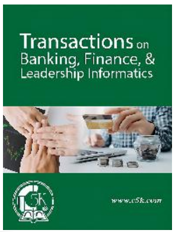
Transactions on Banking, Finance, and Leadership Informatics

"Volume 1, Issue 2, Year 2025"

website: https://www.c5k.com



Research Article

Fraud Transaction Detection using Machine Learning on Financial Datasets

Durga Shahi[1,*], Ali Hassan[2], Muslima Begom Riipa[2], Hafsa Kamal[1], Md Shayakh Alam[3], Arif Ahmed Sizan[1] and Towsif Alam[4]

[1]*Department of Business Administration, Westcliff University, 400 Irvine, CA 92614, USA;*

[2]*Department of Business Administration, International American University, Los Angeles, CA 90010, USA;*

[3]*Department of Engineering Management, Trine University, 1 University Ave, Angola, IN 46703, USA;*

[4]*Department of Marketing Analytics and Insights, Wright State University, 3640 Colonel Glenn Hwy, Dayton, OH 45435, USA;*

**Corresponding Author:* d.shahi.1396@westcliff.edu

ARTICLE INFO

Article history:
9 Sep 2025 (Received)
16 Oct 2025 (Accepted)
25 Oct 2025 (Published Online)

Keywords:
Fraud Detection, Machine Learning, XGBoost, Logistic Regression, and Imbalanced Dataset (SMOTE)

ABSTRACT

Financial fraud poses a significant threat to the digital economy, with credit card fraud being a prevalent challenge. This study evaluates the performance of Logistic Regression (LR) and Extreme Gradient Boosting (XG Boost) models in detecting fraudulent transactions using financial datasets. The study uses practical data from 284,807 transactions, but only 492 are fraudulent; the imbalanced class issue is solved using the Synthetic Minority Oversampling Technique (SMOTE). Our findings show that XG Boost with Random Search selection is better than Logistic Regression in all aspects. XG Boost yielded an accuracy of 99.96%, precision of 95.11%, recall of 79.61%, and F1 score of 86.61%, while for Logistic Regression, the corresponding percentages were 99.92%, 88.1%, 60.5%, and 71.7%. The AUC statistic of 0.98 for XG Boost against 0.97 for LR classified the model as having better discriminant power. The results show that XG Boost is more suitable for real-time fraud detection. However, computational limitations and explainability issues should be considered. For future work, it is suggested that semi-supervised and supervised learning approaches be investigated and work with larger datasets to improve fraud detection in financial systems.

Transactions on Banking, Finance, and Leadership Informatics (TBFLI), C5K Research Publication

1. Introduction

Financial fraud has become an increasingly pervasive issue in today's digital economy, with credit card fraud standing out as one of the most significant challenges. Online shopping and the shift toward cashless transactions have increased the chances of fraudulent activity. The Report (2020) estimates that in 2019, global losses from card fraud will reach $28.65 billion and $35.67 billion by 2023. According to Study (2020), in 2019, the loss due to credit card fraud in the US was an estimated $9.47 billion, which is about one-third of global losses. These statistics are alarming, and we need better fraud detection to prevent these and protect our consumers and financial institutions.

Fraudulent schemes have become much more sophisticated, including skimming, phishing, and identity theft. As the European Central Bank reported in 2020, 73% of fraudulent transactions in 2018 were card-not-present payments, mainly effected online (Bank, 2020). One of the widespread trends illustrates a growing difficulty in monitoring and securing digital transactions, where standard verification methods may fail.

However, many traditional fraud detection systems tend to rely on generating rules based on predefined patterns and thresholds to alert to suspicious activities. While these systems can work to some degree, they aren't good at adjusting to fraudsters' changing tactics and produce high rates of false positives. Not only does it induce customer dissatisfaction, but it also creates very poor operational

*Corresponding author: d.shahi.1396@westcliff.edu (Durga Shahi)

Cite: Durga Shahi, Ali Hassan, Muslima Begom Riipa (2025). Fraud Transaction Detection using Machine Learning on Financial Datasets. *Transactions on Banking, Finance, and Leadership Informatics,* 1(2), pp. 1-XY.

efficiency for financial institutions (Bolton & Hand, 2002). As transaction volumes have grown, the limits of manual or rule-based systems are highlighted, and the requirement for more sophisticated, automated solutions has been strengthened.

Machine learning has been identified as an essential aspect, as it has assisted in handling the challenges arising from fraud detection. Because machine learning deals with large data volumes, algorithms can pick latent features and inconsistencies typical of fraud schemes (Bhattacharyya et al., 2011). These models adapt to new fraud strategies based on the data, making them more adaptive than other models as they experience new strategies in the future. The flexibility and prognosis components of machine learning make it a critical component of contemporary fraud prevention solutions.

Logistic Regression (LR) is a common technique that utilizes a straightforward coefficient estimator for binary classification, such as conjunction or differentiation between fraudulent and actual transactions. Due to non-complex installation, configuration, and precise interpretation, people like using it in many contexts (Yufeng et al., 2004). When analyzing one or more predictor variable data, LR helps estimate the probability of a binary event and assess the risk of fraud. Although it has drawbacks in working with the intricate relationships between multiple forms of fraud, it is easy to use and straightforward to explain.

XGBoost is an extended Gradient boosting algorithm that has received much attention regarding the performance and speed at which it accomplishes classification tasks (Chen & Guestrin, 2016). XGBoost builds a robust predictor model from specific weak models, usually decision trees. It uses the techniques of L1 and L2 norms, respectively, to penalize the model while training to avoid overfitting; it also works well even when large sizes and high dimensions with sparse data – features typically found in large volumes of financial transaction analysis. It is noted that XGBoost can act as a better solution than other machine learning algorithms in different fields; for example, de Sá et al. (2018) considered the field of fraud identification.

This research study aims to compare Logistic Regression and XGBoost in identifying fraudulent transactions in financial datasets. Credibly, a similar study that can be used is the Credit Card Fraud Detection dataset from European cardholders in September 2013 collected from a real dünya environment. Another is that 284,856 transactions with 492 cases of fraud have been found (Dal Pozzolo et al., 2014). The performance of these algorithms will be measured using parameters like accuracy, precision, recall, and area under the ROC curve.

The importance of this study is hereby summed up in the discovery of the suitability of certain models for applying into fraud detecting systems. Since financial organizations are always working to improve their anti-fraud mechanisms, it is important to know the advantages and drawbacks of each of the algorithms described above. Furthermore, by comparing the experimental results of LR and XGBoost, we can determine whether the advantages of simple and interpretable LR are enough or if the powerful features of XGBoost are enough to significantly enhance the capability of distinguishing fraudulent transactions.

Furthermore, this research responds to the problem of unbalanced fraud detection datasets, where the proportion of fraudulent transactions is significantly less than non-fraudulent ones; in the given specific set, it makes up less than 0.172%. If the data sets are skewed towards a particular class, the machine learning models become remarkably insensitive to instances from the minority class (He & Garcia, 2009). Other methods like resampling methods, cost-sensitive learning, or evaluation metrics that will be used for considering class imbalance will be used in the analysis.

The issue of fraud remains a pressing concern for financial institutions, which invest significant resources in developing effective prevention strategies. According to ACFE, 2020, organizations globally lose about 5 percent of their annual revenues to fraud, amounting to more than 4.5 trillion US dollars. This substantial economic impact extends beyond financial losses, affecting brand image and customer confidence, as seen in the case of Fujifilm.

Artificial intelligence, particularly in detecting fraudulent activities, has shifted from type "Reactive" security systems to Proactive security systems. The machine learning models can analyze real physical time data, and significant findings can be shared with decision-makers quickly, which can help organizations defend against different types of threats (Ngai et al., 2011). Thirdly, updating the models with the new data guarantees that the detection systems grow with new fraud strategies.

Some research has focused on using artificial neural networks to detect fraud successfully. In this case, feature selection methods and data preprocessing are employed to prove that logistic regression models for discovering fraudulent credit card transactions were effective, according to Whitrow et al. (2009). Similarly, in the fraud datasets, Pozzolo et al. (2018) found that XGBoost and a cluster of ensemble methods provided a reasonable solution to the efforts focused on handling a class imbalance problem to enhance the detection rates.

Nevertheless, a trade-off problem still exists between the model capacity and the model understandability, which plays an essential role. Thus, in exchange for highly accurate predictions, such models as XGBoost are instead an example of a 'black box,' and making the decision-making procedure transparent is a problem (Lundberg & Lee, 2017). On the other hand, there is technical interpretability in Logistic Regression that will go hand in hand with compliance to set legal standards and thus build stock among the stakeholders.

This research will build on these aspects by investigating not only the forecasting accuracy of both models, which is the ability of the model to predict a target variable correctly, but also the interpretability and usability of the models for a professional user in a real-life financial environment. The study will address questions such as: Is the incremental complexity of XGBoost rewarded with essentially higher detection rates compared to the efficiency of LR? Each model has its way of addressing the issues likely to arise due to imbalanced datasets. What are the consequences for operational and preeminent implementation concerning the computational infrastructure and extent of implementational compatibility?

Through this approach, the research hopes to get a systematic view of the strengths and weaknesses of Logistic Regression and XGBoost in handling this study and detecting fraudulent accounts. The results will be helpful to scholars interested in machine learning applications and practitioners who aim to improve the efficiency of fraud detection models within financial organizations.

2. Literature Review

Protecting customers from fraud is a major theme of investigation that needs to be carried out, particularly with the effects enhanced by electronic payment and online procurement. Through this literature review section, the author provides an overview of the main ML techniques employed in fraud detection: supervised learning, dealing with imbalanced data using SMOTE, and the advantages of the ensemble method, XGBoost.

In a general notion, Abdallah et al. (2016) identify that FDSs face problems such as concept drift, real-time detection, and data skewness. The two highlighted the need for FDS and FPSs to fight these challenges. The authors noted that the FDS and FPSs needed to be in tandem to combat these challenges. Van Vlasselaer et al. (2015) proposed APATE, which integrates transaction and network characteristics in credit card fraud detection. Their studies emphasized the importance of using multiple kinds of data to obtain high AUC scores on detection accuracy.

In fraud detection, Pozzolo et al. (2015) discussed concept drift and class imbalance issues, provided different classifiers for feedback and delayed labels, and combined the results. The cluster average and the moving window adopted in the experiments also enhanced the overall accuracy and the size of the recovered fraud set compared to the simplistic and sequential approach in the reduced real-time environment. Carcillo et al. (2018) tested the active learning strategies in credit card fraud detection when the selection of cardholders for investigation is critical to improving the model, conveniently pointing out the exploration/exploitation dilemma.

Cheng et al. (2020) suggested the use of a spatio-temporal attention-based neural network (STAN) that presents good analytical properties through the connection of spatial and temporal transaction data. They have established that using attention mechanisms improves the identification of suspicious transactions regarding performance-bound parameters compared to other methods. Tiwari et al. (2021) compared several approaches − Hidden Markov Models, Decision Trees, and Support Vector Machines. They recommended that using a diverse family of models, such as Random Forest and XGBoost, will benefit the high-dimensional and nonlinear data category more.

In their paper, Lim et al. (2021) described the machine learning algorithms used in the process of fraud detection, with a focus on the fact that data mining techniques outcompete rule-based traditional methods. The appropriateness of approaches like Artificial Neural Networks (ANNs) and Decision Trees to screen frauds with significant accuracy was also illustrated in the study. Mienye and Jere (2024) offered an understanding of deep learning (DL) approaches in fraud detection, such as CNNs and LSTMs. They said that although DL models are resource-intensive, they produce significantly higher results in identifying fraud patterns.

Singh et al. (2022) were devoted to the tendency of imbalanced data in fraud detection and the methods of oversampling, undersampling, and SMOTE. They discovered that oversampling, together with other ensemble types such as XGBoost, was most effective when working with imbalanced data sets. In the same year, Mînăstireanu and Meșniță also mentioned methods of dealing with imbalanced datasets: cost-sensitive learning and decision trees; they also pointed out that one should choose the right performance indicators for imbalanced cases.

Hajjami et al. (2020) introduced the One-Side Behavioural Noise Reduction (OSBNR) approach to overcome class imbalance and behavioral noise. This method enhanced performance by consolidating close minor class instances while concurrently excluding competitive prominent class examples, increasing the noise in training data. Almhaithawi et al. (2020) used SMOTE and Bayes minimum risk-based cost-sensitive learning technique to improve the detection level for the minority classes and supported their claim wherein SMOTE, along with cost-sensitivity measures, not only improved the savings measure but also showed an impressive result.

Balmakhtar (2021) proposed using ensemble learning hybrid models of EGB and DNNs. Such hybridization showed better results than the single models, and thus, it can be concluded that the best of all can be achieved by combining multiple learning algorithms. Malik et al. (2022) also provided additional proof to support these findings and established that the combination of the Adaboost and LightGBM model, in particular, achieved the best detection accuracy level during the experiments.

Fiore et al. (2019) used GANs to create synthetic data for the minority class in fraud detection, which enhanced classifier performance and provided balanced data for training. Jurgovsky et al. (2018) used LSTMs for sequence classification to show that including transaction sequences helped improve the detection of offline fraud and proposed integrating LSTMs and Random Forest for better overall results.

In their systematic literature review of data mining-based fraud detection (2021), Gupta and Mehta were particular about the superiority of the ML approach over conventional statistical methods where limited labeled data is available. Saheed et al. (2022) showed in a simulation study that for feature selection, PCA can be combined with other supervised learning approaches like KNN and gradient boosting and that this leads to better classification accuracy, in particular when dealing with large feature spaces.

Verma and Tyagi (2022) superposed different supervised learning techniques to differentiate between the two types of fraud and concluded that Logistic Regression and Support Vector Classifiers are the most appropriate when dealing with imbalanced data. The models based on logistic regression yielded the highest accuracy in learning the peculiarity of the fraudulent patterns in conjunction with manual feature engineering.

Elreedy et al. (2024) gave a theoretical explanation for SMOTE, illustrating the problem with its generated sample boasting representativeness. They presented the basis of the distinctions between synthetic samples that can be produced and real minority class samples and highlighted modifications in the oversampling techniques that may enhance the nearness to the actual class distribution.

A hybrid machine learning system, combining supervised and unsupervised learning, was introduced by Vynokurova et al. (2020) for anomaly detection. This system consists of two subsystems: Two, one for anomaly detection through unsupervised learning, and one for anomaly type interpretation through supervised learning. The speed of operation for real-time data was shown to be very high by this approach, and this validated the advantage of combining supervised and unsupervised learning for fraud detection.

In the use of bagging ensemble classifiers for detecting credit card fraud, Zareapoor and Shamsolmoali (2015) developed their work. Complexities inherent to financial datasets, such as large and imbalanced, were remarkably well handled by ensemble learning (i.e., aggregating multiple models to improve performance). Intuitively, the authors found that bagging-based classifiers consistently perform well with accuracy and predictive performance over individual classifiers. This result highlights the usefulness of ensembles in fraud detection, providing an improved generalization by reducing the variance from individual classifiers.

More et al. (2021) proposed an algorithm based on a Random Forest for fraud detection. Their unique approach was to use a learning-to-rank methodology, ranking alerts based on their probability of being fraudulent. The system ranked alerts, allowing investigators to focus on the most promising cases, thus reducing the number of false positives and increasing the efficiency of fraud investigations. This method was also practical, as the study highlighted that ranking algorithms can help investigators allocate resources more effectively for their fraud detection systems.

Hajjami et al. (2020) approached the problem differently by using behavioral noise to describe the problem of imbalanced datasets as a critical challenge in fraud detection. West et al.'s One Side Behavioural Noise Reduction (OSBNR) technique focused on reducing the intersection with the non-fraudulent cases that most likely overwhelm the fraudulent ones. This reduction in behavioral noise significantly affected the classification of fraud detection models. The authors highlighted how much they worked on handling class overlap in the study because even a trim noise level would significantly affect the machine learning models in identifying fraudulent transactions.

The use of hybrid models in fraud detection is achieved here and has been proven to provide effective results, according to Balmakhtar (2021). In their research, Balmakhtar has proposed a combination of two classes of algorithms: Extreme Boosting Gradient (EGB) and Deep Neural Networks (DNN). As discussed earlier, this model achieved better results than individual EGB and DNN models for transaction data, which this study aimed to analyze. The hybrid system improved the solutions offered to identify fraudulent actions taken during financial transactions to optimize the boosting techniques and deep learning.

Likewise, Malik et al. (2022) proposed seven hybrid machine-learning models for credit card fraud and authenticated those models. Of these, the blend of Adaboost and LightGBM gave the best results. They used the two approaches that allowed for the detection of more intricate fraud circumstances that typically involve other algorithms for their identification. From this, it could be concluded that hybrid models are appropriate for the analyzed problem, as flexibility is necessary to address different and changing fraud patterns.

Another excellent contribution was made by Jurgovsky et al. (2018), who implemented LSTM networks over credit card transactions for fraud detection in the transaction sequences. LSTMs are a type of recurrent neural network suitable for sequential data, making them good for fraud detection since transactions are sequential. When comparing the LSTM results with the baseline Random Forest models, the authors found that LSTMs provided a better means of detecting offline transactions, especially where the cardholder was present at the merchant. Therefore, it is proposed that the performance of the fraud detection system can be maximized by using both sequential and non-sequential training methodologies.

Related to the study, there is an attempt to train Generative Adversarial Networks (GANs) to detect fraud. However, Fiore et al. (2019) show that applying GANs in producing synthetic data can be easily used for the minority class. Most of the time, cases of fraudulent payments are overshadowed by genuine ones, or if not, the data sample is skewed in some way, and this creates what is known as a class imbalance, which is detrimental to the performance of the machine learning algorithms. The proposed GANs helped to generate more synthetic data from the minority class, thus effectively counterbalancing the effect of data imbalance on the best-performing fraud detection classifiers. This work demonstrated that GANs offer the possibility of better approaches for addressing class imbalance, hence improving the detection systems.

In their work Verma and Tyagi (2022), they compare and analyse the performance of several supervised machine learning algorithms among them the Logistic Regression and Support Vector Classifiers on imbalanced data sets. They found that their algorithms were capable of dealing with the imbalance common to fraud detection tasks. In fact, they pointed out the role of selecting the right algorithms, which are effective at balancing the skewed class distribution happening routinely in fraud detection scenarios. They find their work contributing to a growing literature advocating for robust algorithms to detect and mitigate fraud issues that present themselves as class-imbalanced problems.

Gupta and Mehta (2021) present a systematic review of data mining techniques and their application in detecting various financial frauds, including credit card fraud. Their meta-analysis reveals that machine learning approaches, particularly those using classification techniques, are more effective than traditional statistical methods. The adaptability of machine learning models to perform well even with limited labeled data, a common scenario in financial fraud detection, is a key finding. This adaptability is crucial when detecting rapidly evolving fraud patterns, as online fraudsters are constantly innovating.

In their study, Suryanarayana et al. (2018) pointed out the usefulness of the logistic regression model in identifying fraud and obtaining better results due to the use of manual feature engineering. This would affirm their idea that although logistic regression is considered one of the fundamental machine learning techniques in the current cohorts, it can provide near state-of-the-art results to cognition when used in conjunction with good feature extraction. The study offered insight into traditional techniques eradication from the modern FRAUD models, significantly when bolstered by domain knowledge and feature engineering.

Lastly, Saheed et al. (2022) applied the principal component analysis, which can be applied to feature selection in fraud detection systems. When including PCA in their experiments and applying supervised learning techniques like KNN and gradient boost, the researchers found that reaching for PCA and picking only the most essential features can improve model performance substantially. This approach not only improved the performance of existing models used in fraud detection but also provided evidence for the effectiveness of using selection techniques for features in improving selected machine learning algorithms that are used in solving specific tasks such as fraud detection.

3. Data and Preprocessing

3.1. Dataset Description

The credit card fraud detection dataset was collected from Kaggle itself, a popular source of machine learning datasets. The particular dataset is the work of a Worldline and the Machine Learning Group of the Université Libre de Bruxelles. We have 284,807 credit card transactions, and this dataset was pre-processed by PCA transformation for confidentiality and performance reasons. The original format of the remaining features is to be labeled as V1 to V28, and the only two features in their original form are the 'Amount' of the transaction and 'Class.'

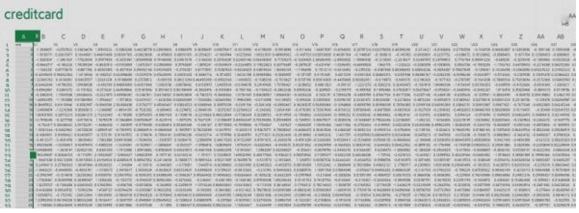


Fig. 1. Screenshot of the excel file.

The 'Class' is the target variable which takes the value of 1 (for fraud) or 0 (for non fraud). One of the primary challenges with this dataset is its significant imbalance: About only 492 of the transactions are fraudulent and makes up about 0.17% of the total dataset. This extreme class imbalance is a key challenge in modeling that necessitates the application of specialized technique such as SMOTE (a synthetic minority over-sampling technique) to be able to be able to train model on the minority class.

3.2. Preprocessing Techniques

The class imbalance handling was also associated with dataset preprocessing, which fitted them, i.e., feature scaling. Since most of these features in this dataset are PCA transformed (as described for the dataset), we must scale these features to allow all features to contribute to the model’s learning process equally. Given the maturity of PCA (Tiwari et al., 2021), the non-transformed features (Time and Amount) were scaled using standardization and normalization techniques to bring them to the same level as the PCA-transformed features.

The extreme class imbalance is a huge challenge in this dataset: we have only 0.17% of fraudulent transactions. In order to counteract this imbalance, the Synthetic Minority

Oversampling Technique (SMOTE) was utilized. It works by generating synthetic versions of minority class examples (fraudulent transactions) using a nearest neighbors' type approach to improve representation for the minority class without replicating instances (Chawla et al., 2002). Figure 2 demonstrates that the original dataset was essentially (i.e., the majority) nonfraudulent transactions in blue and the minority (i.e., least) class represented in red. SMOTE also helped the dataset achieve a more balanced distribution that was favorable for the machine learning model to detect fraudulent patterns rather than drown them into the majority class bias.

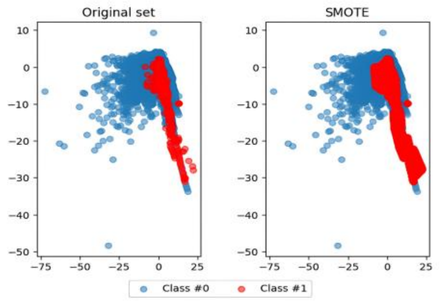


Fig. 2. Before and After Rebalancing Dataset.

Figure 2 illustrates the impact of SMOTE, showing scatter plots before and after applying the technique. The balanced dataset following SMOTE application shows a significant increase in the minority class, leading to better model training and generalization (Singh et al., 2022).

4. Methodology

In this study, we use Logistic Regression and XGBoost to build machine-learning models to detect fraudulent financial transactions. We train and evaluate these models using a dataset that includes both legitimate and fraudulent examples, which enables us to determine how effective these algorithms are. Here, we implement Logistic Regression and XGBoost with respective preprocessing and hyperparameter optimization techniques discussed below.

4.1. Logistic Regression

Logistic Regression is a widely used model for binary classification problems, such as fraud detection, aiming to distinguish between two classes: transactions that are correctly or improperly designated as fraudulent. Logistic Regression is often chosen due to its simplicity and interpretability with a small dataset or when the linear boundary of the dataset is so clear that it can be categorized (Yufeng et al., 2004). This is where we used LR as the baseline model to label transactions as fraudulent or nonfraudulent

The main strength of Logistic Regression is that it is pretty interpretable; we know the impact of each feature on the likelihood of fraud. However, it tends to carry complex data, mainly where nonlinear relationships are essential. Logistic Regression has this limitation and can lead to less exciting fraud detection accuracy if we seek to factor in fraud and do not have fixed fraud patterns that we can rely on. However, we implemented the LR functionality available in sci-kit-learn using a regularisation parameter to minimize overfit while maintaining a library of significant predictive features (Whitrow et al., 2009).

4.2. XGBoost with Random Search

XGBoost is a stronger and robust machine learning technique that has worked well with large datasets and is highly accurate in classification tasks such as fraud detection (Chen & Guestrin, 2016). XGBoost is a method that creates an ensemble of weak classifiers, usually decision trees, to learn from data a strong predictive model able to learn very complex patterns (Xuan et al., 2018).

After implementing our method, XGBoost is trained using random search hyperparameters via three-fold validation to find the optimal set of hyperparameters, such as learning rate, maximum depth, and several separate estimators for each day. We avoided using Grid Search and opted for Random Search because it is less consuming than Grid Search when searching through hyperparameter space (de Sá et al., 2018).

4.3. Model Evaluation Metrics

We compared both models' performance using several important parameters and metrics, especially in fraud detection such as Accuracy, Precision, Recall, F1 score, and area under the curve (ROC-AUC). These metrics provide a comprehensive understanding of each model's ability to classify transactions as fraudulent or non-fraudulent correctly:

- Accuracy measures the percentage of correct predictions made by the model. Logistic Regression gave an accuracy of 92.92%, while XGboost also had an accuracy score of 92.96% for our dataset.

- Fraud detection is very precision-dependent, as it shows how many actual frauds are among predicted frauds. This shows that XGBoost can better minimize false positives (with precision of 95.11%) than Logistic Regression (88.1%).

- Recall, or sensitivity, is a measure of how well the model correctly identifies fraudulent transactions. This study shows that Logistic Regression has a recall of 60.5% while it is proven that XGBoost has a recall of 79.61%, representing the better identification of fraudulent cases.

- The F1 score is a combination of precision and recall. We see that XGBoost detected fraud with a higher F1 score (86.61%) than Logistic Regression (71.7%), and thus, it performs better overall in classifying a given email as fraudulent or legitimate.

- ROC-AUC measures the model's ability to discriminate between classes using different thresholds. Logistic Regression obtained an AUC of 0.97, and XGBoost attained an AUC of 0.98, beating the former.

5. Experiments and Results

5.1. Model Training and Testing

The dataset utilized in this research was split into training and testing sets, adhering to a 70:30 ratio. We applied 70% of the dataset to train our models and kept 30% for evaluation. This approach's main aim was to develop comprehensive models while testing them against data that had never been seen before to assess model generalization ability on unseen instances. The split between the training and test sets guarantees that the model performs reliably and will avoid overfitting the training set. Still, it also provides a means by which the model can be tested on real-world transactions.

Evaluation of the models include Logistic Regression and Extreme Gradient Boosting (XGBoost), each with their advantages and disadvantages when applying to a fraud detection problem. Figure 3 and Figure 5 represent the performance metrics of Logistic Regression model and Figure 4 and Figure 6 represent the performance metrics of XGBoost model.

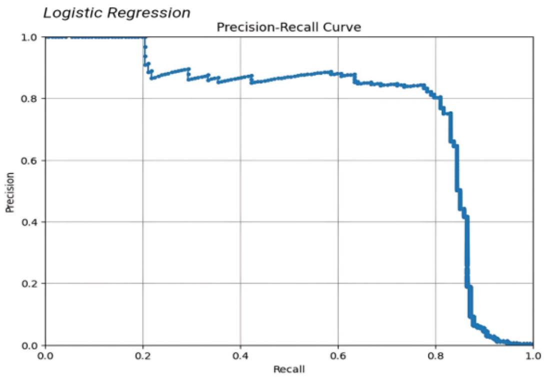


Fig. 3. Precision-Recall Curve of Logistic Regression Model

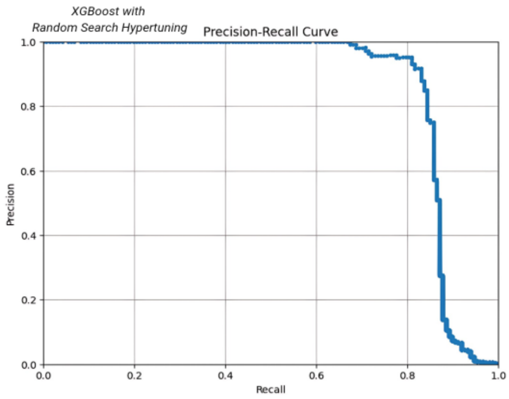


Fig. 4. Precision-Recall Curve of XGBoost with Random Search Model

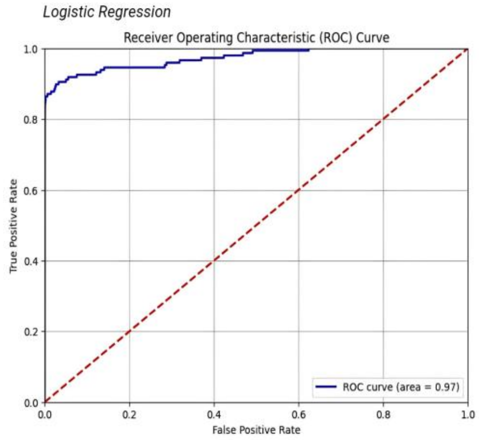


Fig. 5. ROC Curve of Logistic Regression Model

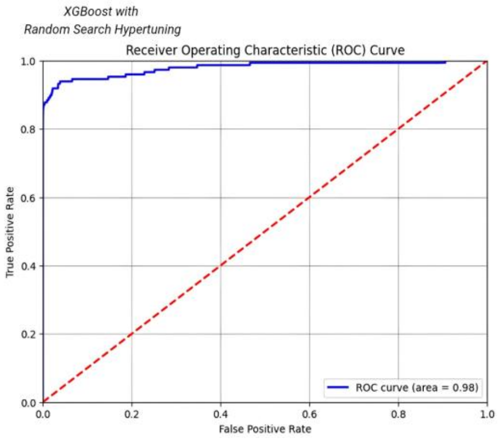


Fig. 6. ROC Curve of XGBoost with Random Search Model

Figure 3 shows the Logistic Regression model's precision-recall curve, which helps us understand how well the model classifies transactions. The logistic regression model correctly classified most transactions as fraud or non-fraud, with an accuracy of 0.9992. However, it is crucial to note that accuracy, in isolation, can be misleading in the context of fraud detection due to the heavily imbalanced nature of the dataset, where non-fraudulent transactions vastly outnumber fraudulent ones.

For Logistic Regression, the precision—which represents the proportion of accurately identified fraud cases compared to the total flagged as fraud—stood at 88.1%. This high precision indicates that the model effectively minimizes false positives, ensuring that the transactions flagged as fraudulent are likely to be truly fraudulent. However, the recall was 60.5%, revealing a moderate limitation of the Logistic Regression model. Recall is essential in fraud detection as it measures the ability of the model to capture all the actual fraudulent instances. The recall rate here suggests that the Logistic Regression model fails to identify almost 40% of fraudulent cases, which could pose a risk if applied in practice. The F1 score, which balances both precision and recall, was

71.7%, highlighting a moderate overall performance of the Logistic Regression model in this specific context.

The ROC curve of Logistic Regression, presented in Figure 5, further highlights the model's discriminative capability. The AUC (Area Under the Curve) value of 0.97 demonstrates the model's effective ability to differentiate between legitimate and fraudulent transactions, but this result should be considered in conjunction with the relatively low recall rate discussed earlier.

However, the XGBoost model outperformed in all the performance metrics. The XGBoost model was about 99.96% accurate, and the Precision-Recall curve for it is shown in Figure 4. XGBoost proves a slight improvement in accuracy compared to Logistic Regression, indicating that it is better at classifying transactions correctly. What's more important, though, is how precise the XGboost model was at 95.11%, much higher than the precision of Logistic Regression. Therefore, it appears that the XGBoost model is more effective in eliminating false positives, which means we have often identified fraudulent transactions of actual fraudsters (Correa Bahnsen et al., 2016).

Similar recall rates were observed, as XGBoost was recalled considerably more often at 79.61%. The result is that compared to Logistic Regression, this figure is a significant step up as it shows that XGBoost can better detect which transactions are fraudulent (and hence can avoid false negatives). The F1 score of 86.61% reinforces that XGBoost delivers higher precision with recall and has a more robust performance to detect fraud. In a financial context, the outcomes of false transactions (false negatives) are very dire, so this balance is significant.

The ROC curve of XGBoost for its AUC shows an AUC value of 0.98, slightly better than that of Logistic Regression, as shown in Figure 6. The larger the value of AUC, the more XGBoost can differentiate a fraudulent and a non-fraudulent transaction for the same classification threshold. This expanded ability is crucial for attaining a flexible tradeoff between sensitivity (true positive rate) and specificity (true negative rate), particularly in imbalanced datasets like this study.

To optimize the performance of the XGBoost model, a Random search using 3-fold cross-validation was conducted (Fig 4). This experimental approach allowed for efficient exploration of hyperparameter space in a manner that does not require the use of all the computational costs of a complete Grid Search. Using the scores, we filtered out parameters like learning rate, maximum depth, and the number estimators to maximize the model's ability to detect fraudulent activities. Results show that using Random Search across multiple metrics brings a significant performance improvement.

When comparing the two models, we noticed that XGBoost is much better than Logistic Regression regarding precision and recall. Although both models provided high accuracy, the other metrics, such as Recall, Precision, and F1 score, reveal the significant keyholes between the two models. XGBoost achieves as low as 5.11% precision and 79.61% recall, suggesting that not only does it reduce the number of false positives, but it also effectively identifies more fraudulent cases, making it a better model for real-world fraud detection (Nielsen, 2016).

The significant improvements observed in the XGBoost model can be attributed to its gradient boosting mechanism, which builds an ensemble of weak classifiers (decision trees) to produce a robust and powerful predictive model. This characteristic, along with hyperparameter tuning through Random Search, enabled XGBoost to outperform Logistic Regression in almost all the metrics considered (Chaudhary et al., 2012).

5.2 Confusion Matrices:

The confusion matrices in Figures 7 and 8 show the models' ability to classify fraudulent and non-fraudulent transactions. The confusion matrices provide a breakdown of four important metrics: True Negatives (TN), False Positives (FP), False Negatives (FN), and True Positives (TP) All of these aspects are important to assess the degree to which the models separate the real and fake transactions.

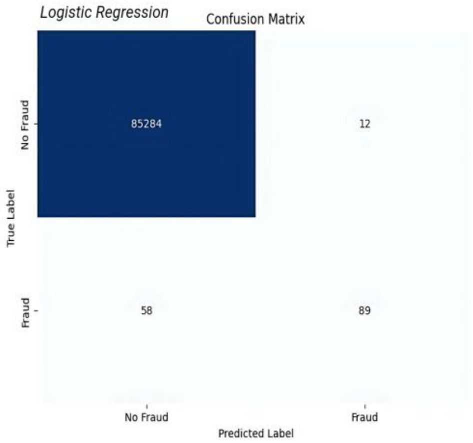


Fig. 7. Confusion Matrix of Logistic Regression Model

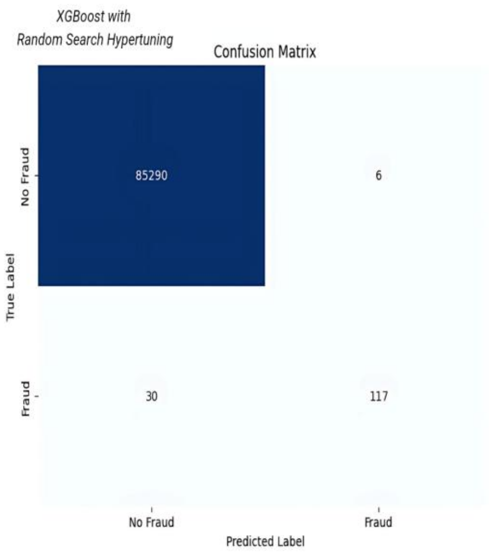


Fig. 8. Confusion Matrix of XGBoost with Random Search Model

The confusion matrix for the Logistic Regression model is shown in Fig 7 below, which displays the model's prognosis results of the transaction as either fraudulent or non-fraudulent. From the given matrix, it is clear that several True Negatives, which is the number of correctly classified non-fraudulent transactions, is 85284. The False Positives (FP), referring to legitimate transaction samples falsely labeled fraudulent, is equivalent to 12. In contrast, the model found 89 TP, which shows the number of fraudulent transactions correctly predicted by the model (Whitrow et al., 2009). However, the False Negatives (FN), which are the fraudulent transactions the model misses, are 58. This indicates that although the Logistic Regression model accurately flagged more non-fraudulent transactions, its recall was relatively less because there were more False Negatives (58), meaning it had missed a few fraud cases (Brownlee, 2016).

Figure 8 shows the confusion matrix for the XGBoost model, and the results indicate an improvement in most metrics compared to Logistic Regression. The True Negatives for XGBoost are 85,290, slightly higher than those of Logistic Regression, indicating fewer legitimate transactions incorrectly flagged as frauds. The False Positives stand at 6, which is a reduction from 12 in the Logistic Regression model, reflecting improved precision in correctly flagging fraudulent activities without false alarms. The True Positives for XGBoost are 117, which is significantly higher compared to the 89 obtained by Logistic Regression. This indicates that XGBoost has a better recall and can more effectively identify fraudulent transactions. Additionally, the number of False Negatives is reduced to 30, which means fewer fraudulent transactions went undetected compared to Logistic Regression (Chen & Guestrin, 2016).

5.3. Model Comparison

In Table 1, the result of comparing Logistic Regression and XGBoost with the Random Search shows that the two models have different performances regarding essential metrics for checking the model's suitability for fraud detection. For the accuracy measurement, Logistic Regression made it to 99.92 percent, which meant that the classifier was valid enough to classify most transactions (Yufeng et al., 2004). However, it achieved a higher precision score of 88.1% and a low recall score of 60.5%, which tells us that although the model did not detect many false positives, it did not capture all the fraudulent transactions. This trade-off is revealed in the F1-score of 71.7%, implying a moderate performance of models on average recall and precision. Irrespective of very high accuracy, other metrics have confirmed the ability of the model to distinguish well between fraud cases and no fraud cases, as indicated by ROC-AUC = 0.97 (Whitrow et al., 2009).

Table 1. Comparative Performance Metrics of Logistic Regression and XGBoost

Metric	Logistic Regression	XGBoost
Accuracy	99.92%	99.96%
Precision	88.10%	95.11%
Recall	60.50%	79.61%
F1 Score	71.70%	86.61%
ROC-AUC	0.97	0.98

On the other hand, XGBoost performed better or almost as well as Logistic regression in nearly all the categories. The presented model achieved an accuracy of 99.96 % and precision of 95.11 %, which is significantly higher than the previous model, while the recall was 79.61 %; it outperforms the last model in terms of minimizing both false positive as well as false negative, which is extremely important in the financial conditions (Chen & Guestrin, 2016). The F1 score of 86.61 presents a good balance between the recognition of fraud and the minimum false positive ratio. A ROC-AUC of 0.98 shows that, on average, across several different threshold levels, XGBoost outperforms the other models and, for this reason, is a more practical approach to fraud detection.

6. Discussion

We found that the XGBoost model outperformed the Logistic Regression model in that part, exhibiting further differences in the strengths and disadvantages of both models in different metrics across the performance. Although chosen for its simplicity and interpretability, Logistic Regression achieved an accuracy of 99.92%, representing only part of the image within a very imbalanced dataset (Yufeng et al., 2004). We found that the precision of Logistic Regression was 88.1%, meaning the model could minimize false positives. However, its recall fell to only 60.5 percent, meaning a lot of fraudulent transactions went unchecked. However, the relatively low recall is problematic for a real-world scenario, as missing fraud cases could lead to substantial financial losses (Whitrow et al., 2009). Logistic Regression has an F1 score of 71.7%, aka, it has difficulty balancing precision and recall. The sound overall discrimination ability shown by the model, as indicated by the ROC-AUC value of 0.97, leaves it lacking the ability to comprehensively cover all fraudulent activities because it cannot reach the recall points.

In contrast, Gradient boosting methods like XGBoost outperformed all the methods across every metric, illustrating that Gradient boosting is effective for large, complex datasets (Chen & Guestrin, 2016). However, Logistic Regression, to some extent, improved the accuracy to 99.96%. Nevertheless, XGBoost's advantages over compared models can be seen for more critical metrics like precision, recall, and F1 score. XGBoost achieved a false positive precision of 95.11%, thus reducing the false positive rate and stabilizing the reliability of classifying a fraudulent transaction while the transaction was not falsely flagged as fraudulent. XGBoost could better capture fraudulent transactions than Logistic Regression (recall of 79.61% greater than Logistic Regression recall). An 86.61% F1 score indicates a better balance between precision and recall, which are extremely important in the fight against fraud, as false negatives, or fraudulent payments, are the most crucial consideration. Another piece of corroboration — the ROC-AUC value of 0.98 — verifies that XGBoost can distinguish fraud transactions from nonfraud transactions.

The Synthetic Minority Oversampling Technique (SMOTE) had a critical impact on the performance of both models, especially considering the class imbalance of the set, in which fraudulent transactions represented 0.17% of the entire set (Chawla et al., 2002). To amplify the training set with frauds, we used SMOTE to generate synthetic examples of the minority class. It also helped both models understand fraudulent patterns without being flooded with the non-fraudulent majority. SMOTE proved helpful for logistic regression, improving recall by 60.5%, but performance was undermined compared to XGBoost. Unlike XGBoost, XGBoost could leverage the better class distribution for 79.61% recall. According to Bian et al. (2016), SMOTE, in conjunction with ensemble learning methods such as XGBoost, typically performs the best on datasets with highly imbalanced data, and our improved metrics support this.

XGBoost performs much better than other methods, but some limitations must be noted regarding model complexity and overfitting risks. One issue with XGBoost is that its decision-making process needs to be revised, a feature that is anti-regulatory compliance in financial institutions (Lundberg & Lee, 2017). Besides, the high dimensionality, complexity of the model, and intimate relationship between the features increase the risk of overfitting. Despite Random Search hyperparameter tuning bringing us one step closer to freeing overfitting through optimization of model parameters, its risk persists as the model is complex.

There is a possibility that a limitation of XGBoost is that of the computational resources and processing time. Training and keeping the XGBoost model for large financial datasets and real-time fraud detection leads to prohibitive computational costs. The 3-fold cross-validation for hyperparameter tuning is computationally expensive and may be a limiting factor for institutions with limited resources. On the contrary, logistic regression is a set of model updates that quickly occur in the systems that necessitate rapid model updates. This is an advantage, but one which is diminished by the lower detection capability of the model, especially for cases of fraud with complex, nonlinear relationships in the data (Ngai et al., 2011).

7. Conclusion

From the analysis of model performances on various evaluated vital metrics, we can see that XGBoost with Random Search has performed well over Logistic Regression in all key metrics. Logistic Regression has an accuracy of 99.92% but a recall of only 60.5%, meaning that it will not be able to capture all the fraudulent transactions and may result in a hazardous gamble in the real Financial World. For instance, on the one hand, XGBoost had a better overall performance accuracy of 99.96%, precision of 95.11%, and recall of 79.61%, and thus led to a much-improved F1 score of 86.61%. We also showed that XGBoost had a superior discriminative capability with ROC-AUC of 0.98 compared to LG's ROC-AUC of 0.97.

These findings have practical implications that indicate that XGBoost is a more desirable model for real-time fraud detection systems. It can minimize false positives and negatives better than other models, essential for maintaining security and customer trust. XGBoost is robust in identifying fraud patterns, especially on highly imbalanced datasets, and is an ideal choice for high-scale financial systems whose cost of undetected fraud may be very high. Although XGBoost is complex and computationally heavy, it may be challenging to utilize in operational deployments for those institutions with limited resources.

In the future, other advanced machine learning models can be experimented with, or hybrid methods can be developed to improve the detection rates further. Second, the scalability and adaptability of these models in more diverse financial settings will be explored with even larger and more complex datasets. Advancing fraud detection mechanisms closer will be contingent on doing so with an understanding of the trade-off between model complexity, interpretability, and performance.

References

Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications*, *68*, 90-113. https://doi.org/https://doi.org/10.1016/j.jnca.2016.04.007

Almhaithawi, D., Jafar, A., & Aljnidi, M. (2020). Example-dependent cost-sensitive credit cards fraud detection using

SMOTE and Bayes minimum risk. *SN Applied Sciences*, *2*(9), 1574. https://doi.org/10.1007/s42452-020-03375-w

Balmakhtar, M. (2021). *Experimental machine learning and deep learning credit card fraud detection* Indiana University of Pennsylvania].

Bank, E. C. (2020). *Report on Card Fraud.* E. C. Bank. https://www.ecb.europa.eu/pub/pdf/cardfraud/ecb.cardfraudreport202305~5d832d6515.en.pdf

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. https://doi.org/https://doi.org/10.1016/j.dss.2010.08.008

Bian, Y., Cheng, M., Yang, C., Yuan, Y., Li, Q., Zhao, J. L., & Liang, L. (2016). Financial fraud detection: a new ensemble learning approach for imbalanced data.

Bolton, R. J., & Hand, D. J. (2002). Statistical Fraud Detection: A Review. *Statistical Science*, *17*(3), 235-249. http://www.jstor.org/stable/3182781

Brownlee, J. (2016). *XGBoost With python: Gradient boosted trees with XGBoost and scikit-learn*. Machine Learning Mastery.

Carcillo, F., Le Borgne, Y.-A., Caelen, O., & Bontempi, G. (2018). Streaming active learning strategies for real-life credit card fraud detection: assessment and visualization. *International Journal of Data Science and Analytics*, *5*(4), 285-300. https://doi.org/10.1007/s41060-018-0116-z

Chaudhary, K., Yadav, J., & Mallick, B. (2012). A review of fraud detection techniques: Credit card. *International Journal of Computer Applications*, *45*(1), 39-44.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *J. Artif. Int. Res.*, *16*(1), 321–357.

Chen, T., & Guestrin, C. (2016). *XGBoost: A Scalable Tree Boosting System* Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA. https://doi.org/10.1145/2939672.2939785

Cheng, D., Xiang, S., Shang, C., Zhang, Y., Yang, F., & Zhang, L. (2020). Spatio-Temporal Attention-Based Neural Network for Credit Card Fraud Detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(01), 362-369. https://doi.org/10.1609/aaai.v34i01.5371

Correa Bahnsen, A., Aouada, D., Stojanovic, A., & Ottersten, B. (2016). Feature engineering strategies for credit card fraud detection. *Expert Systems with Applications*, *51*, 134-142. https://doi.org/https://doi.org/10.1016/j.eswa.2015.12.030

Dal Pozzolo, A., Caelen, O., Le Borgne, Y.-A., Waterschoot, S., & Bontempi, G. (2014). Learned lessons in credit card fraud detection from a practitioner perspective. *Expert Systems with Applications*, *41*(10), 4915-4928. https://doi.org/https://doi.org/10.1016/j.eswa.2014.02.026

de Sá, A. G. C., Pereira, A. C. M., & Pappa, G. L. (2018). A customized classification algorithm for credit card fraud detection. *Engineering Applications of Artificial Intelligence*, *72*, 21-29. https://doi.org/https://doi.org/10.1016/j.engappai.2018.03.011

Elreedy, D., Atiya, A. F., & Kamalov, F. (2024). A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Machine Learning*, *113*(7), 4903-4923. https://doi.org/10.1007/s10994-022-06296-4

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. https://doi.org/https://doi.org/10.1016/j.ins.2017.12.030

Gupta, S., & Mehta, S. K. (2021). Data Mining-based Financial Statement Fraud Detection: Systematic Literature Review and Meta-analysis to Estimate Data Sample Mapping of Fraudulent Companies Against Non-fraudulent Companies. *Global Business Review*, *25*(5), 1290-1313. https://doi.org/10.1177/0972150920984857

Hajjami, S. E., Malki, J., Bouju, A., & Berrada, M. (2020, 19-22 Oct. 2020). A Machine Learning based Approach to Reduce Behavioral Noise Problem in an Imbalanced Data: Application to a fraud detection. 2020 International Conference on Intelligent Data Science Technologies and Applications (IDSTA),

He, H., & Garcia, E. A. (2009). Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*, *21*(9), 1263-1284. https://doi.org/10.1109/TKDE.2008.239

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. https://doi.org/https://doi.org/10.1016/j.eswa.2018.01.037

Lim, K. S., Lee, L. H., & Sim, Y.-W. (2021). A review of machine learning algorithms for fraud detection in credit card transaction. *International Journal of Computer Science & Network Security*, *21*(9), 31-40.

Lundberg, S. M., & Lee, S.-I. (2017). *A unified approach to interpreting model predictions* Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, California, USA.

Malik, E. F., Khaw, K. W., Belaton, B., Wong, W. P., & Chew, X. (2022). Credit Card Fraud Detection Using a New Hybrid Machine Learning Architecture. *Mathematics*, *10*(9).

Mienye, I. D., & Jere, N. (2024). Deep Learning for Credit Card Fraud Detection: A Review of Algorithms, Challenges, and Solutions. *IEEE Access*, *12*, 96893-96910. https://doi.org/10.1109/ACCESS.2024.3426955

More, R., Awati, C., Shirgave, S., Deshmukh, R., & Patil, S. (2021). Credit Card Fraud Detection Using Supervised Learning Approach. *International Journal of Scientific & Technology Research*, *9*, 216-219.

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. https://doi.org/https://doi.org/10.1016/j.dss.2010.08.006

Pozzolo, A. D., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2015, 12-17 July 2015). Credit card fraud detection and concept-drift adaptation with delayed supervised information. 2015 International Joint Conference on Neural Networks (IJCNN),

Pozzolo, A. D., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2018). Credit Card Fraud Detection: A Realistic Modeling and a Novel Learning Strategy. *IEEE Transactions on Neural Networks and Learning Systems*, *29*(8), 3784-3797. https://doi.org/10.1109/TNNLS.2017.2736643

Report, N. (2020). *Global Card Fraud Losses Reach $28.65 Billion.* 1187). The Nilson Report. https://nilsonreport.com/articles/card-fraud-losses-worldwide/

Saheed, Y. K., Baba, U. A., & Raji, M. A. (2022). Big Data Analytics for Credit Card Fraud Detection Using Supervised Machine Learning Models. In *Big Data Analytics in the Insurance Market* (pp. 0). Emerald Publishing Limited. https://doi.org/10.1108/978-1-80262-637-720221003

Singh, A., Ranjan, R. K., & Tiwari, A. (2022). Credit Card Fraud Detection under Extreme Imbalanced Data: A Comparative Study of Data-level Algorithms. *Journal of Experimental & Theoretical Artificial Intelligence*, *34*(4), 571-598. https://doi.org/10.1080/0952813X.2021.1907795

Study, F. R. P. (2020). Update on the U.S. Payments Fraud. Federal Reserve System. https://www.federalreserve.gov/paymentsystems/fr-payments-study.htm

Suryanarayana, S. V., Balaji, G., & Rao, G. V. (2018). Machine learning approaches for credit card fraud detection. *Int. J. Eng. Technol*, *7*(2), 917-920.

Tiwari, P., Mehta, S., Sakhuja, N., Kumar, J., & Singh, A. K. (2021). Credit card fraud detection using machine learning: a study. *arXiv preprint arXiv:2108.10005*.

Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., & Baesens, B. (2015). APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions. *Decision Support Systems*, *75*, 38-48. https://doi.org/https://doi.org/10.1016/j.dss.2015.04.013

Verma, P., & Tyagi, P. (2022). Analysis of Supervised Machine Learning Algorithms in the Context of Fraud Detection. *ECS Transactions*, *107*(1), 7189. https://doi.org/10.1149/10701.7189ecst

Vynokurova, O., Peleshko, D., Bondarenko, O., Ilyasov, V., Serzhantov, V., & Peleshko, M. (2020, 21-25 Aug. 2020). Hybrid Machine Learning System for Solving Fraud Detection Tasks. 2020 IEEE Third International Conference on Data Stream Mining & Processing (DSMP),

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. https://doi.org/10.1007/s10618-008-0116-z

Xuan, S., Liu, G., Li, Z., Zheng, L., Wang, S., & Jiang, C. (2018, 27-29 March 2018). Random forest for credit card fraud detection. 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC),

Yufeng, K., Chang-Tien, L., Sirwongwattana, S., & Yo-Ping, H. (2004, 21-23 March 2004). Survey of fraud detection techniques. IEEE International Conference on Networking, Sensing and Control, 2004,

Zareapoor, M., & Shamsolmoali, P. (2015). Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier. *Procedia Computer Science*, *48*, 679-685. https://doi.org/https://doi.org/10.1016/j.procs.2015.04.201