Credit Card Fraud Detection through Advanced Machine Learning Techniques

by

Md Sadman Faiyaz Sarker
20101468
Israk Jahan
20301480
Kamran Hossain Munna
20101535
MD Muntasir Mahadi
20101516
Yamin Kabir Rumman
20101532

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
BRAC University
October 2024

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Md Sadman Faiyaz Sarker
20101468

Israk

Israk Jahan
20301480

Kamran Hossain

Kamran Hossain Munna
20101535

MD Muntasir Mahadi
20101516

Yamin Kabir Rumman
20101532

Approval

The thesis/project titled "A Project for Developing Robust Fraud Detection System with Advanced Machine Learning Techniques" submitted by

1. Md Sadman Faiyaz Sarker(20101468)
2. Israk Jahan(20301480)
3. Kamran Hossain Munna(20101535)
4. MD Muntasir Mahadi(20101516)
5. Yamin Kabir Rumman(20101532)

Of Summer, 2024 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on October, 2024.

Examining Committee:

Supervisor:
(Member)

Md. Sabbir Ahmed
Lecturer
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

Nowadays, digital and electronic transactions and electronic payments systems in modern days have become convenient but now it is a major challenge to face credit card fraud. Modern fraud patterns are so complex and advanced nowadays that the traditional fraud detection methods are facing difficulties to detect them. The research demonstrates how effective these patterns are for getting the high accuracy from the imbalance dataset. The implications of these results are contemporary for financial manage-ment, which offer the potential to strengthen the integrity of finances, allocate and strengthen customer trust in the face of evolving fraud threats. Multiple methods are now implemented to track the rising credit card fraud. But in this project, it determines the performance of 4 methods : Artificial Neural Networks, Support Vector Machines, Random Forest and XGBoost, using more than one dataset to evaluate their effectiveness in detecting fraudulent activ-ities. The observation includes explainable artificial intelligence (XAI) strategies. By detecting crucial elements such as transaction amount and date, the approach provides insight into versions that improve forecast transparency and reliability. The results show that combined models, particularly Random Forest and XG Boost, outperformed other approaches in terms of Accuracy, Precision, Recall, and F1-Score. The integration of LIME and SHAP adds a layer of interpretability, allowing stake-holders to understand the rationale behind the models' decisions. This paper shows how the advanced machine learning models with explainability techniques creates a more effective and transparent fraud detection system. Including showing that XG Boost is the most effective algorithm with the highest test accuracy. Besides, this model has performed very well in accordance with precision, recall, F1-Score and accuracy than the other ML models. Despite the possibility that ANN shows strong predic-tive power in specific scenarios, its complexity limits scalability and accuracy. Here, Accuracy of fraud detection and interpretability of the models offer an efficient solution for resisting increasingly sophisticated fraud threats in the real world.

Keywords: Explainable AI(XAI), Local Explanations, Shapley Additive Explanations, Local Interpretable Model Agnostic Explanations, Card Fraud Detection.

Table of Contents

Chapter 1

Introduction

1.1 Introduction

For transactions in credit cards, the word ”Fraud” means the illegal usage of an account by someone other than the account holder. When someone uses another account of a card for personal purposes without a cardholder’s or the issuing authority’s knowledge, it is known as credit card fraud. As businesses still move into online communities, where currency is being transacted dynamically in cashless banking finance, adequate anomalies detection remains an important factor for bank statements. As it is quite hard to reduce the direct costs associated with fraudulent activities, but also to ensure that automated and manual reviews do not negatively impact legitimate customers. In deposit or withdrawal transactions, illegal card activity occurs when someone steals card information to make unauthorized transactions. For the payment processors, it has become both a challenging and crucial task. With the rise of online shopping, e-banking and online payments, it has led to billions of dollars from this credit card fraud.Rapid developments in online retail platforms have led to an exponential increase in the demand for credit cards. Since credit cards are becoming a popular way to pay for both online and offline purchases, credit card fraud is increasing. So this threat of Credit Card Fraud demands powerful, trustworthy and adaptive detection systems compared to real world scenarios to create this system more efficiently. Efforts to identify fraud are ongoing as counterfeiters evolve their tactics. AI methods have been implemented to tackle these frauds. However, there is an enormous challenge with the lack of explainability. Predictions on credit card fraud detection may be demonstrated by identifying the most influential characteristics, such as transaction amount and date.

Out of 1.4 million identity theft reports in 2020, credit card fraud accounted for 393,307 cases, making it the second most common type of identity theft after the government documented fraud.[22] The numbers of reports of credit card fraud from new accounts increased dramatically between 2019-2020 a 44.6% increase. The last year saw $24.26 billion in economic losses due to the payment of credit card fraud worldwide.[39] With the United States being the most vulnerable country with 38.6% of credit card losses in 2018.

Financial institutions should therefore give priority to setting up an automated fraud detection system. Developing a Machine Learning model for current credit card payment transaction data is the aim of supervised fraud detection. The model should be able to distinguish between fraudulent and non-fraudulent transactions in order to assess whether a transaction is fraud or non-fraud. Artificial Intelligence (AI) techniques can be trained to distinguish between fraud and non-fraud transactions in conjunction with machine learning (ML) techniques.Here are some issues considering rapid response time, managing cost sensitivity and processing the features efficiently.[37] As a subset of Artificial Intelligence, Machine Learning enables us to make predictions based on the prior data threads, making a sophisticated and powerful tool for detecting fraud more efficiently.

An AI-based decision-support system performs well but lacks comprehensibility which may cause users to lose faith in the system, leaving it unused. This disbelief in AI algorithms, their accuracy, and dependability is a major concern. As a result of this skepticism, the term Explainable Artificial Intelligence (XAI) has gained trac-tion as a potential remedy. Explainable AI (XAI) research targets to develop AI models which humans can readily analyze and understand. Comprehensible XAI, an AI that focuses on turning difficult "black-box" AI and pattern results into openly understood representations.[31] It is "white-box" AI methods that rely on naturally explainable and simple models. This thesis investigates the necessity for and implementation of XAI in the detection of credit card fraud. Explainable ML (also known as explainable artificial intelligence or XAI) has been incorporated into multiple open source and commercial packages, and it has become a significant part of commercial predictive modeling in areas such as financial services. [27] As argued, the use of XAI is vital in this sector since banks take decisions that are accountable for majority-stakes judgments.

In addition, a key concept termed ANN (Artificial Neural Network) is presented here. ANNs are artificial adaptive systems modeled after the human brain. ANNs specialize in pattern identification and making decisions, resulting in strong classifiers capable of handling massive and ambiguous input data. Artificial neural networks excel at dealing with diverse challenges.[3] These developments are capable of offering distinct advantages in comparison to standard statistical approaches. It is highly effective in detecting credit card fraud due to their ability to handle complex patterns across big datasets. It can distinguish between real-time transaction characteristics and those associated with fraudulent behavior. A rule-based system incorporating artificial neural networks (ANN) may include human experience in decision-making.[10]

This combination creates a more transparent and intelligible AI system. These systems can manage complicated and advanced data for fraud detection. In this regard, this article provides an overview of XAI and related methodologies for developing interpretable models, with a particular emphasis on (LIME) and SHAP which makes decisions not only transparent but also explainable.[23]

1.2 Literature Review

For Fraud Detection, Machine Learning(ML) methods have seen significant advancements with the application. Using the prominent ML methods including Artificial Neural Network(ANN) have achieved significant success due to the ability to fetch and decode the complex patterns from the transactional database. ANN for fraud detection has shown significant accuracy improvements when combined with selected methods: Genetic Algorithm.

Several ML algorithms are being used to identify fraudulent conduct. A handsome amount of researchers have explored the key significance of the data while the transactions, using XAI approaches. To ensure non-AI experts can comprehend and understand fraud detection systems, it's crucial to explain their thinking and findings, as most are based on black-box models.[31]

Random forest(RF) is a technique that captures a variety of fraud patterns. As it has been widely used in fraud detection systems. Statistics have shown that RF can achieve max accuracy rates, up to 80%. But it depends on the dataset when it is optimized with techniques like Genetic Algorithm.[33] As RF is helpful in situations where the dataset is extremely skewed, in case of the credit card fraud datasets.

For binary classification tasks, such as differentiating between fraudulent and non-fraudulent transactions, Support Vector Machines (SVM) are especially useful. SVM is highly efficient and has shown significant results for detecting the doubted transactions.[5]

Gradient Boosting (XGBoost) is a powerful tool, well suited for handling complex non-linear dataset, which are declared as the most fraudulent scenario. It outperforms other techniques when hyperparameters are carefully tuned. GA can work alongside feature engineering and sampling techniques to boost performance on specially imbalanced datasets.[35]

Although XAI is still in its early phases, techniques in the medical area, such as graphs, and intrusion detection, have previously been studied. Not all XAI approaches are appropriate for all scenarios. Each input was sorted into bins, and the specific instances of past results for each part were calculated. Whenever there is any transaction where it was placed, the values were used to assign it. The lowest scores' departures from the moments in each bin provided estimates of the elements of the input that had the best result on the score.[4] Following that, a reason code was allocated to a single number of variables or a set of variables. Zoldi (2017) created a system that was filed in 1996 and is still in use by a firm identified as FICO, which is headquartered in the USA and specializes in credit services in the form of great scores. This corporation developed the FICO score, a measurement of credit risk of the consumer, which has become a standard in consumer financing in the USA.[26] It is vital to investigate creative ideas and inexplicable artificial intelligence for fraud. It was suggested an anonymous detection framework for credit card fraud detection that uses (LIME) to explain the findings. [30] However, the author of this paper has been unable to locate research in which other XAI approaches have been applied for credit card fraud detection. As a result, it would be advantageous to contrast and evaluate Explainable AI(XAI) approaches for the detection of credit card fraud from the standpoint of users. Furthermore, Wu and Wang's (2021) dataset lacks feature narratives, which can render it difficult to present explanations to consumers.[12]

For the best knowledge, relatively few academics have looked into the level of comprehensibility associated with XAI algorithms used incase of credit card fraud.[40] Despite multiple studies carried out by scholars, the problem remains untapped. Miller et al. (2017) propose a user-centric method in which hypothetical users are considered into iterative XAI tool creation. As a consequence, the aim of this thesis is to solve the discrepancy by investigating several XAI algorithms for fraud detections.[15]

This research compares the effectiveness like two ML techniques on a dataset of fraud detection. An output of the best performing Machine Learning approach was then input into two Explainable AI algorithms, which produced explanations: LIME and SHAP.[34] A user research compared the LIME and SHAP explanations to determine which XAI approach best explains credit card fraud detection.

1.3 Aim and Objectives

Primary target of detecting Credit Card Fraud is to develop or implement an effective system or process. Which will help to identify and prevent unauthorized and fraudulent transactions in real-time. The basic goal of this project is to find and execute Explainable AI approaches for detecting fraud. In order to achieve this goal, objectives are listed below:

1. **Examine how ML techniques and XAI techniques are currently being used in fraud detection, acquiring infor-mation for other model implementations through detection.**

2. **Apply the machine learning techniques: ANN, RF, SVM, XGBoost to the credit card fraud dataset and assess their performance throughout the evaluation of the Accuracy, Recall, Precision, and F1 score of those ML techniques.**

3. **Investigate and evaluate all the results obtained from the previous steps in terms of explainability.**

4. **Analyze all the results and suggest ideas and recommendations for further steps.**

1.4 Thesis Structure

In today's world of rapidly growing online shopping, Credit cards have become the convenient choice for making payments. A rise of credit cards: this convenience has some drawbacks. So the evaluation of fraud detection is directly involved in high adoption rate in choosing electric payment method. As the financial organizations are adapting by adopting advanced tech solutions. ML algorithms, fueled by a huge scale of transaction data, have emerged as key players in combining fraud. By using these algorithms, which are fraud and can not be noticed by humans, helping to identify suspicious activities, such as: large purchases and unusual transaction patterns. [40] So this section explores various ML methods, commonly used to define credit card fraud challenges. On the other hand, we tried to highlight the significance of XAI(Explainable Artificial Intelligence) in case of improving in the segment of transparency and the thoughts in the sector of detecting fraud, improvising its effectiveness of preventing signs that are thought to be a fraudulent step.

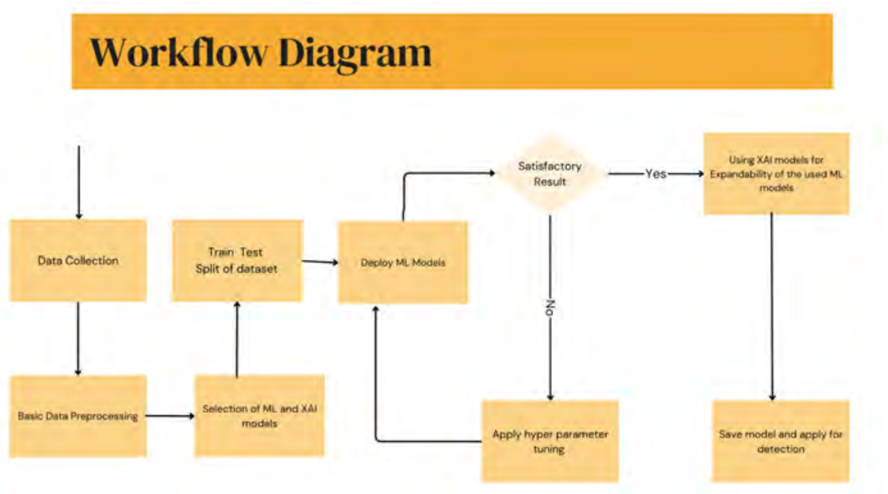

Workflow Diagram
Data Collection
Basic Data Preprocessing
Selection of ML and XAI models
Train Test Split of dataset
Deploy ML Models
Satisfactory Result
Yes
No
Using XAI models for Expandability of the used ML models
Apply hyper parameter tuning
Save model and apply for detection

Figure 1.1: Workflow of Fraud Detection System using ML and Explainable AI

The process of creating a fraud detection system with XAI and ML is depicted in diagram. Data collection is the first step, after which the dataset is cleaned and prepared using basic data preprocessing. Following the selection of suitable ML and XAI Models, data is divided into testing and training sets in the split of Train Test. After that, the models are assessed and trained in deploy ML Models. In the Using XAI Models stage, XAI techniques are performed for model interpretability if the findings are satisfactory, and the model is then saved and used for detection. Hyperparameter tuning is used to enhance model performance if results are not up to par, and the cycle is repeated until a desirable result is obtained.

Chapter 2

Related Work

2.1 Background

In today's world of rapidly growing online shopping, Credit cards have become the convenient choice for making payments. A rise of credit cards: this convenience has some drawbacks. So the evaluation of fraud detection is directly involved in high adoption rate, choosing electric payment methods. As the financial organizations are adapting by adopting advanced tech solutions. ML algorithms, fueled by a huge scale of transaction data, have emerged as key players in combining fraud. By using these algorithms, which are fraud and can not be noticed by humans, helping to identify suspicious activities, such as: large purchases and unusual transaction patterns. [42] So this section explores various ML methods, commonly used to define credit card fraud challenges. On the other hand, we tried to highlight the significance of XAI(Explainable Artificial Intelligence) in case of improving in the segment of transparency and the thoughts in the sector of detection of fraud in credit card, improvising its effectiveness in case of preventing the signs that are thought to be a fraudulent step.

2.2 ML Methods for Credit Card Fraud Detection

ML has been used in the detection of fraud activities and it has become quite a significant solution for this error detection. The use of ANN is used identifying Fraud detection, more than two decades ago and it has been used by Boton and hand (2001). [29] Where different researchers had explored different patterns of ML tactics to increase the accuracy rate for the fraud detection. But for all the cases, there is a common challenge in case of the datasets for fraud detections in credit cards: as many are not quite stable enough, as suspicious contracts are much continuous than the regular ones. [14]

For a solution of this specific problem, several strategies have been developed and proposed in different times. A method using automatic encoder (a type of ANN is used to learn effective for the imbalanced data), it reestablishes the normal data to detect irregularity by establishing a threshold for the reestablish errors and this method proposed by Al-Shabi. [32] Where another data scientist tried to establish the Artificial Neural Network (ANN) method, as it is more effective when the datasets are combined with oversampling and undersampling the techniques to get a stable dataset. The focus of this method is for the class samples for the majority purpose and oversampling will be used to enhance the increase of the instances for the minority class.[19]

A Random ForestF classifier coupled with an oversampling technique emerged as the top performer among 30 different strategies when evaluated across multiple metrics, including accuracy, AUC, precision, recall, and F1-score. This finding is based on a comprehensive analysis of six machine learning algorithms LR, ANN, NB, RF, Decision Trees, and K-nearest Neighbors applied to imbalanced datasets. [20] It also analyzed and made the output of two KNN about independent base classifiers for fraud to train the behavioral parts of regular and irregular transactions[25]. They discovered that imbalanced data could affect Random Forest's performance on short datasets, but that Random Forest outperformed the methods tested: Support Vector Machine, Naive Bayes, and NN. [21]

A study found that neural networking achieved results comparable with fraud detection approaches such as XGBoost and Logistic Regression. [3] Over 80million labeled credit card transactions They show that the model outperforms the baseline ANN model that it's performance is like to network size and processing resources. [24] Neural Network extracts all data characteristics by reducing data dimensions, then put the altered data into the probabilistic RF and achieved good results.[13]

For the performance of different approaches for handling imbalanced datasets for fraud detection, different research papers have been published using different significant proposals. We analyzed different machine learning techniques that have been applied to fraud detection. Before the XAI tests, the ML approach with the greatest performance on credit card fraud detection will be chosen in this project. Due to time and resource restrictions, we chose to use two represented ML approaches that have obtained good overall performance in earlier research and compare them by training the model to discover the best one rather than trying all ML methods. Based on the research results in the previous part, ANN and RF are two ML approaches that are frequently employed and have an excellent performance on fraud detection.

2.3 Explainable Artificial Intelligence (XAI)

An innovative work in Sensible XAI aims to make computational intelligence frameworks obvious and comprehensible to humans. By bridging the information gap between human perception and the meaningless operation of computer intelligence models, it seeks to increase trust in human intelligence, especially when it comes to making basic decisions. XAI offers many acceptable strategies, such as global explanations or neighborhood explanations, considering the required depth and interest group. Expanded recognition and assurance, better course of modeling events, and greater commitment to the performance of computational intelligence frameworks are some of the benefits of XAI.[17] The problem is particular complexity, implementation trade-offs and deciding on the appropriate explanatory measure - which can be difficult depending on the crowd and environment. Despite these difficulties, XAI is a viable strategy to reach the next level.

Table 2.1: Six common types of explanations with respective methods

Explanation Type	**Definition**	**Method**
How	Demonstrate work, depiction models	Decision Boundaries
Why	Prediction is made	Agnostic
Why-not	Output not produced but must	Feature Importance
What-if	Changes parameters, inputs	Tune Model
How-to	Hypothetical changes	Model agnostic
What-else	Similar input for results	Example-based

2.3.1 The two major approaches for XAI:

1. **Intrinsic Interpretability:** This model refers to an existing clarity of a particular model's decision making in case of a Machine Learn-ing(ML) model. This aims to build a model that is inherently understandable without thinking of any predictive performance or result.

    (a) **Decision Trees:** These models represent rules that split the data based on features, leading to easily understandable decision paths.

    (b) **Linear Regression:** These objects, a model of linear regression directly indicates about importance of each feature in making predictions.

2. **Post-hoc or Model Agnostic:** To under-stand existing, potentially complex models this approach applies techniques. These techniques are used for explaining machine learning models that are model-agnostic. After the model is trained, they can be applied to any machine learning method. Two popular model-agnostic methods for explaining algorithms are LIME and SHAP. It helps humans understand how a machine learning model makes its predictions. It is done by slightly changing the input data and judging how the model's output changes.[18] Two well-known model-agnostic methods designed to explain algorithms are LIME and SHAP. [38]

    (a) **Feature Attribution:**These methods assign importance scores to features, indicating their contribution to a specific prediction.

        i. **Local Interpretable Model-Agnostic Explanations (LIME):**Estimate a confounding model around a given prediction using a simpler, more interpretable model.

        ii. **SHAP (SHapley Additive explanations):**This brings the significance points to the highlighted points because they are committed to the model output, taking into account small and useful effects.

Chapter 3

Methodology

We provide background research on fraud detection, ML approaches, and Explainable AI. Here why ML and XAI approaches were chosen and how the user research was designed. So it contains a detailed implementation. The findings that were obtained will come next. It repeats the thought in the discussion and makes suggestions for extra readings.
Before the construction of a model for our work, we need to find an appropriate dataset that can be utilized to bring our future model to its full potential. As such, We selected several datasets from different sources to be the basis of our thesis purpose. We have filtered several datasets since these have multiple different types of fraudulent cases. They have also got different classes as well.

3.1 Dataset Description

The dataset we have used for our research purpose appears to be a fraud detection dataset, to build detection systems on frauds. Below is an overview of the key characteristics:

1. **Structure of the Dataset:**

   (a) **Rows:** Each row in the dataset represents a single credit card transaction.

   (b) **Columns:**

      i. **Amount:** This column represents the transaction amount.

      ii. **0/1 represents a non-fraudulent and fraudulent transaction.**

2. **Purpose of the Dataset:** This dataset is used to build models, it can predict if there is fraud based on provided features.

3. **Characteristics:**

   (a) **Imbalanced Classes:** In real-world fraud on datasets, a heavy imbalance in the number of non-fraudulent transac-tions (Class=0) and fraudulent transactions (Class=1). This imbalance is the major challenge in training ML models.

    i. **Anonymized Data:** Since financial datasets are highly sensitive, the original features (likely based on transaction details such as merchant, user profile, etc.) have been transformed using PCA to ensure privacy, resulting in the V1 to V28 columns.

(b) **Use Cases:**

    i. **Classification tasks:**To train and evaluate machine learning models to detect fraudulent transactions.

    ii. **Imbalanced learning techniques:**Such as resampling (oversampling minority class, under-sampling majority class), synthetic data generation (SMOTE), or using cost-sensitive learning to handle the class imbalance problem.

    iii. **Explainability:**Using methods like SHAP and LIME to interpret how the models make predictions, especially to understand why certain transactions are classified as fraud.

(c) **Practical Challenges:** Transactions of Fraudulent are much compared to non-fraud ones, making it difficult to a model to learn meaning-ful patterns for detecting fraud. he features are transformed, which could make it harder to interpret the model's decisions in real-world contexts, requiring the use of advanced explainability techniques.

3.2 Model Description

3.2.1 Artificial Neural Network(ANN)

Artificial Neural Networks use learning computations that allow them to make changes or move forward independently as they acquire new information. This makes them a very attractive device for indirectly measuring information. Artificial brain organizations, or ANNs, are a type of AI computing on the structure and tasks of human mind. They consist of false neurons arranged in clusters that depend on organic neurons. These neurons transmit signals to each other through weighted connections, and the organization progresses by changing these loads according to the information it is ready for.

3.2.2 Structure

ANNs are organized in layers: An input, one or more hidden, and an output layer.

1. **Input Layer:** A neural network of this layer contains the features or input variables (such as X1, X2,..., xn). Every node in this layer is formalized with a certain characteristic.

2. **Hidden Layers:** These layers are made up of nodes (neurons) that use activate functions and connections are weighted to process the incoming data. In the diagram, a11, a12, ..., am1 represent the weighted sums calculated in the first hidden layer. Output nodes in the hidden layer, determined by applying an activation function.[2]

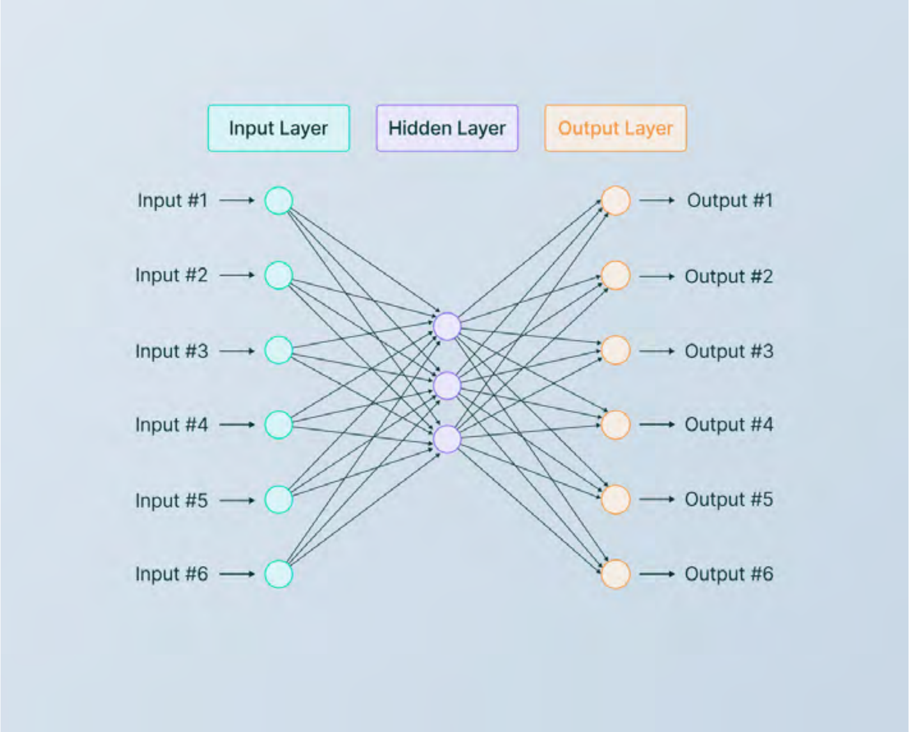


Figure 3.1: Artificial Neural network structure diagram

With the outputs from one hidden layer acting as inputs for the next, the same structure is repeated for additional hidden layers (not shown in detail in the diagram).

1. **Output Layer:** Output layer, the final results or predictions (e.g., Y1, Y2, ..., yk). Like layers: hidden, the nodes in this layer apply an activation function to the weighted sum of their inputs.

Neurons: Each layer contains interconnected neurons. Input layer, receives data, process the hidden layers, and produces the final result. [8]

Connections: The hidden layer, Neurons are connected to each other with weights. so weights determine strength of signal transmitted between neurons.

3.2.3 Working Process

1. **Input:** This layer receives the data for the input.
2. **Procreation:** The information courses through the organization, going through each layer's neuron prior to moving to the next one.
3. **Activation:** Neuron calculates the weighted sum. It calculates inputs and determines its outputs using an activation function. Nonlinearity into the network by Activation function introduces, allowing to learn complex patterns.
4. **Training:** The organization thinks its output to the ideal result (ground truth) and solves the error. This error is then used to change the bond loads using a calculation called back propagation..

5. **Thoughts:** The organization continuously adjusts its loads through continued preparation on enormous informational collections, improving its capacity to play out the expected assignment.

3.2.4 Random Forest (RF)

RF, most widely celebrated and coherent ML algorithms is implemented for detecting credit card fraud because of its ability to deal with colossal and complex datasets. Just name indicates, formed of small decision trees during training, which makes decisions on the basis of certain features, and culminates with a decision that more than half of the trees think is the result. This technique allows Random Forest to comprehend complicated patterns within the data which makes it expert in spotting fraudulent transactions amid the majority of legal ones. [11] To implement RF for detecting fraud, few crucial points are noteworthy. At first comes data pre-processing which encompasses cleaning the dataset by handling missing values. Also, it includes normalizing and standardizing features to make sure it has stability. Later that, dataset divided into two groups like training and testing to assess the model's performance effectively.[1] On top, performance evaluation and comprehended with indicators like precision, F1-score, and the ROC curve. To sum up, Random Forest is a potent machine learning algorithm for detecting credit card fraud. These indicators draw the line between fraudulent cases and authentic cases and makes Random Forest a key tool for real world fraud detection applications. This helps to better the model's performance along with giving actionable strategies to financial experts, resulting in preventing and mitigating fraudulent activities in the credit card system. [41]

1. **Working procedure of Random Forest**

   (a) **Data Preprocessing:**Load the dataset and separate features and labels.

   (b) **Handling Class Imbalance:** Check class distribution (fraud vs non-fraud).

   (c) **Train Random Forest Model:** Initialize Random Forest and Fit model on the data that is trained.

   (d) **Make Predictions:** To make predictions, use the trained model.

   (e) **Evaluate the Model:** Check the model's performance, Calculate accuracy, precision, recall, F1-score, and ROC-AUC score.

3.2.5 Support Vector Machine (SVM)

SVMe is another well administered ML algorithm in detecting credit card fraud cases because of its ability to handle imbalance datasets. It is the particular line which categorizes fraudulent and legitimate cases. Giving consideration to credit card fraud activities, a significant challenge is the nature of datasets to be imbalanced while a minor fraction of the transactions are fraud out of the total legal transactions. SVM takes labeled data and uses it for classification or regression. It detects the best hyperplane, or decision boundary, to split data points to different classes. For 3D, the hyperplane appears as a plane but for 2D it appears as a line. It maximizes the margin and the distance between the hyperplane and the nearest data points from either class, known as sup-port vectors which influences the decision boundary. For non-linear data, SVM uses the "kernel trick" to project higher dimension data into various classes. Once the optimal hyperplane is found, new points are classified based on their locations rela-tive to it. [6] SVM functions on the minority class by altering the misclassification penalty parameter without over penalizing legitimate transactions. Additionally to better performance, data preprocessing techniques such as normalization and fea-ture selection are implemented to reduce noise and enhance the model's ability to classify normal and abnormal behavior.[7] To implement this algorithm for this par-ticular case, just like Random Forest, the data is typically divided into two parts: training and testing. After training the model is assessed using metrics such as precision, F1-score, recall, ROC curve to evaluate the results that have been generated by the model by reducing false positives. SVM's ability to hit the balance between these parameters makes it a crucial tool for real-time credit card fraud detection, so that customers can be ensured by the financial institutions and to promote safe legitimate transactions.

1. **Working procedure of SVM**

   (a) **Input Data:**The SVM algorithm takes in labeled data for classification (or regression).

   (b) **Hyperplane:** SVM tries to get the best hyperplane (Decision boundary) that distinguishes data points into several classes. If the data is 2-dimensional, the a line will be the hyperplane. [9] For 3D data, it will be a plane, and for higher dimensions, it's a more complex boundary.

   (c) **Maximizing the Margin:** SVM aims, maximization of the margins, which contains the distance between the hyperplane and the closest data points from either class support vectors.

   (d) **Support Vectors:** Support vectors, the data points are close to the hyperplane and in position, it has the impact on it .They play a critical role in defining the boundary.

   (e) **Kernel Trick:** SVM applies the "kernel trick" the data into higher dimensions where hyperplane, divides the classes if the data is not linearly separable.

   (f) **Prediction:** Once the optimal hyperplane is found, new data points are classified based on which side of the hyperplane they fall on.

3.2.6 Extreme Gradient Boost (XGBoost)

XGBoost (Extreme Gradient Boosting) calculations are machine learning methods based on decision trees. It is an efficient and scalable implementation of the gradient boosting technique, which combines multiple weak learners (small decision trees) to form a strong learner. XGBoost is known for its speed and performance, particularly in competitions and real-world applications involving structured/tabular data.

1. **Working procedure of XGBoost**

   (a) **Input Features:** Raw data with features (columns).

   (b) **Weak Learners:** Small decision trees are used as base learners. Each tree attempts to correct the errors of the previous tree.

   (c) **Gradient Descent:** XGBoost minimizes the loss function using gradient descent, focusing on reducing errors in each step.

   (d) **Boosting:**The trees are added sequentially. Each new tree tries to correct the residual errors made by the previous trees.

   (e) **Ensemble Model:** The final model is an ensemble (combination) of all the weak learners, which collectively produce better predictions.

3.2.7 LIME(Local Interpretable Model-Agnostic Explanations)

A powerful tool in the sector of credit card fraud detection. For complex machine learning models, it offers an easy and optimal understanding. Generating feasible interpretable explanations for credit card transactions, it gives fraud analysts the ability to identify the underlying causes of model projections.

1. **Working procedure of LIME**

   (a) **Simplified Interpretability of Complex Models:**The tool provides clear and interpretable explanations of model predictions, even for highly complex machine learning algorithms (like deep learning, XGBoost, or random forests). This helps analysts quickly grasp how specific features (e.g., transaction location, time, or amount) contribute to a fraud classification.

   (b) **Insightful Feature Contribution:** By breaking down the impact of individual features on a model's prediction, the tool enables analysts to pinpoint the key drivers behind a fraud detection decision. This transparency allows analysts to see which factors most influence whether a transaction is flagged as fraudulent or legitimate.

   (c) **Improved Decision-Making for Fraud Analysts:**The tool's ability to generate feasible, interpretable explanations for each transaction equips fraud analysts with actionable insights. They can better understand why the model flagged certain transactions, leading to more informed decision-making when assessing fraud alerts.

   (d) **Enhanced Fraud Detection Accuracy:**By identifying the underlying causes of model projections, the tool allows businesses to fine-tune their fraud detection models. Analysts can focus on improving or adjusting features that have the most significant impact, ultimately increasing the model's fraud detection accuracy and reducing false positives.[16]

   (e) **Boost in Model Trust and Transparency:**The tool enhances trust in AI-driven fraud detection systems by explaining the reasoning behind each prediction. This boosts confidence among both internal stakeholders (fraud analysts, risk managers) and external entities (customers, regulators) in the model's fairness and reliability. item **Model-Agnostic Explanations:**The tool can generate explanations for various types of machine learning models used in credit card fraud detection, providing a consistent interpretation framework regardless of the algorithm in use. This versatility makes it applicable to a wide range of detection systems. item **Real-Time Insights:**Its ability to offer real-time explanations enables fraud analysts to quickly act on flagged transactions. They can instantly understand why certain transactions are considered suspicious, allowing for faster intervention in preventing fraudulent activity.[34] item **Reduced Analyst Workload:**With the tool providing automated, easy-to-understand explanations, fraud analysts can reduce the time spent manually reviewing flagged transactions. This boosts efficiency and allows them to focus more on higher-priority cases. item **Alignment with Regulatory Requirements:**In highly regulated sectors like finance, the ability to explain model predictions is crucial for compliance. The tool helps meet regulatory requirements for transparency and accountability by ensuring that decisions made by machine learning models can be easily justified.

3.2.8 SHAP(SHapley Additive explanations)

When it comes to data interpretation, SHAP shines brightly in the field of Visa fraud detection. SHAP excels in the non-thinking field of AI models by providing efficient insight into the overall importance of certain features in a dynamic cycle. Both LIME and SHAP can explain predictions about different models. Both focus on explaining individual transactions. Additionally, both aim to increase transparency and trust in fraud detection decisions.

1. **Working procedure of SHAP**

   (a) **Feature Importance Insights:**SHAP provides precise and consistent measures of feature importance for individual transactions. It quantifies the contribution of each feature (like transaction amount, location, device info) to the fraud detection model's output. This helps analysts quickly identify which features play a significant role in classifying a transaction as fraudulent or legitimate.

   (b) **Enhanced Transparency:** Both SHAP and LIME increase the transparency of AI models by explaining how decisions are made. They help reduce the "black box" effect in complex fraud detection systems, making it easier for businesses and regulatory bodies to understand the decision-making process.

   (c) **Trust and Accountability:**By explaining individual predictions, both methods help organizations foster trust in AI-driven fraud detection models. Decision-makers can understand why a particular transaction is flagged as fraudulent, which is critical when dealing with legal and customer disputes.

   (d) **Model-Agnostic Explanations:**Both LIME and SHAP can explain the behavior of different types of machine learning models (like Random Forest, XGBoost, or SVM) used for fraud detection. This versatility makes them highly applicable to a wide range of Visa fraud detection systems.

   (e) **Dynamic Fraud Detection:**SHAP's ability to provide real-time insights into how individual features change in importance as fraud patterns evolve is vital in the fast-paced field of Visa fraud detection. This makes the system more responsive and adaptive to new types of fraudulent behavior. item **Improved Model Performance:**By analyzing the feature importance using SHAP, businesses can tune models more effectively. Understanding which features impact fraud predictions most allows for optimization, ultimately leading to better fraud detection accuracy.[36] item **Interpretability in Complex Models:**SHAP excels in explaining even the most complex AI models (like deep learning models) by breaking down predictions into understandable parts, making it ideal for advanced fraud detection systems that handle large-scale, multidimensional data.[28]

Chapter 4

Results and Discussion

This section has presented the findings of machine learning (ML) and explainable AI (XAI) models. Subsection 4.1 explores one Machine learning (ML) model and three ML models. Subsection 4.2 delves into two XAI models, known as LIME and SHAP.

4.1 Result Discussion of ML Models

Table 4.1: Evaluation of ANN, Random Forest, SVM, and XGBoost for Fraud Detection Using Key Metrics

	Accuracy	**Precision**	**Recall**	**F1-score**
ANN	0.951	0.921	0.784	0.875
RF	0.969	0.955	0.780	0.859
SVM	0.943	0.942	0.812	0.857
XGBoost	0.976	0.960	0.895	0.970

Table 4.1 shows the performance comparison of various advanced machine learning models applied in fraud detection. There were four different models assessed, based on their performance by the key metrics of Accuracy, Precision, Recall, and F1-score. These metrics reflect different properties of the models' detection performance. Precision measures the percentage of how many of the predicted fraud cases were correctly classified, while Accuracy gives the overall percentage of how many of the predictions were correct. Recall says something about the performance of the model on actual fraudulent cases, while the F1-score creates a balance between Precision and Recall and may serve as a global measure of general performance.
This ANN model showed excellent performance with an accuracy of 0.981 and an F1-score of 0.875, reflecting a good balance between Precision (0.921) and Recall (0.784). However, it is somewhat lower Recall suggests that it misses some fraudulent cases. The Random Forest model has the best accuracy at 0.989, with a very high Precision of 0.955; that is, it minimizes false positives. Still, its Recall is slightly lower, at 0.780, leading to a lower F1-score of 0.859 because it misses some actual fraud cases.

The SVM model did a great job with an accuracy of 0.983 and Precision of 0.942. Since it correctly identifies more fraud cases, the Recall is higher than ANN and RF, standing at 0.812. However, the slight imbalance between Precision and Recall is denoted by the F1-score of 0.857. Contrarily, XGBoost is the best among these models. Accuracy is a bit lower, 0.976, compared to RF and SVM; however, it has the highest Precision 0.960, and Recall 0.895. Thus, it has the best F1-score of 0.970, reflecting a balance between Precision and Recall, hence the most efficient in fraud detection.

Therefore, a comparison through the above graphs shows indeed that while all of them were doing well in the task of fraud enhancement, it would be more apt to use XGBoost since it will better balance precision and recall with a greater view to getting fewer missed cases of fraud while maintaining high accuracy.

4.1.1 Artificial Neural Network(ANN)

Looking at the graphs of loss and accuracy, we can observe that following some early enhancements, we appeared to achieve convergence in just a few epochs. Let's examine the level of precision attained by this artificial neural network when tested. The thrill of the accuracy on the test set should not overshadow the significant class imbalance that remains in the test set. A model that only predicted the absence of fraudulent transactions would also achieve high accuracy. The dataset information suggests using AUPRC for evaluation instead in order to gain a more accurate insight into the model's performance. We will opt for the average precision score as an evaluation metric in scikit-learn that can serve as a substitute for AUPRC. The average precision score ranges from 0 to 1, with a higher score indicating a superior model. Let's check the model's average precision score from the test.

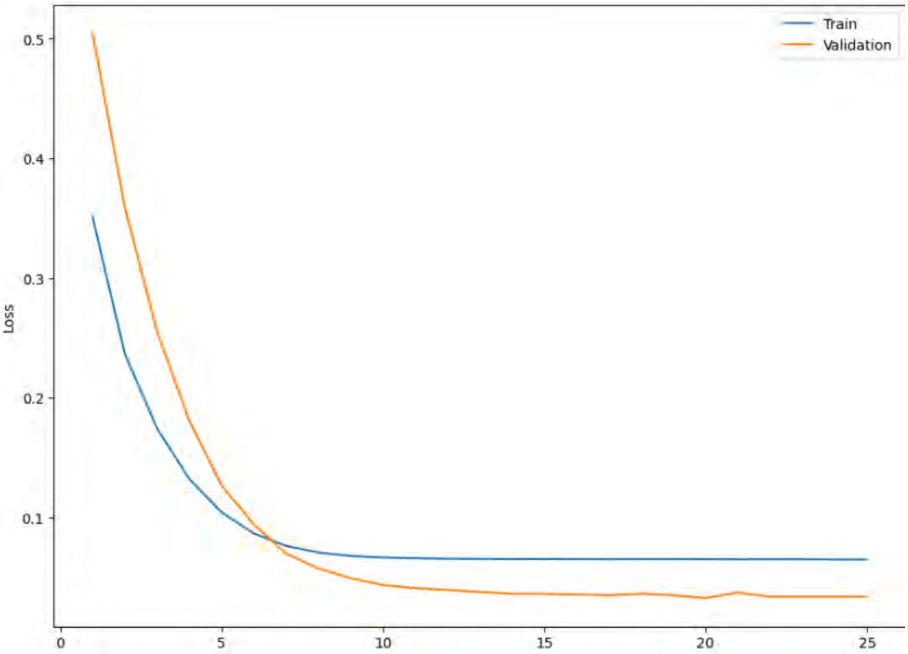


Figure 4.1: Training and Validation Loss Curve for ANN-based Fraud Detection

4.1.2 Random Forest (RF)

Using a random forest model, we successfully detected fraudulent credit card transactions with precision. The variables most associated with fraud, ranked from highest to lowest correlation, are V17, V14, V10, V12, and V11. But random forest achieved a cross-validated Accuracy score of 0.989.

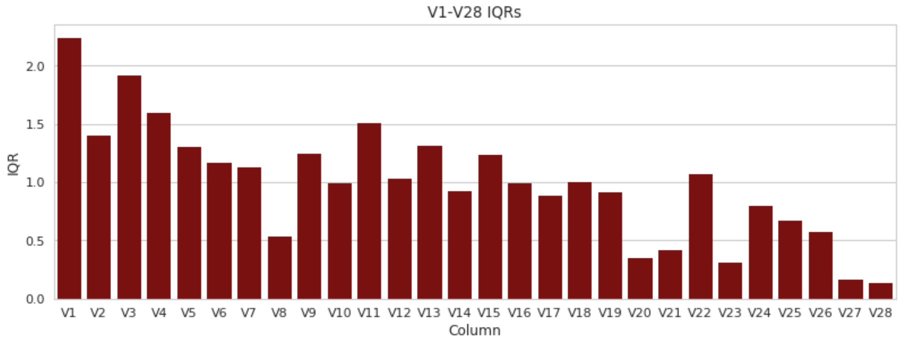


Figure 4.2: The IQRs of V1-V28 are on a similar scale as the standard deviations

1. **Divide the data with a random, stratified train/test split, where the test set accounts for 20% of the data.**
2. **Utilizing the power transform helps eliminate skewness in the data by modifying the transaction amounts.**
3. **The normalization of mean and variance for every feature within a machine learning process.**

4.1.3 Support Vector Machine (SVM)

We can see that the study using the reduced data is far from irrelevant, which means that the last step of the previously computed PCA could have been done in a more efficient way. Indeed one of the main question we have with the PCA once we calculated the principals components direction, is how many of this component are we gonna keep.This means that some of the 30 dimensions are do not discriminate classes that much. Eventually the accuracy came with almost like 94%. Firstly, we observed that the frequency of frauds is not affected by time. Additionally, most fraud cases involve small sums of money.

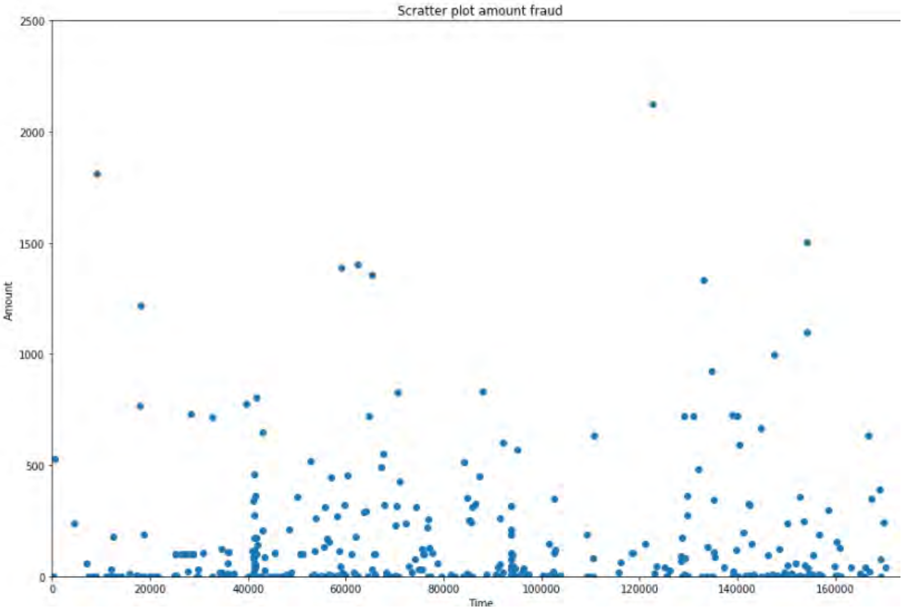


Figure 4.3: Scatter plot amount fraud through SVM

4.1.4 Extreme Gradient Boosting(XGBoost)

Here, XGBoost stands out as the best-performing model. From the above results our algorithm achieved auc-roc (i.e. area under the precision-recall curve) score of 0.979. It shows great accuracy and F1-score, showing its effectiveness in detecting true positives while reducing false positives. Despite achieving high accuracy, ANN's lower recall indicates it could potentially overlook certain positive instances. On the contrary, SVM has high sensitivity but lower specificity, which could result in inaccurate positive predictions. Random Forest provides an evenly distributed performance in terms of precision and recall, making it a viable option if both measures hold significance. Nevertheless, XGBoost has better accuracy and F1-score, making it the preferred choice in this particular situation. The model chosen should be based on the particular needs of the application. When reducing incorrect positive predictions is important, it is advised to use XGBoost. If the main focus is on capturing every positive case, SVM could be an option, but the precision must be assessed thoroughly to prevent too many false positives.

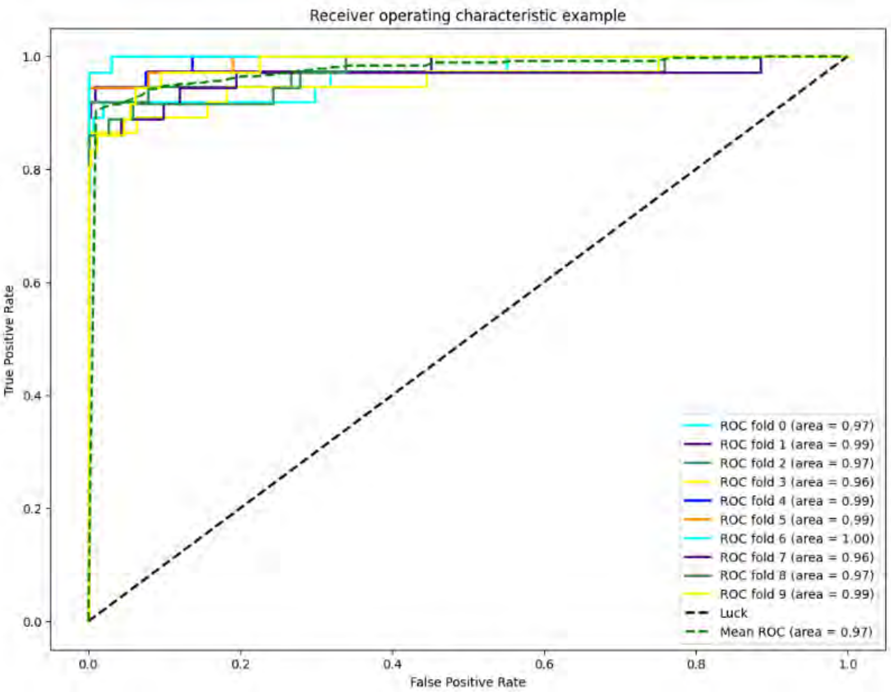
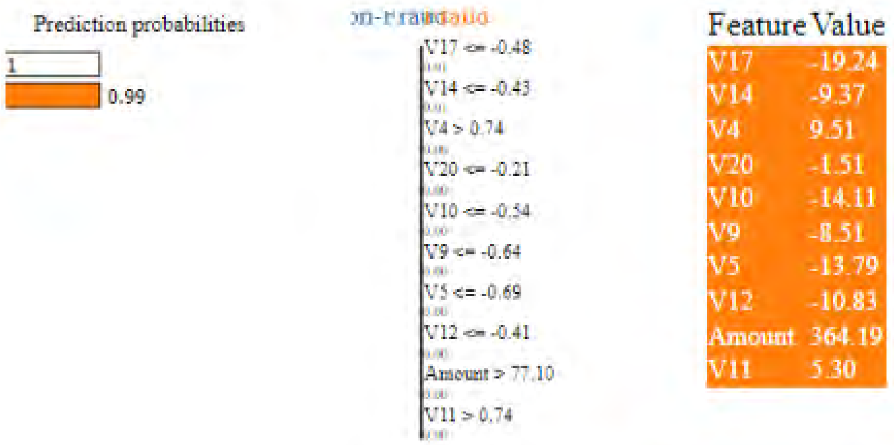


Figure 4.4: Classification through XGBoost

4.2 Result Discussion of Explainable AI (XAI) Models

Here the explanation can look like one fraudulent transaction using LIME. It is summarized like a diagram, showcasing the parameters that influence the model's prediction and the specific parameter indicating a fraudulent and non-fraudulent classification.

4.2.1 LIME



Feature	Value
V17	-19.24
V14	-9.37
V4	9.51
V20	-1.51
V10	-14.11
V9	-8.51
V5	-13.79
V12	-10.83
Amount	364.19
V11	5.30

Figure 4.5: Prediction explained by LIME

To the left of the figure, the probability of the particular transaction being fraudulent/non-fraudulent is shown. In the middle, the most important parameters for the transaction being fraudulent (orange) and not fraudulent (blue) are shown. The model predicts a 0.99 probability for the instance being fraudulent (orange bar). This is a strong indication that the transaction is classified as fraud. There is a 1% chance that the transaction is classified as non-fraud. This LIME result explains that the model heavily relied on the values of V17, V14, V4, V20, V10, V9, V5, V12, Amount, and V11 to predict that the transaction is fraudulent. The transaction was assigned a very high probability (0.99) of being fraudulent, largely influenced by extreme negative values in features like V17, V14, V10, and V5, along with the Amount being significantly high (364.19). This result demonstrates the explainability of the model's prediction, showing which features led to the decision, which is key when validating fraud detection models with transparency.

4.2.2 SHAP

It displays a SHAP generated explanation and offers a visualization of a global explanation. How the credit card fraud detection model operates generally regarding key parameters. The initial diagram presents a summary of the key components of the model. The hue mirrors the worth of the object. Features (red is high, blue is low). Illustrated explanations of the AI's predictions for a specific transaction. Red blocks indicate characteristics that raise the likelihood of a transaction being fraudulent, while Blue blocks signify characteristics that reduce the chances of fraud. The presence of blue blocks signals a decrease in the features. potential for deceit. The model confidently predicts fraud with a value of 0.99, while the baseline prediction is 0.001723 if no features are considered. Red features positively influenced the fraud prediction, as shown in the SHAP visualization breaking down the model's process to reach a 0.99 fraud probability. This SHAP explanation illustrates the importance of specific features like V17, V12, and V14 with significant negative values in predicting fraudulent transactions. It highlights the quantification and visualization of individual feature contributions for understanding the model's classification.

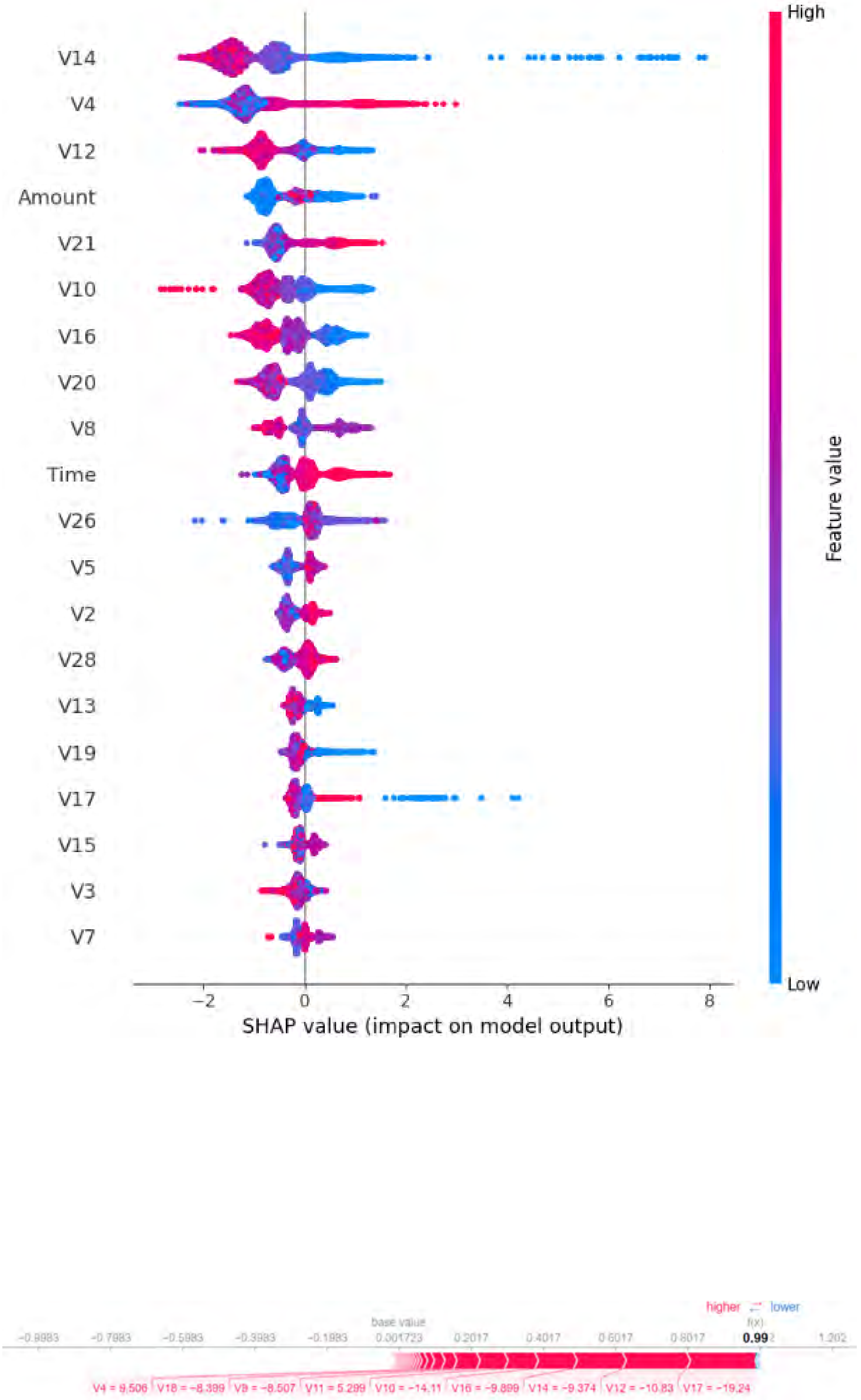


Figure 4.6: Prediction explained by SHAP globally

Chapter 5

Analysis

5.1 Discussion

In this part, from a different perspective of Machine Learning Methods, Explainable Artificial Intelligence methods as well as the user study design. There are certain limitations on this project, for future work, it is necessary to have more balanced and updated data. As this sector changes time to time and ethical aspects of the research are also discussed.

Machine Learning(ML) This thesis focuses on addressing imbalanced datasets in credit card fraud detection. Algorithm ANN results explained using SHAP and LIME due to a higher accuracy value. The study of ML methods, and more research is needed to compare ANN with other methods. Rerunning with additional ML methods could enhance outcomes.

Explainable AI(XAI) In this undertaking, we expect that exploration members will rate the logic given by LIME and SHAP clarifications higher than in situations where no obvious reasons are given. Utilizing LIME and SHAP further developed reasonableness, however just somewhat, as a matter of fact. A potential justification for this might be that the reasonableness of LIME and SHAP isn't completely reflected in the review. The explanation that we didn't do the meeting is that we don't possess the ability to arrive at the bank workers yet and we need more chances to contact the bank representatives.

By employing both LIME and SHAP, two complementary approaches to elucidate the model's decisions. Increasing trust on the model through:

1. **Offering straightforward clarifications (LIME).**
2. **Guaranteeing thorough, uniform justifications (SHAP).**

This mix can reduce the "black-box" issue commonly linked with machine learning models by guaranteeing that every decision made by the system can be explained, traced, and trusted.

LIME visually illustrates the exact route the model took for a particular transaction, indicating the key features that had the most impact. On the contrary, SHAP provides a more accurate feature explanation and guarantees that the weight and consistency of features are maintained in all predictions. Together, they offer a complete perspective on the functioning of the fraud detection model, guaranteeing stakeholders trust the predictions and comprehend why specific transactions are identified as fraud.

5.1.1 Limitations

We faced limits on time and resources in this study. The dataset used is synthetic because real credit card transaction data with detailed features is hard to find online due to confidentiality. To get a better dataset, cooperating with banks for real transaction data is suggested for future work. With bank cooperation, user studies can move from questionnaires to in-depth interviews, gathering more valuable feedback.

5.1.2 Future WorkPlan

Exploring additional eXplainable Artificial Intelligence (XAI) methods for clarity is also valuable. Gathering user feedback on preferred methods is essential, and the user study should be short but informative, including clear charts. Involving not only bank workers but also AI experts in the study can provide a wide range of data. Interviews for richer feedback and considering additional metrics like transparency are suggested for future work.

5.2 Conclusion

This project aimed to explore XAI methods for credit card fraud detection using a synthetic dataset due to confidentiality. The dataset size was limited by computing resources. ANN, KNN, SVM were chosen for their past performance, and oversampling addressed dataset imbalance. ANN slightly outperformed only in accuracy score.

ANN was selected for better performance and evaluated with LIME and SHAP explanations through a user study with bank workers in China. Understandability, satisfaction, sufficiency, and trustworthiness were measured. Overall, both explanations improved outcomes' explainability, with LIME scoring higher in understandability, sufficiency, and satisfaction, while SHAP scored higher in trustworthiness.

For future research, obtaining a real high-quality dataset through collaboration with banks is suggested. Improvements in questionnaire design and presentation, possibly with more visual elements, could enhance participation and understanding. Extended user studies, including interviews, could gather more valuable feedback through bank collaboration.

Bibliography

[1] L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.

[2] I. Shafi, J. Ahmad, S. I. Shah, and F. M. Kashif, "Impact of varying neurons and hidden layers in neural network architecture for a time frequency application," in *2006 IEEE International Multitopic Conference*, IEEE, 2006, pp. 188–193.

[3] C. A. Paasch, *Credit card fraud detection using artificial neural networks tuned by genetic algorithms.* Hong Kong University of Science and Technology (Hong Kong), 2008.

[4] C. Whitrow, D. J. Hand, P. Juszczak, D. Weston, and N. M. Adams, "Transaction aggregation as a strategy for credit card fraud detection," *Data mining and knowledge discovery*, vol. 18, pp. 30–55, 2009.

[5] P.-F. Pai, M.-F. Hsu, and M.-C. Wang, "A support vector machine-based model for detecting top management fraud," *Knowledge-Based Systems*, vol. 24, no. 2, pp. 314–321, 2011.

[6] M. Hejazi and Y. P. Singh, "Credit data fraud detection using kernel methods with support vector machine," *Journal of Advanced Computer Science and Technology Research*, vol. 2, no. 1, pp. 35–49, 2012.

[7] X. Wang, H. Shao, N. Japkowicz, *et al.*, "Using svm with adaptively asymmetric misclassification costs for mine-like objects detection," in *2012 11th International Conference on Machine Learning and Applications*, IEEE, vol. 2, 2012, pp. 78–82.

[8] H.-G. Han, L.-D. Wang, and J.-F. Qiao, "Efficient self-organizing multilayer neural network for nonlinear system modeling," *Neural Networks*, vol. 43, pp. 22–32, 2013.

[9] S. Amarappa and S. Sathyanarayana, "Data classification using support vector machine (svm), a simplified approach," *Int. J. Electron. Comput. Sci. Eng*, vol. 3, pp. 435–445, 2014.

[10] A. U. S. Khan, N. Akhtar, and M. N. Qureshi, "Real-time credit-card fraud detection using artificial neural network tuned by simulated annealing algorithm," in *Proceedings of international conference on recent trends in information, telecommunication and computing, ITC*, Citeseer, 2014, pp. 113–121.

[11] M. Zakariah *et al.*, "Classification of large datasets using random forest algorithm in various applications: Survey," *International Journal of Engineering and Innovative Technology (IJJEIT)*, vol. 4, no. 3, 2014.

[12] Y. Dai, J. Yan, X. Tang, H. Zhao, and M. Guo, "Online credit card fraud detection: A hybrid framework with big data technologies," in *2016 IEEE Trustcom/BigDataSE/ISPA*, IEEE, 2016, pp. 1644–1651.

[13] O. S. Yee, S. Sagadevan, and N. H. A. H. Malim, "Credit card fraud detection using machine learning as data mining technique," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 10, no. 1-4, pp. 23–27, 2018.

[14] I. Benchaji, S. Douzi, and B. El Ouahidi, "Using genetic algorithm to improve classification of imbalanced datasets for credit card fraud detection," in *Smart Data and Computational Intelligence: Proceedings of the International Conference on Advanced Information Technology, Services and Systems (AIT2S-18) Held on October 17–18, 2018 in Mohammedia 3*, Springer, 2019, pp. 220–229.

[15] D. Wang, Q. Yang, A. Abdul, and B. Y. Lim, "Designing theory-driven user-centric explainable ai," in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 2019, pp. 1–15.

[16] V. Balayan, "Human-interpretable explanations for black-box machine learning models: An application to fraud detection," M.S. thesis, Universidade NOVA de Lisboa (Portugal), 2020.

[17] E. Glikson and A. W. Woolley, "Human trust in artificial intelligence: Review of empirical research," *Academy of Management Annals*, vol. 14, no. 2, pp. 627–660, 2020.

[18] C. Molnar, G. König, J. Herbinger, *et al.*, "General pitfalls of model-agnostic interpretation methods for machine learning models," in *International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers*, Springer, 2020, pp. 39–68.

[19] A. Muaz, M. Jayabalan, and V. Thiruchelvam, "A comparison of data sampling techniques for credit card fraud detection," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 6, 2020.

[20] N. Rtayli and N. Enneya, "Enhanced credit card fraud detection based on svm-recursive feature elimination and hyper-parameters optimization," *Journal of Information Security and Applications*, vol. 55, p. 102 596, 2020.

[21] H. Tingfei, C. Guangquan, and H. Kuihua, "Using variational auto encoding in credit card fraud detection," *IEEE Access*, vol. 8, pp. 149 841–149 853, 2020.

[22] D. Virmani and N. Kaushik, "M, mathur v, saxena s (2020) analysis of cyber attacks and security intelligence: Identity theft," *Indian Journal of Science and Technology*, vol. 13, no. 25, pp. 2529–2536, 2020.

[23] L. Antwarg, R. M. Miller, B. Shapira, and L. Rokach, "Explaining anomalies detected by autoencoders using shapley additive explanations," *Expert systems with applications*, vol. 186, p. 115 736, 2021.

[24] J. I.-Z. Chen and K.-L. Lai, "Deep convolution neural network model for credit-card fraud detection and alert," *Journal of Artificial Intelligence*, vol. 3, no. 02, pp. 101–112, 2021.

[25] T. K. Dang, T. C. Tran, L. M. Tuan, and M. V. Tiep, "Machine learning based on resampling approaches and deep reinforcement learning for credit card fraud detection systems," *Applied Sciences*, vol. 11, no. 21, p. 10 004, 2021.

[26] Y. Ji, *Explainable ai methods for credit card fraud detection: Evaluation of lime and shap through a user study*, 2021.

[27] D. V. Kute, B. Pradhan, N. Shukla, and A. Alamri, "Deep learning and explainable artificial intelligence techniques applied for detecting money laundering–a critical review," *IEEE access*, vol. 9, pp. 82 300–82 317, 2021.

[28] H. Sahlaoui, A. Nayyar, S. Agoujil, M. M. Jaber, *et al.*, "Predicting and interpreting student performance using ensemble models and shapley additive explanations," *IEEE Access*, vol. 9, pp. 152 688–152 703, 2021.

[29] O. Voican, "Credit card fraud detection using deep learning techniques.," *Informatica Economica*, vol. 25, no. 1, 2021.

[30] T.-Y. Wu and Y.-T. Wang, "Locally interpretable one-class anomaly detection for credit card fraud detection," in *2021 International Conference on Technologies and Applications of Artificial Intelligence (TAAI)*, IEEE, 2021, pp. 25–30.

[31] J. Chaquet-Ulldemolins, F.-J. Gimeno-Blanes, S. Moral-Rubio, S. Muñoz-Romero, and J.-L. Rojo-Álvarez, "On the black-box challenge for fraud detection using machine learning (ii): Nonlinear analysis through interpretable autoencoders," *Applied Sciences*, vol. 12, no. 8, p. 3856, 2022.

[32] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud: A review of anomaly detection techniques and recent advances," *Expert systems With applications*, vol. 193, p. 116 429, 2022.

[33] J. Zhou, S. Huang, T. Zhou, D. J. Armaghani, and Y. Qiu, "Employing a genetic algorithm and grey wolf optimizer for optimizing rf models to evaluate soil liquefaction potential," *Artificial intelligence review*, vol. 55, no. 7, pp. 5673–5705, 2022.

[34] E. R. Mill, W. Garn, N. F. Ryman-Tubb, and C. Turner, "Opportunities in real time fraud detection: An explainable artificial intelligence (xai) research agenda," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 5, pp. 1172–1186, 2023.

[35] T. L. Serongwa, "Credit card fraud detection system using extreme gradient boosting machine and isolated forest," Ph.D. dissertation, University of the Witwatersrand, Johannesburg, 2023.

[36] P. Agrawal, R. Gnanaprakash, and S. H. Dhawane, "Prediction of biodiesel yield employing machine learning: Interpretability analysis via shapley additive explanations," *Fuel*, vol. 359, p. 130 516, 2024.

[37] V. Chang, B. Ali, L. Golightly, M. A. Ganatra, and M. Mohamed, "Investigating credit card payment fraud with detection methods using advanced machine learning," *Information*, vol. 15, no. 8, p. 478, 2024.

[38] M. T. Islam, K. Ashraf, M. H. Hosen, S. Nawar, and S. Asgar, “Predictive modeling of anxiety levels in bangladeshi university students: A voting-based approach with lime and shap explanations,” in *2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems (iCACCESS)*, IEEE, 2024, pp. 01–06.

[39] H.-k. Kwon, “Recent transformation of the us economy,” in *Openness and Coordination: National Economies of the US, Japan, and Germany in a Globalized World*, Springer, 2024, pp. 79–117.

[40] F. M. Talaat, A. Aljadani, M. Badawy, and M. Elhosseini, “Toward interpretable credit scoring: Integrating explainable artificial intelligence with deep learning for credit card default prediction,” *Neural Computing and Applications*, vol. 36, no. 9, pp. 4847–4865, 2024.

[41] Y. Zhang, B. Suleiman, M. J. Alibasa, and F. Farid, “Privacy-aware anomaly detection in iot environments using fedgroup: A group-based federated learning approach,” *Journal of Network and Systems Management*, vol. 32, no. 1, p. 20, 2024.

[42] R. Udayakumar, A. Joshi, S. Boomiga, and R. Sugumar, “Deep fraud net: A deep learning approach for cyber security and financial fraud detection and classification,”