-world financial environments.

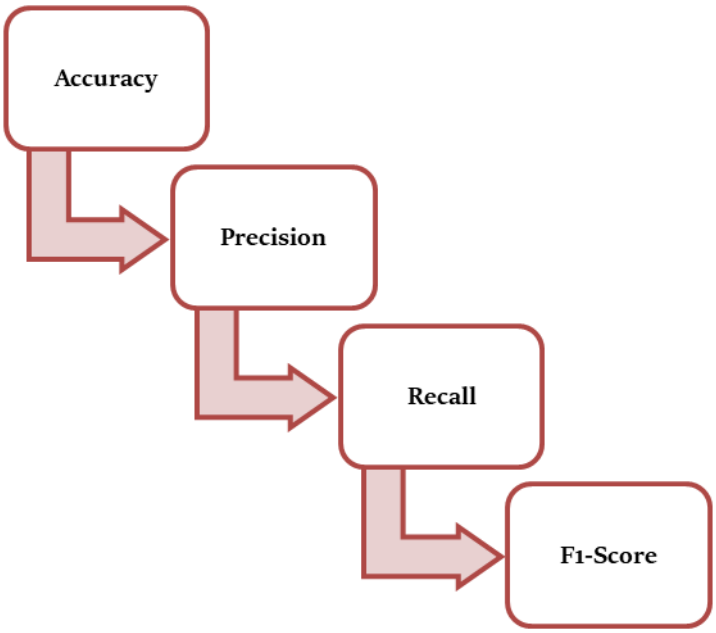

Fig 3 - Performance Evaluation Metrics

ISSN: 3071-642X
PP: 01-17.

3.4.1. Accuracy

The accuracy is a measure of the overall correctness of the fraud detection model and it is calculated as the percentage of correctly classified transactions divided by the number of transactions. It takes into account fraudulent and legal transactions that are accurately forecasted. In the simplest of terms, accuracy is the ratio of transactions that have been correctly identified as fraudulent (True Positives) plus those that have been correctly identified as legitimate (True Negatives) to the total number of transactions. The higher the accuracy, the more the model performs well. In other fraud detection datasets, however, fraudulent transactions are not a large part of the population, and accuracy alone may not be a good indicator of how effective the model is, since a model with high accuracy could simply guess most of the transactions are safe.

3.4.2. Precision

Precision is the ratio of correctly identified frauds to the number of transactions identified as frauds. It tests the Model's capacity to prevent false fraud alerts. Precision is the ratio of transactions correctly identified as fraudulent (True Positives) to the number of transactions predicted to be fraudulent (True Positives + False Positives). The higher the precision, the more likely the model is to be right when it says that a transaction is fraudulent. This ratio is significant in financial systems since too many false alarms will cause inconvenience to the customer and cost the financial company more in investigating the issues.

3.4.3. Recall

Recall or sensitivity/detection rate is a measure of how effective the model is at detecting real fraudulent transactions. It is determined by comparing the number of fraud cases identified correctly (True Positives) and dividing it by the number of actual fraud cases (True Positives + False Negatives). A very high recall value means that the model is able to capture most fraudulent activity and has few missed fraud cases. Recall is also an important metric in fraud detection systems as it could lead to financial loss if a fraudulent transaction is not detected.

3.4.4. F1-Score

The F1-Score is a balanced evaluation measure which takes both precision and recall into account. It is defined as the harmonic mean of precision and recall, giving it an overall evaluation of the model's performance at identifying fraud, while minimizing false detections. The F1-Score is particularly helpful in the case of an imbalanced data set, where one class is much larger than the other. High F1-Score means that the model performs well in terms of both accuracy and minimal false alarms. Hence, it's broadly applied as a key performance indicator for fraud detection research and implementations.

4. RESULTS AND DISCUSSION

4.1. Experimental Setup

The experimental setup aimed to assess the effectiveness and reliability of the proposed machine learning-based fraud detection system in realistic scenarios. The experiments were

ISSN: 3071-642X
PP: 01-17.

conducted with a dataset of banking transactions which included both normal and fraudulent transactions. The dataset included multiple attributes of the transactions, including the transaction amount, time, customer behavior indicators, and whether the transaction was fraudulent or not (the class label). The dataset was preprocessed before model development, involving data cleaning, missing values imputation, data deduplication, and normalization of numerical data to enhance data consistency and the model performance. Also, feature engineering techniques were used to derive relevant information and to boost the prediction of the machine learning algorithms. The data was split into two equal parts of 80% and 20% and used for training and testing the models. The data was split into 80% training data and 20% testing and performance evaluation data. The training dataset was utilized to allow the machine learning algorithms to learn patterns and relationships among the attributes of transactions and the fraud labels. In this stage, the models tweaked their internal settings to make them correctly identify authentic and fraudulent transactions. Finally, the test set (unseen during training) was used to evaluate the models' ability to generalize and to benchmark their performance on test transactions. Different machine learning algorithms such as Logistic Regression, Decision Tree, Random Forest and XGBoost were implemented and compared in the same experiment setup. The accuracy, precision, Recall and F1-Score are all the usual classification parameters used for the assessment. These metrics were able to give a full picture of how well each model could detect fraud and reduce false alarms and missed frauds. Experiments were carried out on a machine-learning environment of Python, along with commonly used libraries like Pandas, NumPy, Scikit-learn, and XGBoost. The experimental configuration allowed a fair comparison of the different algorithms and was found to offer a reliable based from which to deduce the effectiveness of the fraud detection framework proposed.

4.2. Performance Comparison

Four machine learning algorithms (Logistic Regression, Decision Tree, Random Forest, XGBoost) were used to test the performance of the proposed fraud detection system. Accuracy, Precision, Recall and F1-Score were used as performance metrics for comparison of the models, which are commonly used to evaluate classification performance. The results of the experiments show that ensemble learning methods (Random Forest and XGBoost) performed better than single classifiers. A comparison of two algorithms is used to show the advantages and disadvantages of each one in the detection of fraudulent transactions while reducing classification errors to a minimum.

Table 1: Performance Comparison

Algorithm	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Logistic Regression	93.5	91.2	89.8	90.5
Decision Tree	95.1	93.4	92.7	93.0
Random Forest	98.2	97.6	96.9	97.2
XGBoost	97.5	96.8	95.9	96.3

ISSN: 3071-642X
PP: 01-17.

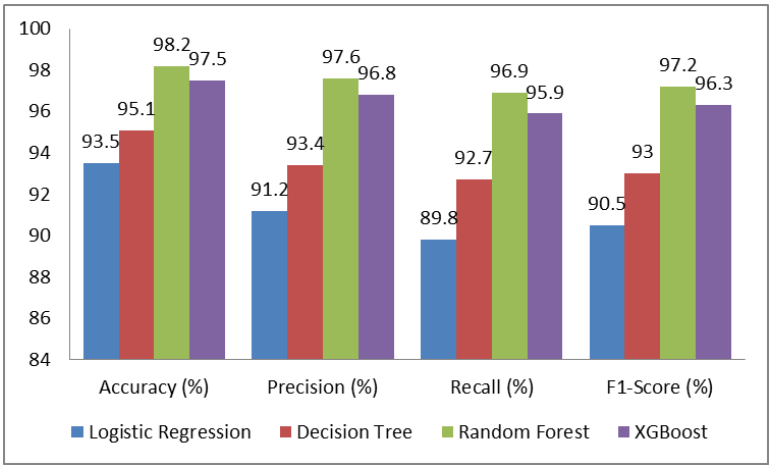

Fig 4 - Performance Comparison

4.2.1. Logistic Regression

Logistic Regression showed an accuracy of 93.5%, which means that it was able to correctly classify a significant number of transactions. The precision was 91.2%, that is, the majority of transactions classified as fraud were fraud cases. The recall score of 89.8% indicates that it was able to identify a good portion of the fraudulent transactions, while some of the transactions were not identified. The F1-Score is 90.5%, which indicates a good balance between precision and recall. Logistic Regression is a simple model, easy to interpret, and has low computational needs, and is therefore a reliable baseline model, but is not effective for fraud patterns that are very complicated and not linear.

4.2.2. Decision Tree

The Decision Tree model performed well when compared to Logistic Regression with an accuracy rate of 95.1%. The accuracy rate of 93.4% suggests that it has greater potential to accurately detect fraudulent transactions and minimize false alarms. The recall score of 92.7% indicates that the model was able to identify the majority of fraud cases it encountered in the dataset. The Decision Tree at F1-Score of 93.0% achieved a good balance between reliability of fraud detection and classification. The model's ability to capture non-linear relationships and give interpretable decision rules has helped in improving its performance in fraud detection.

4.2.3. Random Forest

Overall, the Random Forest algorithm yielded the best results when compared to all other algorithms evaluated. The accuracy of it was 98.2%, which was a very good classification performance. The precision score of 97.6% shows that almost all the transactions marked as fraudulent were accurately identified. Likewise, the recall rate of 0.969 indicates that the model was able to identify almost all the frauded transactions in the dataset. The F1-Score of 97.2% also indicates a good trade-off between precision and recall. The reason for this is due to the fact that Random Forest is an ensemble learning method that uses multiple decision trees to decrease overfitting and increase prediction accuracy.

ISSN: 3071-642X
PP: 01-17.

4.2.4. XGBoost

XGBoost also performed very well in the fraud detection task, with an accuracy of 97.5%. The accuracy of the model in predicting frauds was 96.8%, which is a high accuracy. It has a recall score of 95.9%, which is a great indicator of the effectiveness of its fraud identification capabilities and its reduction of fraud opportunities going undetected. The balanced and reliable classification ability of the model is highlighted by the F1-Score of 96.3% Despite somewhat inferior performance in this experiment, XGBoost still showed a high level of competition thanks to its state-of-the-art boosting algorithm, ability to work with complex data and its excellent generalization capabilities.

4.3. Discussion

It is evident from the experimental results that the machine learning techniques are effective in detecting the banking transaction as a fraud. In comparison, the ensemble classifiers like the Random Forest and XGBoost achieved a remarkable performance compared to the traditional classifiers like the Logistic Regression and Decision Tree. This improvement is due to the ensemble models that can be used to combine several learners and get their predictions to minimize classification errors and enhance model robustness. Among the most challenging problems is fraud detection, as fraudulent transactions can take on the characteristics of subtle, dynamic and hard-to-separate patterns from the regular ones. In such cases ensemble methods often prove to be very effective because they can learn complex interactions between transaction features and be more flexible in modelling changes in the data. Random Forest had the best accuracy of 98.2% and remarkable values of precision, recall and F1 score. The results show that the model was very effective in correctly identifying the fraudulent transactions, with very few false positives and false negatives. This is because Random Forest uses multiple decision trees, each trained on a subset of the data and features, which helps it outperform the other models. The model uses the predictions of many trees to avoid overfitting and to achieve better generalization, so that it can make good predictions for new transaction records it hasn't seen before. Additionally, the Random Forest algorithm is capable of processing high-dimensional data, non-linear relationships, and large datasets, which are prevalent in financial transaction data. The accuracy of XGBoost was also high at 97.5% which further ensured that the boosting techniques worked well in fraud detection applications. Logistic Regression and Decision Tree were found to be satisfactory, but they were less effective as they could not model highly complex fraud patterns as compared to the other algorithms. The results indicate that using only simple classification models might not be enough for present fraud detection systems, which fraudsters are constantly refining. Considering the overall results, the proposed machine learning framework was validated and the need for ensemble learning approaches in the construction of accurate, reliable, scalable fraud detection systems capable of supporting real-time financial security and risk management processes were highlighted.

5. CONCLUSION

Fraud in online banking has become a significant problem for financial institutions in the online world. As online transactions, mobile banking services and digital payment methods are increasingly becoming the way to conduct business, there are new opportunities for cybercriminals to

ISSN: 3071-642X
PP: 01-17.

exploit vulnerabilities and engage in fraudulent activities. The traditional fraud detection models using pre-defined rules and manual monitoring are no longer effective to deal with the evolving and sophisticated fraud attacks. Fraud schemes are continually changing, and there is an increasing demand for intelligent, automated and adaptive fraud detection systems that are able to detect fraudulent activity from a broad spectrum of potential sources with sufficient precision and efficiency. To address this challenge, the current research introduced a machine learning fraud detection framework with the aim of improving the security of online banking transactions by analyzing and modeling data.

The proposed framework incorporated several important stages, such as data collection, data preprocessing, feature engineering, model training, and fraud classification. Different machine learning techniques like Logistic Regression, Decision Tree, Random Forest and XGBoost were applied and compared with the metrics like Accuracy, Precision, Recall and F1-Score. The results of the experiments showed that machine learning methods can be used to identify legitimate and fraudulent transactions with good efficacy since they learn the complex behaviour from the historical information. Random Forest model has shown the best performance with an accuracy of 98.2% with decent precision and recall scores and F1 scores. Encapsulated in its ensemble learning mechanism, it was able to capture the intricate relationships within transaction data and reduce the classification errors, which made it highly suitable for fraud detection applications.

The study also underscored the benefits of leveraging machine learning models for fraud detection, such as greater scalability, agility, decreased manual effort, and help with real-time monitoring. These capabilities can enable financial institutions to detect transactions that may be suspicious, mitigate the risk of financial losses, and build trust with their customers. Advanced deep learning models, explainable AI, federated learning, and real-time streaming analytics could be explored in future works to further enhance detection performance. Overall the results validate that machine learning is a very effective, reliable and sustainable solution to secure the online banking systems and to protect financial assets from the constantly emerging and growing cyber attacks.

REFERENCES

[1] Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2015). Credit card fraud detection: A realistic modeling and a novel learning strategy. IEEE Transactions on Neural Networks and Learning Systems, 29(8), 3784–3797.

[2] Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. Decision Support Systems, 50(3), 602–613.

[3] Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. arXiv preprint arXiv:1009.6119.

[4] West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. Computers & Security, 57, 47–66.

[5] Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. Decision Support Systems, 50(3), 559–569.

[6] Carcillo, F., Le Borgne, Y. A., Caelen, O., Bontempi, G., & Mazzer, Y. (2019). Combining unsupervised and supervised learning in credit card fraud detection. Information Sciences, 557, 317–331.

ISSN: 3071-642X
PP: 01-17.

[7] Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. Information Sciences, 479, 448–455.

[8] Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P. E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. Expert Systems with Applications, 100, 234–245.

[9] Roy, A., Sun, J., Mahoney, R., Alonzi, L., Adams, S., & Beling, P. A. (2018). Deep learning detecting fraud in credit card transactions. Systems and Information Engineering Design Symposium (SIEDS), IEEE, 129–134.

[10] Sahin, Y., & Duman, E. (2011). Detecting credit card fraud by decision trees and support vector machines. Proceedings of the International MultiConference of Engineers and Computer Scientists, 1, 442–447.

[11] Bahnsen, A. C., Aouada, D., Ottersten, B., & Stojanovic, A. (2016). Feature engineering strategies for credit card fraud detection. Expert Systems with Applications, 51, 134–142.

[12] Whitrow, C., Hand, D. J., Juszczak, P., Weston, D., & Adams, N. M. (2009). Transaction aggregation as a strategy for credit card fraud detection. Data Mining and Knowledge Discovery, 18(1), 30–55.

[13] Le Borgne, Y. A., & Bontempi, G. (2005). Machine learning for credit card fraud detection: Practical issues and challenges. International Conference on Computational Intelligence for Financial Engineering, IEEE, 62–68.

[14] Randhawa, K., Loo, C. K., Seera, M., Lim, C. P., Nandi, A. K., & Lal, A. N. (2018). Credit card fraud detection using AdaBoost and majority voting. IEEE Access, 6, 14277–14284.

[15] Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. Journal of Network and Computer Applications, 68, 90–113.