A DEEP LEARNING ALGORITHM IMPLEMENTATION FOR CREDIT CARD FRAUD ANALYSIS AND DETECTION

A PROJECT REPORT

Submitted by

DHAVAMANI V **512220106001**

KUMAR S **512220106002**

PRAGADEESWARAN K **512220106005**

In the partial fulfillment for the award of the degree

Of

BACHELOR OF ENGINEERING

IN

ELECTRONICS AND COMMUNICATION ENGINEERING

S.K.P ENGINEERING COLLEGE

TIRUVANNAMALAI - 606 611

ANNA UNIVERSITY: CHENNAI 600 025

MAY 2024

BONAFIDE CERTIFICATE

Certified that this project report **"A DEEP LEARNING ALGORITHM IMPLEMENTATION FOR CREDIT CARD FRAUD ANALYSIS AND DETECTION" is** the Bonafide work of "**V. DHAVAMANI, S. KUMAR, K. PRAGADEESWARAN"** who carried out the project work under my supervision.

SIGNATURE

Dr.E.RAJENDRAN, B.E, M.Tech.,Ph.D.,

HEAD OF THE DEPARTMENT,

DEPARTMENT OF ECE,

SKP ENGINEERING COLLEGE,

TIRUVANNAMALAI.

SIGNATURE

Dr. S. BASKARAN, ME.,Ph.D.,

PRINCIPAL / SUPERVISIOR,

DEPARTMENT OF ECE,

SKP ENGINEERING COLLEGE,

TIRUVANNAMALAI.

Submitted for the anna university project work viva-voce held

on _____

INTERNAL EXAMINER

EXTERNAL EXAMINER

ACKNOWLEDGEMENT

The successful completion of any task would be incomplete without meaning the people who made it possible and whose constant guidance and encouragement secured us this success.

We express our sincere thanks and heartiest gratitude to our beloved Chairman Thiru. **K. KARUNANITHI, B.E., MBA.,** and our beloved Principal **Dr. S. BASKARAN, M. E, Ph.D.** and who gave us permission and encouraged us throughout the course to pursue the new goals and ideas in doing our project.

No words of gratitude will suffice for the unquestioning support extended to us by our Head of the department **Dr.E.RAJENDRAN, B.E,M.Tech,Ph.D.** for being ever supporting force during our project work.

We find no words to express our immense pleasure in thanking us Internal guide **Dr. S. BASKARAN, M.E, Ph.D.** for his valuable training and guidance for the project. His dedicated guidance made us to complete this project successfully.

Also, we express our thanks to all the Staff members and Lab assistant of Our department provides necessary information.

Finally, we extend our sincere thanks to our parents and friends who act as backbone throughout the project.

ABSTRACT

In today's world, high dependency on internet technology has enjoyed increased credit card transactions, but credit card fraud has also accelerated as an online and offline transaction. Financial fraud is a growing concern with far-reaching consequences for the government, corporate organizations, and the finance industry. Hence, this project proposes A Deep Learning Algorithm Implementation for Credit Card Fraud Analysis and Detection.

This work was developed by the Python Jupyter software. Initially, the input dataset is initiated by a preprocessing technique. This process is handled by data cleaning, which helps clean the datasets and, additionally, handles the missing values. Under pre-processing, the process started with data visualization process. Which process is used to clearly show the patterns in the dataset and improve integrity. Next, the data splitting process handles the data set and divides it for the purpose of the regression process. Deep learning algorithms are used in this work. The deep learning technique that handles the MLP algorithm is to predict fraudulent transactions and normal transactions. The final step is predicting whether the output of identification for fraudulent transactions is achieved in the model evaluation process.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

MLP	-	Multilayer Perceptron
ROC	-	Receiver Operating Characteristic
GUI	-	Graphical User Interface
DL	-	Deep Learning
MCC	-	Merchant Category Code
NRI	-	Never Received Issue
IG	-	Information Gain
LSTM	-	Long Short-Term Memory
CCFD	-	Carbon Contracts for Difference
CNP	-	Card-Not-Present
CP	-	Card-Present
GAN	-	Generative Adversarial Network
VAE	-	Variational Autoencoder
PCA	-	Principal Component Analysis
GRU	-	Gated Recurrent Unit
SMOTE	-	Synthetic Minority Oversampling Technique
ENN	-	Edited Nearest Neighbor
RNN	-	Recurrent Neural Network
ML	-	Machine Learning
ESMOTE	-	Ensemble Synthesized Minority Oversampling techniques
AUC	-	Area under the Curve
POS	-	Point-Of-Sale
CVM	-	Card Verification Method
AVS	-	Address Verification Systems
PIN	-	Personal Identification Number
FDS	-	Fraud Detection System
CNN	-	Convolutional Neural Network
PAN	-	Primary Account Number
ANN	-	Adaptive Neural Networks

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

Credit card fraud is a huge ranging term for theft and fraud committed using or involving at the time of payment by using this card. The purpose is to purchase goods without paying, or to transfer unauthorized funds from an account. A credit card that has been stolen, lost, or counterfeited might result in fraud. Card-not-present fraud, or the use of your credit card number in e-commerce transactions has also become increasingly common as a result of the increase in online shopping. Increased fraud, such as CCF, has resulted from the expansion of e-banking and several online payment environments, resulting in annual losses of billions of dollars. In this era of digital payments, CCF detection has become one of the most important goals. As a business owner, it cannot be disputed that the future is heading towards a cashless culture. As a result, typical payment methods will no longer be used in the future, and therefore it will not be helpful for expanding a business. Customers will not always visit the business with cash in their pockets. It is now placing a premium on debit and credit card payments. As a result, companies will need to update their environment to ensure that it can take all types of payments. In the next years, this situation is expected to become much more severe. Credit card fraud is added to identity theft. As per the information from the United States Federal Trade Commission, the theft rate of identity had been holding stable during the mid-2000s, but it was increased by 21 percent in 2008. Even though credit card fraud, that crime which most people associate with ID theft, decreased as a percentage of all ID theft complaints in 2000, out of 13 billion transactions made annually, approximately 10 million or one out of every 1300 transactions turned out to be fraudulent. Also, 0.05% (5 out of every 10,000) of all monthly active

accounts was fraudulent. Today, fraud detection systems are introduced to control one-twelfth of one percent of all transactions processed which still translates into billions of dollars in losses. Credit Card Fraud is one of the biggest threats to business establishments today. However, to combat fraud effectively, it is important to first understand the mechanisms of executing a fraud. Credit card fraudsters employ a large number of ways to commit fraud. In simple terms, Credit Card Fraud is defined as "when an individual uses another individuals' credit card for personal reasons while the owner of the card and the card issuer are not aware of the fact that the card is being used". Card fraud begins either with the theft of the physical card or with the important data associated with the account, including the card account number or other information that necessarily be available to a merchant during a permissible transaction. Card numbers, generally the Primary Account Number (PAN), are often reprinted on the card, and a magnetic stripe on the back contains the data in machine-readable format. It contains the following Fields:

- Name of card holder
- Card number
- Expiration date
- Verification/CVV code
- Type of card

There are more methods to commit credit card fraud. Fraudsters are very talented and fast-moving people. In the Traditional approach, to be identified is application Fraud, where a person will give the wrong information about himself to get a credit card. There is also the unauthorized use of Lost and Stolen Cards, which makes up a significant area of credit card fraud. There are more enlightened credit card fraudsters, starting with those who produce Fake and Doctored Cards; there are also those who use Skimming to commit fraud. It will get this information held on either the

magnetic strip on the back of the credit card, or the data stored on the smart chip is copied from one card to another. Site Cloning and False Merchant Sites on the Internet are becoming a popular method of fraud for many criminals with a skilled ability for hacking. Such sites are developed to get people to hand over their credit card details without knowing it has been swindled.

1.2 CREDIT CARD FRAUD

Figure 1.1 Credit Card Fraud

Credit Card Misrepresentation is one of the greatest dangers to business and business foundations today. Just, Master card Misrepresentation is characterized as, "when an individual uses another individual‟ Visa for individual utilize while the proprietor of the card and also the card backer don't know about the thing that the card is being utilized." various frameworks/models, process and preventive measures will stop Master card extortion and lessen monetary dangers. Banks and Visa organizations have accumulated a lot of Master card account exchanges. It uses a number of deep learning algorithms for detecting CCF. However, in this study, we choose the CNN model and its layers to determine if the original fraud is the normal transaction of qualified datasets. Some transactions are common in datasets that have been labelled fraudulent and demonstrate questionable transaction behavior. The Charge card is a plastic

card issued to a number of clients as one of the methods of instalment. It enables cardholders to buy merchandise and enterprises in light of the cardholder's guarantee.

1.2.1 DIFFERENT TYPES OF CREDIT CARD FRAUD

Different Types of Credit Card Fraud Misrepresentation discovery frameworks come into situation when the fraudsters surpass the extortion aversion frameworks and begin false exchanges. Alongside the advancements in Data Innovation and upgrades in the correspondence channels, misrepresentation is spreading everywhere throughout the world with aftereffects of vast measures of false misfortune. Anderson (2007) has recognized and depicted the distinctive sorts of extortion. Charge card cheats can continue in a wide range of courses, for example, basic burglary, fake cards, Never Got Issue (NRI), application misrepresentation and on the web/Electronic extortion (where the card holder is absent). Master card misrepresentation recognition is appallingly troublesome, yet in addition regular issue for arrangement.

Card Not Present Transaction (CNP)

Figure 1.2 Card Not Present Transaction

In the event that a card isn't physically present when a client makes a buy, the shipper must depend on the cardholder, or somebody indicating to

be thus, introducing card data in a roundabout way, regardless of whether via mail, phone or over the Internet.

1.2.2 IDENTIFY THEFT

The Identity Theft is divided into two classifications:

Application fraud

Application fraud happens when a man utilizes stolen or counterfeit archives to open a record in someone else's name. Culprits take records, for example, service bills and bank explanations to develop helpful individual data

Account takeover

This theft happens when an illegal poses as an intelligent customer, gains control of an account and then makes unofficial transactions

1.2.3 SKIMMING

An electronic strategy for catching a casualty's close to home data utilized by personality criminals. The skimmer is a little gadget that outputs a Visa and stores the data contained in the attractive strip. Skimming can occur amid a true-blue exchange at a business.

1.3 CHALLENGES IN CREDIT CARD FRAUD DETECTION

The challenge would be to acknowledge fraudulent transactions such that merchant accounts 'customers' are not charged for transactions that it didn't perform. The infield of credit card fraud detection several challenges still needs to address. Some of them are covering here.

Major challenges of identifying credit card fraud seem to be:

Enormous information

Enormous information has been stored on even a daily basis as well as the design construction should be quick enough yet to react properly to a fraud.

Imbalanced information

Imbalanced information, i.e. its majority of transactions (98.9%) is also not fraud, making it impossible to identify fraudulent activity.

Misclassified information

Misclassified information might be another major issue because not all fraudulent activity is captured or recorded.

Adaptive methods

Adaptive methods used among fraudsters against the system.

1.4 THE PROBLEM OF FRAUD DETECTION

Fraud is as old as humanity itself and can take an unlimited variety of different forms. Moreover, the development of new technologies provides additional ways in which criminals commit fraud, for instance in e-commerce the information about the card is sufficient to perpetrate a fraud. The use of credit cards is prevalent in modern day society and credit card fraud has kept on growing in recent years. Financial losses due to fraud affect not only merchants and banks (e.g. reimbursements), but also individual clients. If the bank loses money, customers eventually pay as well through higher interest rates, higher membership fees, etc. Fraud also affects the reputation and image of a merchant causing non-financial losses that, though difficult to quantify in the short term, become visible in the long period. For example, if a cardholder is a victim of fraud with a certain company, they no longer trust their business and choose a competitor. The actions taken against

fraud can be divided into fraud prevention, which attempts to block fraudulent transactions at source, and fraud detection, where successful fraud transactions are identified a posteriori. Technologies that have been used in order to prevent fraud are Address Verification Systems (AVS), Card Verification Method (CVM) and Personal Identification Number (PIN). AVS involves verification of the address with zip code of the customer while CVM and PIN involve checking of the numeric code that is keyed in by the customer. For prevention purposes, financial institutions challenge all transactions with rule-based filters and data mining methods as neural networks. Fraud detection is, given a set of credit card transactions, the process of identifying if a new authorized transaction belongs to the class of fraudulent or genuine transactions. A Fraud Detection System (FDS) should not only detect fraud cases efficiently, but also be cost-effective in the sense that the cost invested in transaction screening should not be higher than the loss due to frauds.

In order to minimize costs of detection it is important to use expert rules and statistical-based models (e.g. Machine Learning) to make a first screen between genuine and potential fraud and ask the investigators to review only the cases with high risk. Typically, transactions are first filtered by checking some essential conditions (e.g. sufficient balance) and then scored by a predictive model. The predictive model scores each transaction with high or low risk of fraud and those with high risk generate alerts. Investigators check these alerts and provide feedback for each alert, i.e. true positive (fraud) or false positive (genuine). This feedback can then be used to improve the model. A predictive model can be built upon experts' rules, i.e. rules based on knowledge from fraud experts, but these require manual tuning and human supervision. Alternatively, with Machine Learning (ML) techniques. Efficiently discover fraudulent patterns and predict transactions that are most likely to be fraudulent. ML techniques consist in inferring a

prediction model on the basis of a set of examples. The model is in most cases a parametric function, which allows predicting the likelihood of a transaction to be fraud, given a set of features describing the transaction. In the domain of fraud detection, the use of learning techniques is attractive for a number of reasons.

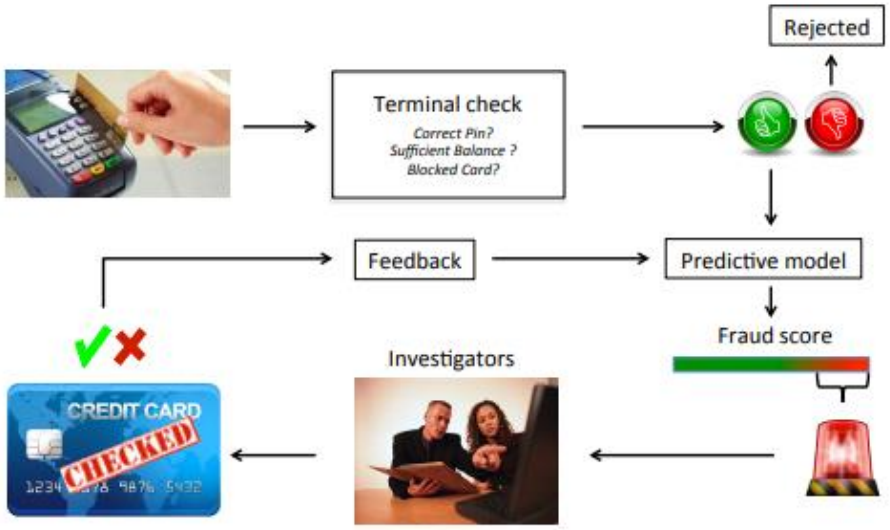

Figure 1.3 The Credit Card Fraud Detection process

Credit card frauds occur in various ways: just to mention some, can have stolen card fraud, cardholder-not-present fraud and application fraud:

Stolen card fraud I

It is the most common type of fraud where the fraudster usually tries to spend as much as possible and as quickly as possible. The detection of such a fraud typically relies on the discovery of an unexpected usage pattern of the credit card (generally unexpectedly important) with respect to the common practice.

Cardholder-not-present fraud I

It is often observed in e-business. Here the fraudster needs the information about a credit card but not the card itself. This fraud demands

prompt detection since, unlike the previous case, the official card owner is not aware that his own data has been stolen.

Application fraud

Corresponds to the application for a credit card with false personal information. This kind of fraud occurs more rarely since it could be detected during the application by checking the information of the applier, contrary to other frauds that cannot be anticipated.

1.5 BEHAVIORAL ANALYTICS

Figure 1.4 Behavioral Analytics

Banks analyze customer behavior and spending patterns to create profiles of normal card usage. Any deviation from the established patterns can trigger alerts for potential fraud. For example, if a customer suddenly starts making multiple high-value transactions, it indicates fraudulent activity. Advanced fraud detection systems leverage AI and machine learning algorithms to analyze large volumes of transaction data. These systems learn from historical data to identify patterns and anomalies associated with fraudulent transactions. As new fraud techniques emerge, the algorithms adapt and update their detection capabilities.

1.6 FRAUD SCORING

Financial institutions assign a risk score to each transaction based on various factors such as transaction amount, location, merchant reputation, and customer behavior. Transactions with high-risk scores are flagged for manual review or additional authentication steps.

Figure 1.5 Fraud Scoring

1.7 CARD BLOCKING AND ALERTS

Banks utilize real-time monitoring to identify potentially fraudulent transactions. In some cases, it blocks or temporarily freezes a card if suspicious activity is detected. Additionally, customers can receive real-time alerts via SMS, email, or mobile apps for every transaction made with their debit card. This allows them to quickly identify unauthorized activities and report them to the bank. It's important to note that while these methods are effective in detecting and preventing debit card fraud, no system is foolproof. It's always advisable for customers to exercise caution, regularly monitor their account statements, and report any suspicious activity to their bank immediately.

Figure 1.6 Card Blocking and Alert

1.8 CREDIT CARD WORKING PROCESS

The working process of a credit card involves several steps, from application to transaction processing.

Figure 1.7 Credit Card Working Process

Here's a general overview of how a credit card works:

Application

The first step is applying for a credit card. Individuals typically submit an application to a bank or credit card issuer. The application includes personal information, income details, and credit history.

Credit Card Approval

The bank or credit card issuer reviews the application and assesses the applicant's creditworthiness. It evaluates factors such as credit score, income, and existing debt. If approved, the applicant receives the credit card.

Credit Limit

The credit card comes with a predefined credit limit, which is the maximum amount the cardholder can borrow or spend using the card. The credit limit is determined by the issuing bank based on the applicant's creditworthiness.

Card Activation

The cardholder needs to activate the credit card before it can be used. This is usually done by calling a specific phone number or through an online activation process provided by the issuer.

Cardholder Verification

When making a purchase, the cardholder presents the credit card to the merchant. The merchant verifies the card's authenticity by checking the hologram, card number, and the cardholder's signature (in the case of a physical card).

Transaction Authorization

Once the merchant confirms the card's validity, it initiates the transaction by swiping or inserting the card into a point-of-sale (POS) terminal or by entering the card details in an online payment gateway. The transaction details, including the purchase amount, merchant ID, and card information, are sent to the card issuer for authorization.

Authorization Request

The merchant's payment processor sends an authorization request to the card issuer, asking for approval to proceed with the transaction. The issuer evaluates the request based on the cardholder's available credit, credit limit, and any other risk factors.

Authorization Response

The card issuer reviews the authorization request and sends a response back to the merchant's payment processor. The response can be either approval or decline. If approved, the issuer reserves the authorized amount from the cardholder's available credit.

Transaction Settlement

After receiving the authorization approval, the merchant completes the sale. The transaction is settled, which involves transferring the funds from the cardholder's credit line to the merchant's account. Settlement typically occurs at the end of the day or in a batch process.

Credit Card Statement

The card issuer generates a monthly statement that details all the transactions made with the credit card during the billing cycle. The statement includes the transaction amounts, dates, merchant names, and other relevant information. It also indicates the minimum payment due and the payment due date.

Repayment

The cardholder is responsible for repaying the credit card issuer for the purchases made. The cardholder can choose to pay the full outstanding balance or make a minimum payment, which accrues interest on the remaining balance. If the cardholder does not pay the full balance, interest charges are applied to the remaining amount.

Credit Card Fees

Credit cards come with various fees, such as annual fees, late payment fees, cash advance fees, and foreign transaction fees. These fees are outlined in the credit card agreement and can vary depending on the card issuer and type of card.

It's important for credit card users to manage their credit responsibly, make timely payments, and keep track of their spending to avoid accumulating excessive debt or falling victim to fraud.

1.9 CREDIT CARD PROCESS

The credit card process involves several steps, from application to transaction processing. Here's a general overview of how the credit card process works:

Application

The first step is to apply for a credit card. Individuals can apply online, through the mail, or by visiting a bank or credit card issuer's branch. The application typically requires personal information, such as name, address, income details, and employment information.

Creditworthiness Assessment

Figure 1.8 Creditworthiness Assessment

The credit card issuer reviews the application and assesses the applicant's creditworthiness. This involves checking the applicant's credit history, credit score, income level, and debt-to-income ratio. The issuer determines whether to approve or decline the application based on this assessment.

Credit Limit Determination

If the application is approved, the credit card issuer assigns a credit limit to the cardholder. The credit limit is the maximum amount the cardholder can borrow or spend using the credit card. The assigned limit is based on factors such as the applicant's creditworthiness, income, and the issuer's policies.

Figure 1.9 Credit limit determination

Card Issuance

Once the application is approved, the credit card issuer produces and sends the physical credit card to the cardholder's registered address. The card typically includes the cardholder's name, the credit card number, the expiration date, and the card's security code (CVV/CVC).

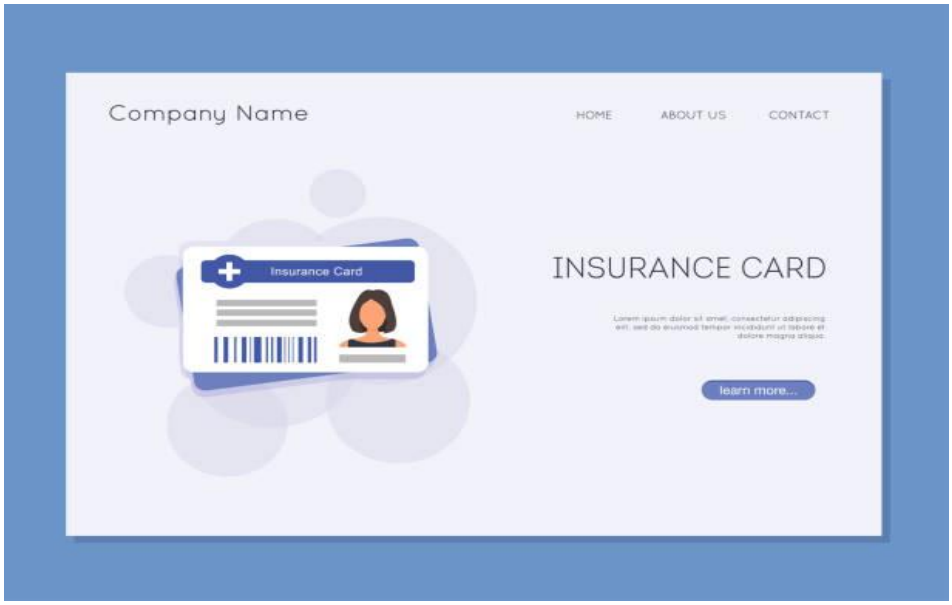

Figure 1.10 Credit Card Insurance

Card Activation

Upon receiving the credit card, the cardholder needs to activate it before it can be used for transactions. Activation is typically done by calling a specific phone number or through an online activation process provided by the issuer. The cardholder needs to provide personal information and the card details for verification.

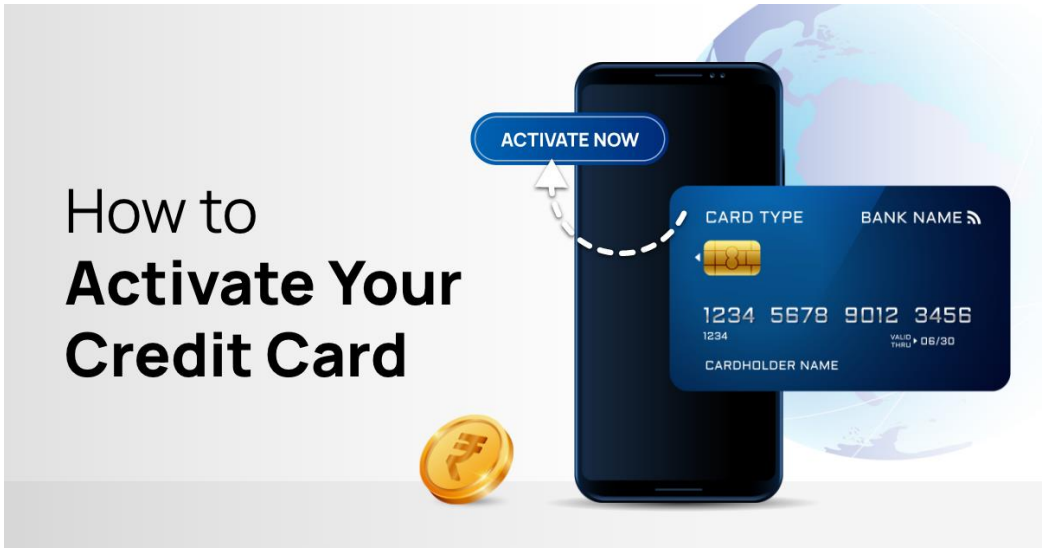

Figure 1.11 Credit Card Activation

Card Usage

Once the credit card is activated, the cardholder can use it to make purchases or obtain cash advances. The card can be swiped, inserted into a point-of-sale (POS) terminal, or used for online transactions. The cardholder needs to provide the necessary card details, including the card number, expiration date, and security code, for each transaction.

Figure 1.12 Credit Card usage

Transaction Processing

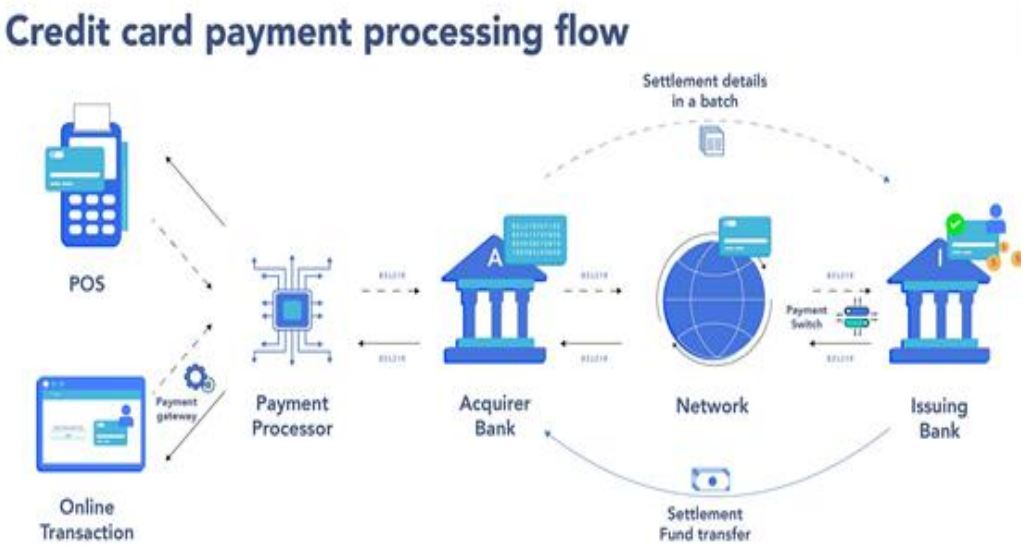

Figure 1.13 Transaction Processing

When the cardholder makes a purchase, the merchant initiates the transaction by sending the transaction details to a payment processor or acquiring bank. The payment processor verifies the transaction and sends an authorization request to the credit card issuer.

Authorization

The credit card issuer receives the authorization request and verifies the transaction details, including the cardholder's available credit, credit limit, and any other risk factors. Based on this verification, the issuer sends an authorization response to the payment processor, approving or declining the transaction.

Transaction Settlement

If the transaction is approved, the payment processor settles the transaction, and the funds are transferred from the issuer to the merchant's account. Settlement typically occurs within a few business days, depending on the payment processor and the merchant's agreement.

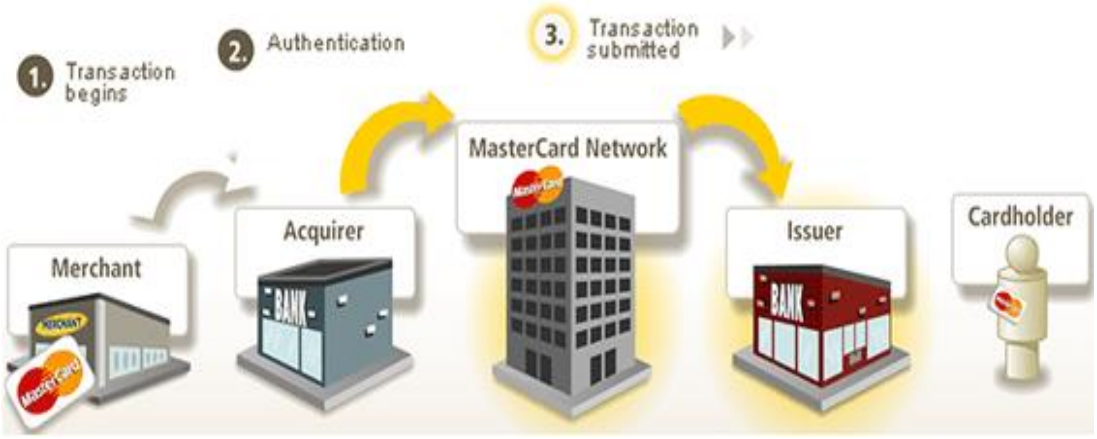

Figure 1.14 Transaction Settlement

Billing Cycle and Statement

The credit card issuer sets a billing cycle, usually monthly, during which the cardholder's transactions are compiled. At the end of the billing cycle, the issuer generates a credit card statement, which provides a summary of all the transactions made during that period. The statement includes the transaction details, dates, merchant names, and the total amount owed.

Repayment

The cardholder is required to make payments to the credit card issuer to settle the outstanding balance. The issuer provides a minimum payment

amount, which is a percentage of the total balance due. However, paying only the minimum payment incurs interest charges on the remaining balance. To avoid interest charges, the cardholder can choose to pay the full balance by the payment due date.

Fees and Interest Charges

Credit cards have associated fees, such as an annual fee, late payment fees, cash advance fees, foreign transaction fees, or balance transfer fees. Additionally, if the cardholder carries a balance beyond the grace period, interest charges are applied to the remaining amount.

1.10 ADVANTAGES

- More accurate.
- Effective Technique.
- Reduce financial losses.
- Improve compliance.
- Increase customer trust and loyalty.

1.11 APPLICATIONS

- Object detection image.
- Immune response abnormalities.
- Classification.

1.12 OBJECTIVE

- To predict whether a credit card transaction is fraudulent or not.
- To efficiently classify credit card frauds accurately using a multilayer perceptron (MLP) classifier.
- To evaluate the precision and recall value by using a deep learning classification model.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

Now these days digital, statistics are very easily available throughout the world because of digital online availability. All the information that also has a large volume, wide range, frequency, as well as importance is stored from small to large organizations over the cloud. The whole information is available from massive amounts of sources such as followers on social media, customer order behaviours, likes, and shares. White-collar crime is the ever-increasing problem with-reaching consequences for the finance sector, business institutions as well as governments. Fraud can indeed be described as illegal deceit to gain financial benefit. Enhanced card transactions had already appreciated a heavy emphasis on communication technology. When credit card transactions are by far the most prevalent form of transaction for offline and online payments, raising the rate of card fraud accelerates as well. Machine learning is the innovation of this century that eliminates conventional strategies and also can function on huge datasets where humans can't immediately access. Strategies of machine learning break within two important categories; supervised learning versus unsupervised learning; Tracking of fraud can also be achieved any form and only be determined how to use as per the datasets. Supervised training includes anomalies to always be identified as before. Many supervised methods are being used over the last few decades to identify credit card fraud. The major obstacle in implementing ML for detecting fraud seems to be the presence of extremely imbalanced databases. Most payments are legitimate in several available evidence sets, with such an extremely small number of fraudulent ones. The significant challenges to investigators are designing the accurate as well as efficient fraud prevention framework that will be low on false positives but efficiently identifies fraud activity.

Fawaz Khaled Alarfaj *et al* [2022] proposed in People can use credit cards for online transactions as it provides an efficient and easy-to-use facility. With the increase in usage of credit cards, the capacity of credit card misuse has also enhanced. Credit card frauds cause significant financial losses for both credit card holders and financial companies. In this research study, the main aim is to detect such frauds, including the accessibility of public data, high-class imbalance data, the changes in fraud nature, and high rates of false alarm. The relevant literature presents many machines learning based approaches for credit card detection, such as Extreme Learning Method, Decision Tree, Random Forest, Support Vector Machine, Logistic Regression and XG Boost. In addition, have performed experiments by balancing the data and applying deep learning algorithms to minimize the false negative rate. The proposed approaches can be implemented effectively for the real-world detection of credit card fraud. A slight value overhead the higher bound value or considers the threshold an anomaly. The most popular models in use. Using the algorithms, all these models built workably provided by the sci-kit-learn package. It beneficial when constructing a sequential model incrementally to show the summary of the model thus far with the current output. Strongly emphasise the research that reported fraud detection in the problem of class imbalance. Many techniques are used to detect credit cards. Traditional algorithms are not very well suited for large datasets. A CNN is a DL method; it can deeply relate to three-dimensional data, such as image processing. The CNN does not require heavy data pre-processing for training. In a real-world transaction, fraudulent and no fraudulent classes are not balanced due to the nature of the problem.

Emmanuel Ileberi *et al* [2021] described and the advance in technologies such as e-commerce and financial technology (FinTech) applications have sparked an increase in the number of online card transactions that occur on a daily basis. As a result, there has been a spike in

credit card fraud that affects card issuing companies, merchants, and banks. It is therefore essential to develop mechanisms that ensure the security and integrity of credit card transactions. In this research, implement a machine learning (ML) based framework for credit card fraud detection using a real-world imbalanced dataset that were generated from European credit cardholders. The models were evaluated using the accuracy, the recall, the precision, the Matthews Correlation Coefficient (MCC), and the Area under the Curve (AUC). Moreover, the proposed framework was implemented on a highly skewed synthetic credit card fraud dataset to further validate the results that were obtained in this research. The attributes V1 to V28 do not have specific feature names due to data security and integrity reasons. The SMOTE method generates synthetic data points that are not a direct replica of the minority class instance. RC metrics is not enough to assess the performance of our method it made quite a few errors in predicting fraudulent transactions. The ELM is less complex in comparison to the MLP., the impact of credit card fraud affects institutions such as card issuers, merchants, and small businesses. Investigate the AdaBoost to increase the quality of classification on a highly skewed credit card fraud dataset. The major contribution of this research work can be summarized., the authors implemented an under-sampling technique to solve the issue of class imbalance that exist in the dataset that was used.

Fuad A. Ghaleb *et al* [2023] brief that recent increase in credit card fraud is rapidly has caused huge monetary losses for individuals and financial institutions. Most credit card frauds are conducted online by illegally obtaining payment credentials through data breaches, phishing, or scamming. Many solutions have been suggested to address the credit card fraud problem for online transactions. However, the high-class imbalance is the major challenge that faces the existing solutions to construct an effective detection model. Most of the existing techniques used for class imbalance

overestimate the distribution of the minority class, resulting in highly overlapped or noisy and unrepresentative features, which cause either overfitting or imprecise learning. In this study, a credit card fraud detection model (CCFDM) is proposed based on ensemble learning and a generative adversarial network (GAN) assisted by Ensemble Synthesized Minority Oversampling techniques (ESMOTE-GAN). Multiple subsets were extracted using under-sampling and SMOTE was applied to generate less skewed sets to prevent the GAN from modelling the noise. These subsets were used to train diverse sets of GAN models to generate the synthesized subsets. A set of Random Forest classifiers was then trained based on the proposed ESMOTE s cost intensive, due to the need for human intervention; thus, it is not suitable for the huge volume of transactions even if the over sampling leads to a balance in the dataset, the internal distribution of the minority class might become unrepresentative, due to the unpredictable behaviour of the fraud. Despite these drawbacks, the research community has widely adopted SMOTE. Under-sampling the minority class leads to unrepresentative features leading to impractical solutions. The imprecise prediction by GAN lead to generating inaccurate samples. The ability to denoise the features leads to oversampling using auto encoding and hence an over fitting problem.

Ibomoiye Domor Mienye *et al* [2023] Credit cards play an essential role in today's digital economy, and their usage has recently grown tremendously, accompanied by a corresponding increase in credit card fraud. Machine learning (ML) algorithms have been utilized for credit card fraud detection. ML classifiers to achieve optimal performance. In order to solve this problem, this paper proposes a robust deep-learning approach that consists of long short-term memory (LSTM) and gated recurrent unit (GRU) neural networks as base learners in a stacking ensemble framework, with a multilayer perceptron (MLP) as the meta-learner. Meanwhile, the hybrid

synthetic minority oversampling technique and edited nearest neighbour (SMOTE-ENN) method is employed to balance the class distribution in the dataset. The experimental results showed that combining the proposed deep learning. LSTM and GRU-based RNNs to solve the vanishing gradient problem and have shown good performances in different sequence classification oversampling methods create balanced training sets by duplicating samples in the minority class, which could result in overfitting. Deep learning-based ensemble models have rarely been employed for credit card fraud detection, even though combining deep learning-based techniques such as LSTM in an ensemble model could result in more robust models. RNNs are not limited to a unidirectional data flow. They can loop through several layers and temporarily memorize information that can be used later.

Yuanming Ding *et al* [2023] presented a rapid spread of mobile banking and e-commerce has coincided with a dramatic increase in fraudulent online payments in recent years. Although machine learning and deep learning are widely used in credit card fraud detection, the typical credit card transaction data set is unbalanced, and the fraud data is much less than the normal transaction data, limiting the effectiveness of traditional binary classification algorithms. To overcome this issue, researchers oversample minority class data and utilize ensemble learning classification algorithms. However, oversampling still has disadvantages. The method is tested on an open credit card dataset, with the experimental results demonstrating that the oversampling method utilizing the improved VAEGAN is superior to the oversampling method of Generative Adversarial Network (GAN), Variational Autoencoder (VAE), and Synthetic. ‘‘Time’’ and ‘‘Amount’’, that have not undergone PCA conversion. Scaling factor A changes the density value corresponding to the value of each selected random variable and does not change the expected sum after the product Variance. The GAN method promotes the F1 value when the expansion ratio is less than three,

and the effect is not as good as the baseline model when the expansion ratio is greater than three. The improved VAEGAN does not perform equally well to SMOTE on the classification indicator AUC but has improved performance compared to other oversampling and baseline models. The improved VAEGAN does not perform equally well to SMOTE on the classification indicator AUC but has improved performance compared to other oversampling and baseline models. Normal transaction data, limiting the effectiveness of traditional binary classification algorithms.

Wang Ning *et al* [2023] described a popularity of online transactions, credit card fraud incidents are occurring more and more frequently, and adaptive enhancement (Adaboost) models are most often used in credit card fraud detection, so how to improve the robustness of the traditional Adaboost algorithm has become a hot issue. A large part of the reason for the poor robustness of the traditional Adaboost algorithm is that the base classifier is selected in a way that is uniquely oriented to the error rate. The self-paced learning selected. The self-adaptive threshold finding algorithm selected in this paper can well mitigate the influence of human experience on model training. The traditional Adaboost algorithm suffers from problems such as overfitting and poor robustness. These improvements do not fundamentally solve the problem of overfitting in Adaboost algorithm. The direction of decision-making is not changed, thereby reconstructing the weak classifier weight update method. The reliance on grid search not only leads to a significant increase in model training time, but also makes it more difficult to set the parameter search range. The classifier that minimizes the generalization error cannot be selected, and the speed of model convergence decrease, and even lead to a decrease in the overall model performance. Improvements do not fundamentally solve the problem of overfitting in Adaboost algorithm.

Huang Tingfei *et al* [2020] brief that Machine learning approaches are widely used to analyse and detect the increasingly serious problem of credit card fraud. However, typical credit card datasets present imbalanced classification situations because of severely skewed class distributions. Although researchers have some strategies to deal with these imbalances, disadvantages remain. It outperforms recent oversampling methods based on generative adversarial network (GAN) models. After submitting the extended dataset to the baseline for training, the test of the VAE model performs well on indicators including precision, F-measure, accuracy and specificity. Credit card fraud not only brings huge economic losses to financial institutions and banks, but also trouble and stress to the lives of individuals who are affected. The parameters of the generator and discriminator cannot be obviously optimized. The loss function values of the discriminator and generator are not changed obviously, the model is considered to have reached an actual balance. The final output of the model is not image data, using a convolution network as the neural network module of VAE can't achieve better performance. The comprehensive samples generated by the model are not all positive data, so the injection of comprehensive samples in the training set will causes an increase in false negatives. Aspects are not only larger than the baseline, but also improved by about 3% than the optimal.

Bertrand Lebichot *et al* [2021] demonstrated Credit card fraud jeopardizes the trust of customers in e-commerce transactions. This led in recent years to major advances in the design of automatic Fraud Detection Systems (FDS) able to detect fraudulent transactions with short reaction time and high precision. Present and discuss 15 transfer learning techniques (ranging from naive baselines to state-of-the-art and new approaches), making a critical and quantitative comparison in terms of precision for different transfer scenarios. Quantifying the distance between domains is not

trivial. The source and target domain cannot be revealed for confidentiality reasons. Model should use features that cannot discriminate between the source and target domains. Expected, when the ratio r decreases, the accuracy of all methods decreases as well. In order to compare the different strategies, it is then important to consider the trend of the accuracy for decreasing r (the slower the deterioration the better). If increase the number of target labels, the adoption of adversarial or augmented feature strategies is recommended. Restrained to consider the most relevant features (notably the ones whose importance computed by Random Forest is in the top third). That humans take advantage of previously learned skills (e.g. recognize apples or playing the piano) to speed up the learning of somewhat related tasks (recognize pears or playing the organ). The advantage of being at the same time self-labelling and instance-based (see Section III). As in the original paper, only the EE classifier is included as a base learner. The abundant labelled data present in the source domain to estimate the conditional probabilities in the source domain and the priors. Working in an unbalanced world.

Rafael San Miguel Carrasco *et al* [2020] developed in Fraud detection systems support advanced detection techniques based on complex rules, statistical modelling and machine learning. Triggered by these systems still require expert judgement to either confirm a fraud case or discard a false positive. Reducing the number of false positives that fraud analysts investigate, by automating their detection with computer-assisted techniques, can lead to significant cost efficiencies. Alert reduction has been achieved with different techniques in related fields like intrusion detection. The performance achieved by each neural network setting is presented and discussed. The optimal setting allowed to capture of total fraud cases with less alerts. Obtained alert reduction rate would entail a significant reduction in cost of human labour, because alerts classified as false positives by the

neural network wouldn't require human inspection. Oversampling methods like SMOTE have proven to improve accuracy in supervised models. Lastly, recent research concludes that combining multiple outlier scores can negatively impact accuracy. Clustering customers ahead of the classification task improves accuracy for certain types of customers. Not suitable for the unsupervised approach used in DAE. Did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Daniele Lunghi *et al* [2023] described an Imbalanced learning jeopardizes the accuracy of traditional classification models, particularly for what concerns the minority class, which is often the class of interest. This paper addresses the issue of imbalanced learning in credit card fraud detection by introducing a novel approach that models fraudulent behaviour as a time-dependent process. The main contribution is the design and assessment of an oversampling strategy, called 'Adversary-based Oversampling', which relies on modelling the temporal relationship among frauds. The strategy is implemented by two learning approaches: first, an innovative regression-based oversampling model that predicts subsequent fraudulent activities based on previous fraud features. Second, the adaptation of the state-of-the-art TimeGAN oversampling algorithm to the context of credit card fraud detection. An oversampling approach incorporating time-dependent modelling of frauds provides competitive results, measured against common fraud detection metrics, compared to traditional oversampling algorithms. Did not find sufficiently accurate results. Instead, realized that the behaviour of fraudsters does indeed exhibit time dependent patterns. Implementations do not allow maintaining the structure in the data, as the interpolation between admissible observations not be an admissible observation. It does not explicitly estimate the fraudsters' behaviour. Moreover, TimeGAN ADV-O is a generative model, meaning that the artificial frauds are not explicitly built on the ones in the original training

data. It must be noted this does not prove that fraudsters do not adapt their behaviour to the card it has access to, but that do not have any hint in this direction from our data. All performances are similar and significantly improved compared to the baseline, further showcasing the importance of resampling techniques for fraud detection.

N. K. Trivedi *et al* [2020] proposed a large volume, wide range, frequency, as well as importance is stored from small to large organizations over the cloud. The whole information is available from massive amounts of sources such as followers on social media, customer order behaviours, likes, and shares. White-collar crime is the ever-increasing problem with-reaching consequences for the finance sector, business institutions as well as governments. Fraud can indeed be described as illegal deceit to gain financial benefit. Enhanced card transactions had already appreciated a heavy emphasis on communication technology. When credit card transactions are by far the most prevalent form of transaction for offline and online payments, raising the rate of card fraud accelerates as well.

O.S. Yee *et al* [2018] developed the application of machine learning techniques spreads widely throughout computer sciences domains such as spam filtering, web searching, ad placement, recommender systems, credit scoring, drug design, fraud detection, stock trading, and many other applications. Machine Learning classifiers operate by building a model from example inputs and using that to make predictions or decisions, rather than following strictly static program instructions. There are many different types of machine learning approaches available with the intentions to solve heterogeneous problems. Due to the nature of this study, which was focused on classification, the discussion that follows is based on this topic. Machine learning classification refers to the process of learning to assign instances to predefined classes. Formally, there are several types of learning such as supervised, semi-supervised, and unsupervised, reinforcement, transduction

and learning to learn. To conduct supervised based machine learning classification, the discussions about the rest of the methods are discarded from further elaboration. In most classification studies, supervised based learning is favoured more than other methods due to the ability to control the classes of the instances with the interventions of human. In supervised learning, the classes of the instances would be labelled prior to feeding into classifiers. Then, by using certain evaluation metrics, the performances of the classifiers could be measured.

A. Roy *et al* [2018] presented the the various parameters that are used to construct the model to find the optimal combination of parameters to detect fraudulent activity. Deep Learning algorithms are a class of machine learning algorithms that use multiple non-linear processing units for feature extraction and transformation. These processing units discover intermediate representations in a hierarchical manner. The features discovered in one layer form the basis for processing of the succeeding layer. In this way, Deep Learning algorithms learn intermediate concepts between raw input and target. Deep Learning initially found use in automatic speech recognition, image recognition and natural language processing. More recently, Deep Learning algorithms have been used in areas that require the prediction of human behaviour such as customer relationship management, recommendation systems and mobile advertising. However, all these applications require choosing between different Deep Learning topologies (i.e., the structure of the model) as well as the parameters that are used to construct the model.

A. Thennakoon *et al* [2019] proposed an extreme increase in fraudulent transactions that affect the economy dramatically. Credit card fraud can be classified into several categories. The two types of frauds that can be mainly identified in a set of transactions are Card-not-present (CNP) frauds and Card-present (CP) frauds. Those two types can be described

further by bankruptcy fraud, theft/counterfeit fraud, application fraud, and behavioural fraud. The study aims at addressing four fraud natures that belong to the CNP fraud category described above and repose a method to detect those frauds real time. Machine learning is this generation's solution which replaces such methodologies and can work on large datasets which is not easily possible for human beings. Machine learning techniques fall into two main categories; supervised learning and unsupervised learning. Fraud detection can be done in either way and only can be decided when to use according to the dataset. Supervised learning requires prior classification to anomalies. During the last few years, several supervised algorithms have been used in detecting credit card fraud.

Bhanusri *et al* [2020] illustrated the prevailing data mining concerns people with credit card fraud detection model based on data mining. Since the problem is approached as a classification problem, classical data mining algorithms are not directly applicable. This project is to propose a credit card fraud detection system using supervised learning algorithm. Supervised algorithms are evolutionary algorithms which aim at obtaining better solutions as time progresses. Credit card is the most popular mode of payment. As the number of credit card users is rising world-wide, the identity theft is increased, and frauds are also increasing. In the virtual card purchase, only the card information is required such as card number, expiration date, secure code, etc. Such purchases are normally done on the Internet or over telephone. To commit fraud in these types of purchases, a person simply needs to know the card details. The mode of payment for online purchase is mostly done by credit card. The details of credit card should be kept private. To secure credit card privacy, the details should not be leaked. Different ways to steal credit card details are phishing websites, steal/lost credit cards, counterfeit credit cards, theft of card details, intercepted cards etc. For security purpose, the above things should be avoided. In online fraud, the

transaction is made remotely and only the card's details are needed. A manual signature, a PIN or a card imprint are not required at the purchase time. In most of the cases the genuine cardholder is not aware that someone else has seen or stolen his/her card information. The simple way to detect this type of fraud is to analyse the spending patterns on every card and to figure out any variation to the "usual" spending patterns. Fraud detection by analysing the existing data purchase of cardholder is the best way to reduce the rate of successful credit card frauds. As the data sets are not available and also the results are not disclosed to the public. The fraud cases should be detected from the available data sets known as the logged data and user behaviour. At present, fraud detection has been implemented by a number of methods such as data mining, statistics, and artificial intelligence.

S. Patil *et al* [2018] described how to improve fraud detection accuracy with growing number of transactions done by user per second. The increase in number of users and online transactions has brought heavy workloads to these systems. The size of day-to-day transaction and past historical transactions has increased to several PB in recent years. Processing of data for model building, model training and predictive modelling on incoming transaction with minimum delay is very hard to achieve for current CCFD system. So as to overcome this problem proposed a solution of interfacing SAS with Hadoop framework and building self-adaptive analytical framework model on top of it for fraud detection. Have also compared the performance of various algorithm of fraud detection and tune the analytical server with most optimal model for fraud detection.

S. P Maniraj *et al* [2019] presented a case where a person uses someone else's credit card for personal reasons while the owner and the card issuing authorities are unaware of the fact that the card is being used. Fraud detection involves monitoring the activities of populations of users in order to estimate, perceive or avoid objectionable behaviour, which consist of

fraud, intrusion, and defaulting. This is a very relevant problem that demands the attention of communities such as machine learning and data science where the solution to this problem can be automated. This problem is particularly challenging from the perspective of learning, as it is characterized by various factors such as class imbalance. The number of valid transactions far outnumber fraudulent ones. Also, the transaction patterns often change their statistical properties over the course of time. These are not the only challenges in the implementation of a real-world fraud detection system, however. In real world examples, the massive stream of payment requests is quickly scanned by automatic tools that determine which transactions to authorize. Machine learning algorithms are employed to analyse all the authorized transactions and report the suspicious ones. These reports are investigated by professionals who contact the cardholders to confirm if the transaction was genuine or fraudulent.

D. Varmedja *et al* [2019] proposed some of the key parts of the company's future business. Because of that, companies need to store that data, to process it and what is really important, to keep it safe. Without securing data, a lot of it can be used by other companies or even worse, it can be stolen. In most cases, financial information is stolen, which can harm whole company or individual. There are several types of frauds. Check Fraud occurs when person forges a check or pays for something with check knowing that there is not enough money. Internet sales is fraud where fraudster sale fake items or counterfeit items or taking payment without delivering the item. There are a couple more, such as charities fraud, identity theft, credit card fraud, debt elimination, Insurance fraud and others. Due to increasing popularity of cashless transactions, one of the most common frauds are credit card frauds. Credit card fraud refers to the situation where fraudster uses credit card for their needs while owner of that credit card is not aware of that. Fraudulent transactions conducted using credit cards

acquired worldwide amounted to €1.8 billion in 2016. Although there is a tremendous volume increase in credit card transactions, the amount of frauds is proportionally the same or have decreased due to sophisticated fraud detection systems. However, fraudsters are constantly coming up with new ways to steal information.

S. Xuan *et al* [2018] presented credit cards are widely used due to the popularization of ecommerce and the development of mobile intelligent devices. Card-not-present transactions (i.e., online transaction without a physical card) is more popular, especially all credit card operations are performed by web payment gateways, e.g., PayPal and Alipay. Credit card has made an online transaction easier and more convenient. However, there is a growing trend of transaction frauds resulting in a great loss of money every year. It is estimated that losses are increased yearly at double digit rates by 2020. Since the physical card is not needed in the online transaction environment and the card's information is enough to complete a payment, it is easier to conduct a fraud than before. Transaction fraud has become a top barrier to the development of e-commerce and has a dramatic influence on the economy. Hence, fraud detection is essential and necessary. Fraud detection is a process of monitoring the transaction behaviour of a cardholder in order to detect whether an incoming transaction is done by the cardholder or others. Generally, there are two kinds of methods for fraud detection: misuse detection and anomaly detection. Misuse detection uses classification methods to determine whether an incoming transaction is fraud or not. Usually, such an approach has to know about the existing types of fraud to make models by learning the various fraud patterns. Anomaly detection is to build the profile of normal transaction behaviour of a cardholder based on his/her historical transaction data and decide a newly transaction as a potential fraud if it deviates from the normal transaction behaviour.

However, an anomaly detection method needs enough successive sample data to characterize the normal transaction behaviour of a cardholder.

J. Jurgovsky *et al* [2018] proposed to evolve and migrate to the internet and money is transacted electronically in an ever-growing cashless banking economy, accurate fraud detection remains a key concern for modern banking systems. It is not only to limit the direct losses incurred by fraudulent transactions, but also to ensure that legitimate customers are not adversely impacted by automated and manual reviews. In payment industry, fraud on card occurs when someone steals information from the card to do purchases without the permission and the detection of these fraudulent transactions has become a crucial activity for payment processors. A typical fraud detection system is composed of an automatic tool and a manual process. The automatic tool is based on fraud detection rules. It analyses all the new incoming transactions and assigns a fraudulent score. The manual process is made by fraud investigators. Focus on transactions with high fraudulent score and provide binary feedback (fraud or genuine) on all the transactions it analysed. The fraud detection systems can be based on expert driven rules, data driven rules or a combination of both types of rules. Expert driven rules try to identify specific scenarios of fraud discovered by the fraud investigators. A scenario of fraud can be a cardholder does a transaction in a given country and, in the two next weeks, (s) he does another transaction for a given amount in another given country." If this scenario is detected in the stream of transactions, then the fraud detection system will produce an alert. Data driven rules are based on machine learning algorithms. It learns the fraudulent patterns and try to detect them in a data-stream of new incoming transactions. The very commonly employed machine learning algorithms include logistic regression, support vector machines and random forests. Fraud detection is a challenging machine learning problem for many reasons: (i) the data distribution evolves over time because of seasonality and new

attack strategies, (ii) fraudulent transactions represent only a very small fraction of all the daily transactions and (iii) the fraud detection problem is intrinsically a sequential classification task. In this method, address primarily the last issue by using LSTM networks as machine learning algorithms in fraud detection systems.

CHAPTER-3

EXISTING SYSTEM

3.1 INTRODUCTION

In recent years, there has been a significant increase in the volume of financial transactions due to the expansion of financial institutions and the popularity of web-based e-commerce. Fraudulent transactions have become a growing problem in online banking, and fraud detection has always been challenging. Along with credit card development, the pattern of credit card fraud has always been updated. Fraudsters do their best to make it look legitimate, and credit card fraud has always been updated. Fraudsters do their best to make it look legitimate. They try to learn how fraud detection systems work and continue to stimulate these systems, making fraud detection more complicated. Therefore, researchers are constantly trying to find new ways or improve the performance of the existing methods. People who commit fraud usually use security, control, and monitoring weaknesses in commercial applications to achieve their goals. However, technology can be a tool to combat fraud. To prevent further possible fraud, it is important to detect the fraud right away after its occurrence. Fraud can be defined as wrongful or criminal deception intended to result in financial or personal gain. Credit card fraud is related to the illegal use of credit card information for purchases in a physical or digital manner. In digital transactions, fraud can happen over the line or the web, since the cardholders usually provide the card number, expiration date, and card verification number by telephone or website. There are two mechanisms, fraud prevention and fraud detection, that can be exploited to avoid fraud-related losses. Fraud prevention is a proactive method that stops fraud from happening in the first place. On the other hand, fraud detection is needed when a fraudster attempts a fraudulent transaction. Fraud detection in banking is considered a binary classification problem in which data is classified as legitimate or fraudulent. Because

banking data is large in volume and with datasets containing a large amount of transaction data, manually reviewing and finding patterns for fraudulent transactions is either impossible or takes a long time. Therefore, machine learning-based algorithms play a pivotal role in fraud detection and prediction. Machine learning algorithms and high processing power increase the capability of handling large datasets and fraud detection in a more efficient manner. Machine learning algorithms and deep learning also provide fast and efficient solutions to real-time problems.

3.2 METHODOLOGY

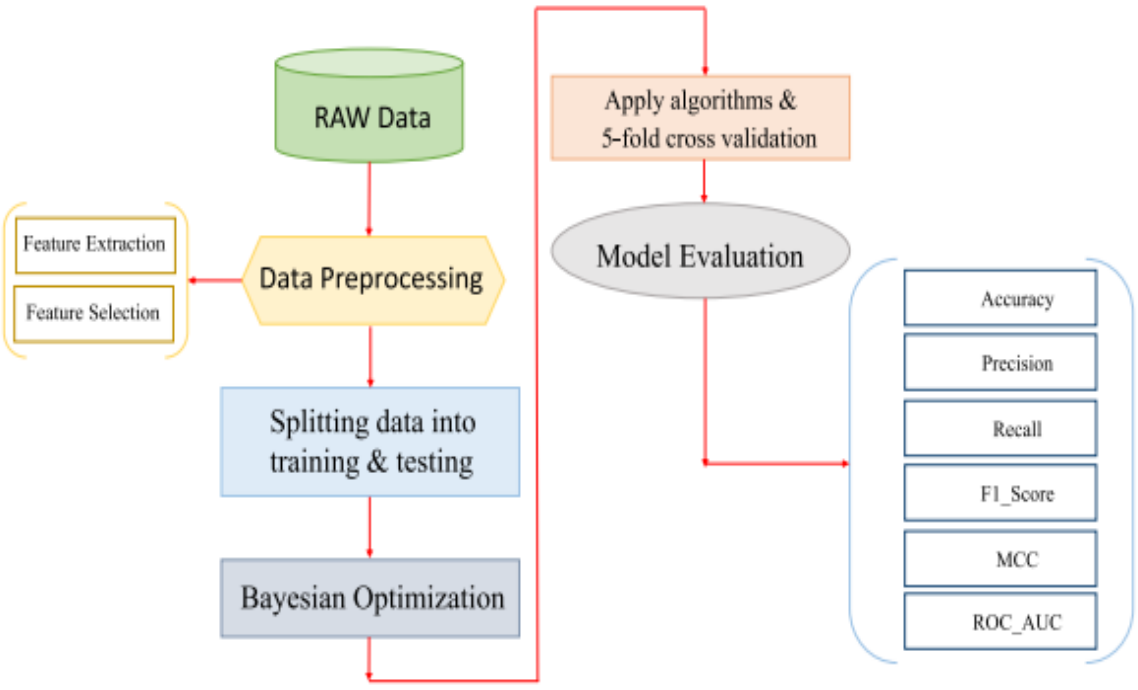

Figure 3.1 Block Diagram for the Existing System

First, apply the desired pre-processing the data and further dividing the data into two sections: training and testing. Followed by performing Bayesian optimization on the training data to find the best Hyperparameter that lead to the improvement of performance. Weight-tuning as a preprocess for unbalanced data, as well as CatBoost and XGBoost to improve the performance of the LightGBM method by accounting for the voting mechanisms, are used. The cross-validation

method is used to obtain a performance comparison in an unbalanced set and then examine the algorithms using different evaluation metrics.

Data Pre-Processing

In real datasets for credit card fraud detection, unbalanced data is expected. This data imbalance causes performance issues in machine learning algorithms, and having a class with the majority of the samples influences the evaluation results. Therefore, in many studies, under sampling and over-sampling methods are used to solve the data imbalance problem. Using under-sampling methods leads to data loss. Besides, using over-sampling methods leads to the production of duplicate data that doesn't provide information.

Feature Extraction

The ''time' feature includes the time (in seconds) elapsed between each transaction and the first transaction. To make the most of the feature, we expand it to extract the transaction hour feature, which gives us more information than the time feature itself.

Feature Selection

The features are unknown except for ''Time'' and ''Amount'', and we have no additional information. Feature selection tries to find a subset of features that improve the classifier's performance on effectively detecting credit card fraud. The information gain (IG) method is used to select the most important features that lead to a dimension reduction of the training data. Information gain functions by extracting similarities between credit card transactions and then awarding the greatest weight to the most significant features based on the class of legitimate and fraudulent credit card transactions. The information gain method has been proven to be computationally efficient and shows leading performance in terms of precision.

Algorithms

Hyperparameters have a significant effect on the performance of machine learning models. We refer to optimization as the process of finding the best set of hyperparameters that configure a machine learning algorithm during its training. Recently, it was shown that the Bayesian method is capable of finding the optimized values in a much smaller number of training courses compared with evolutionary optimization methods. In Bayesian optimization algorithm to tune the hyperparameters that lead to computational time reduction and performance improvement.

3.3 RESULTS AND DISCUSSIONS

The stratified 5-fold cross validation method and the boosting algorithms with the Bayesian optimization method to evaluate the performance of the framework. Extract the hyperparameters and evaluate each algorithm individually before using the majority voting method.

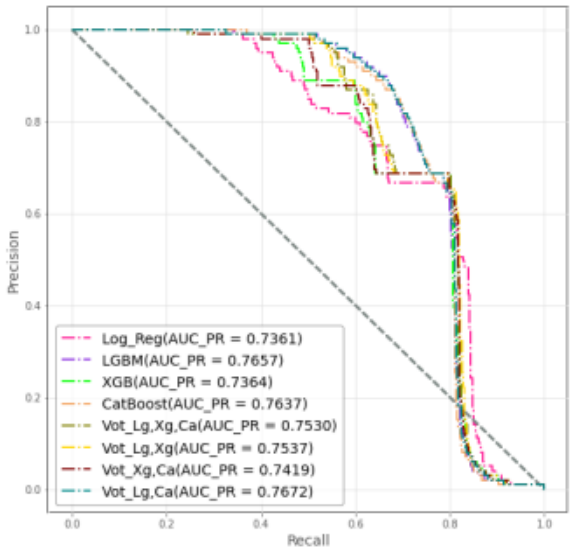

Figure 3.2 Precision_Recall curve.

The precision-recall curve is illustrated in Figure 3.2 and shows the system performance in a more precise manner compared with the ROC-AUC

curve. However, the results cannot be cited because false negatives are far from the view of this diagram.

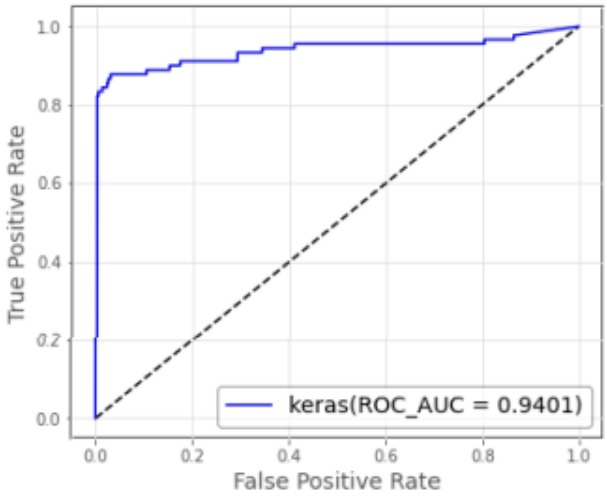

Figure 3.3 ROC curve of deep learning.

According to the digits obtained in figure 3.3, deep learning has achieved better performance compared with individual algorithms and majority voting ensemble learning. The MCC and F1-score metrics have values of 0.8129 and 0.8132, respectively.

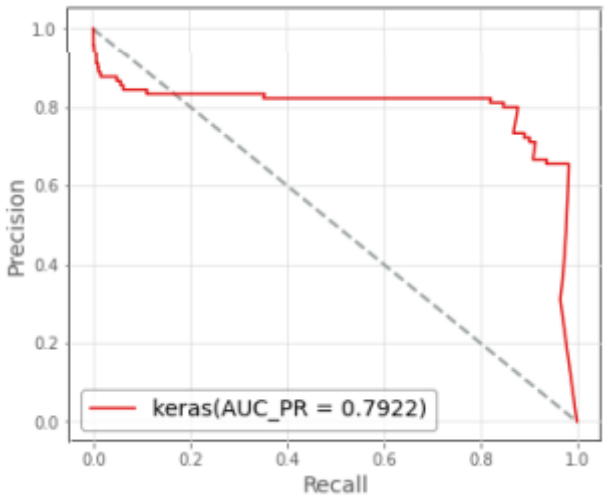

Figure 3.4 Precision- recall curve of deep learning

The diagram of the Precision-Recall curve is shown in Figure 3.4 and shows the value as 0.7922.

3.4 CONCLUSION

The credit card fraud detection problem in real, unbalanced datasets hence, this work presents a machine learning approach to improve the performance of fraud detection. Our experimental results showed that the LightGBM method improved the fraud detection cases by 50% and the F1-score by 20% compared with the recently presented method. Improve the performance of the algorithm with the help of the majority voting algorithm. It also improved the criteria by using the deep learning method. The assurance of the results of MCC for unbalanced data proved that, compared to other criteria of evaluation, it's stronger.

3.5 DRAWBACKS OF EXISTING SYSTEM

- However, this method causes an increase in the false-positive rate, which is not acceptable in banking for customer orientation.
- Additionally, the results cannot be cited because false negatives are far from the view of this ROC-AUC curve diagram.

CHAPTER-4

PROPOSED SYSTEM

4.1 INTRODUCTION

Fraud refers to obtaining goods/services and money by illegal way. Fraud deals with events which involve criminal motives that, mostly, are difficult to identify. Credit cards are one of the most popular objectives of fraud but not the only one. Credit card fraud, a wide-ranging term for theft and fraud committed or any similar payment mechanism as a fraudulent resource of funds in a transaction. Credit card fraud has been expanding issue in the credit card industry. Detecting credit card fraud is a difficult task when using normal process, so the development of the credit card fraud detection models has become of importance whether in the academic or business organizations currently. Furthermore, role of fraud has been changed suddenly during the last few decades along with advancement of technologies. Credit Card Fraud is one of the biggest threats to business and commercial establishments today. Simply, Credit Card Fraud is defined as, "when an individual uses another individual‟ credit card for personal use while the owner of the card as well as the card issuer are not aware of the thing that the card is being used." A number of systems/models, process and preventive measures will help to stop credit card fraud and reduce financial risks. Banks and credit card companies have gathered large amounts of credit card account transactions. The Credit Card is a plastic card issued to number of users as one of the modes of payment. It allows cardholders to purchasing goods and services based on the cardholder's promise. In China, credit card users are growing rapidly, but only a very few credit card holders use credit cards for paying for day-to-day purchase comparatively with confidence and a sense of security. Reason is credit card holder has not enough confidence to trust upon the payment system. Secure credit services of banks and

development of E-business a reliable fraud detection system is essential to support safe credit card usage, Fraud detection based on analysing existing purchase data of cardholder (current spending behaviour) is a promising way for reducing the rate of credit card frauds. Fraud detection systems come into scenario when the fraudsters exceed the fraud prevention systems and start fraudulent transactions. Credit card frauds can be proceeded in many different ways such as simple theft, counterfeit cards, Never Received Issue (NRI), application fraud and online/Electronic fraud. Credit card fraud detection is dreadfully difficult, but also common problem for solution. As there is limited amount of data with the transactions being confided, for example, transaction amount, merchant category code (MCC), acquirer number and date and time, address of the merchant. Various techniques in Knowledge Discovery, such as decision tree, neural network and case-based reasoning have broadly been used for forming several fraud detection systems/ models. These techniques usually need adequate number of normal transactions and fraud transactions for learning fraud patterns.

4.2 PROPOSED WORK

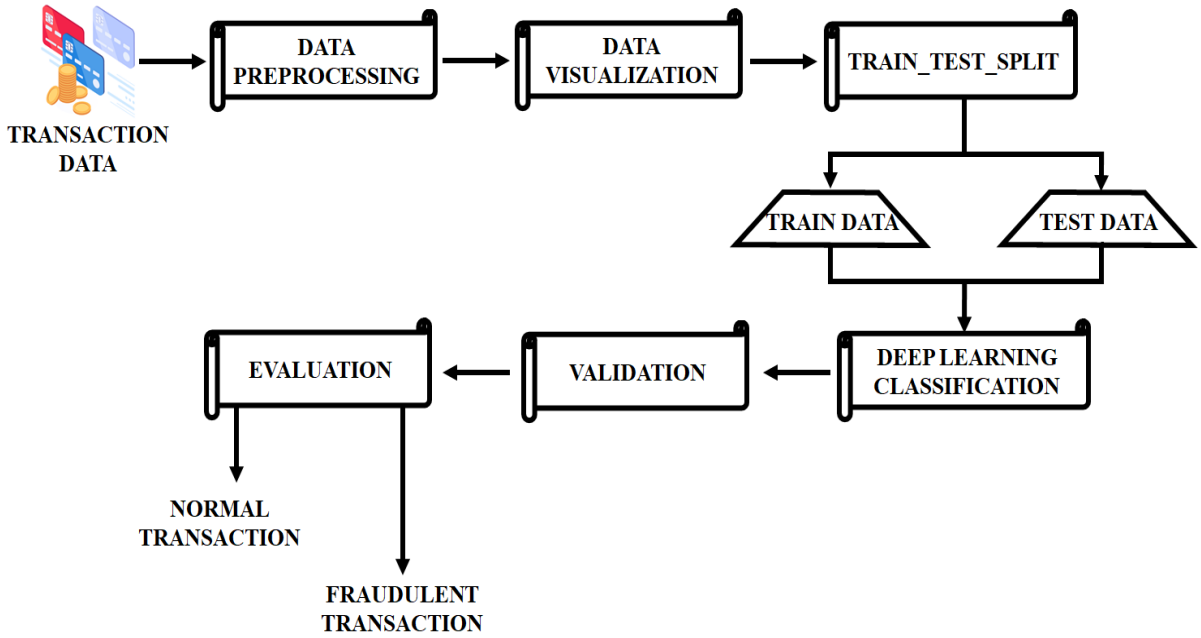

Figure 4.1 Block Diagram for Proposed System

First, the input dataset is allowed to undergo preprocessing. Which involves cleaning and analyzing the input dataset and handling the missing values. Following data pre-processing, the next stage of data visualization Matplotlib and Seaborn tools are used in the data visualization process. These modules are used for observing, exploring, and understanding the data in detail. Next, the data splitting process handles the data set and divides training and testing data for the purpose of the regression and identification process. Then the datasets are allowed to be classified using the classification techniques of deep learning. The deep learning technique handles the MLP algorithm. It works to handle robustness, missing data, and feature importance analysis against overfitting. Which makes it a popular option for a variety of uses, including prediction and detection issues. Additionally, the classification techniques provide excellent specificity and sensitivity. Following that, data validation is employed to improve integrity. Finally, the predicted fraudulent transaction is identified in the evaluation block. Overall, the proposed project was developed in Python, employing the Jupyter software.

4.3 DATA PREPROCESSING

Data preprocessing, a component of data preparation, describes any type of processing performed on raw data to prepare it for another data processing procedure. It has traditionally been an important preliminary step for the data mining process. The need for data preprocessing is there because good data is undoubtedly more important than good models and for which the quality of the data is of paramount importance. Therefore, companies and individuals invest a lot of their time in cleaning and preparing the data for modeling. The data present in the real world contains a lot of quality issues, noise, inaccurate, and not complete. It may not contain relevant, specific attributes and could have missing values, even incorrect and spurious values. To improve the quality of the data, preprocessing is essential. In

preprocessing, the collected datasets are given as input to the machine, which is then trained accordingly. Preprocessing helps to make the data consistent by eliminating any duplicates, irregularities in the data, normalizing the data to compare, and improving the accuracy of the results.

4.4 DATA VISUALIZATION

Data visualization is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from. The main goal of data visualization is to make it easier to identify patterns, trends and outliers in large data sets. Data visualization is a field in data analysis that deals with visual representation of data. It graphically plots data and is an effective way to communicate inferences from data.

Using data visualization, it can get a visual summary of our data. With pictures, maps and graphs, the human mind has an easier time processing and understanding any given data. Data visualization plays a significant role in the representation of both small and large data sets, but it is especially useful when we have large data sets, in which it is impossible to see all of our data, let alone process and understand it manually. These processes are handled in two modules: Matplotlib and Seaborn.

Matplotlib:

Matplotlib is a comprehensive library for creating static, animated, and interactive visualizations in Python. Matplotlib makes easy things easy and hard things possible.

Seaborn:

Seaborn is a library for making statistical graphics in Python. It is built on top of Matplotlib, and which works on exploring and understanding the data.

4.4.1 Types of Data Visualizations

Tables

This consists of rows and columns used to compare variables. Tables can show a great deal of information in a structured way, but they can also overwhelm users that are simply looking for high-level trends.

Pie charts and stacked bar charts

These graphs are divided into sections that represent parts of a whole. They provide a simple way to organize data and compare the size of each component to one other.

Line charts and area charts

These visuals show change in one or more quantities by plotting a series of data points over time and are frequently used within predictive analytics. Line graphs utilize lines to demonstrate these changes while area charts connect data points with line segments, stacking variables on top of one another and using color to distinguish between variables.

Histograms

This graph plots a distribution of numbers using a bar chart (with no spaces between the bars), representing the quantity of data that falls within a particular range. This visual makes it easy for an end user to identify outliers within a given dataset.

Scatter plots

These visuals are beneficial in revealing the relationship between two variables, and they are commonly used within regression data analysis. However, these can sometimes be confused with bubble charts, which are used to visualize three variables via the x-axis, the y-axis, and the size of the bubble.

Heat maps

These graphical representation displays are helpful in visualizing behavioral data by location. This can be a location on a map, or even a webpage.

Tree maps

Which display hierarchical data as a set of nested shapes, typically rectangles. Tree maps are great for comparing the proportions between categories via their area size.

4.5 CLASSIFICATION

Data classification is the process of organizing data into categories that make it easy to retrieve, sort and store for future use. A well-planned data classification system makes essential data easy to find and retrieve. Data classification is broadly defined as the process of organizing data by relevant categories so that it may be used and protected more efficiently. On a basic level, the classification process makes data easier to locate and retrieve.

4.5.1 MULTILAYER PERCEPTRON (MLP)

An artificial neural network model with feed forward architecture that maps sets of input data onto a set of desired outputs iteratively, through the process of learning. A MLP consists of an input layer of neurons, one or more hidden layers of neurons and an output layer of neurons, where each layer is fully connected to the next layer. It is a feed forward artificial neural network model that maps input sets to output proper sets. A multilayer perceptron (MLP) is made up of multiple layers, each of which is connected to the next. Each node is a processing element or a neuron that has a nonlinear activation function except the input nodes. The final stage is the prediction process, output is generated by the output layer using the data processed in the hidden layer. It employs back propagation, a supervised learning technique, to train the network.

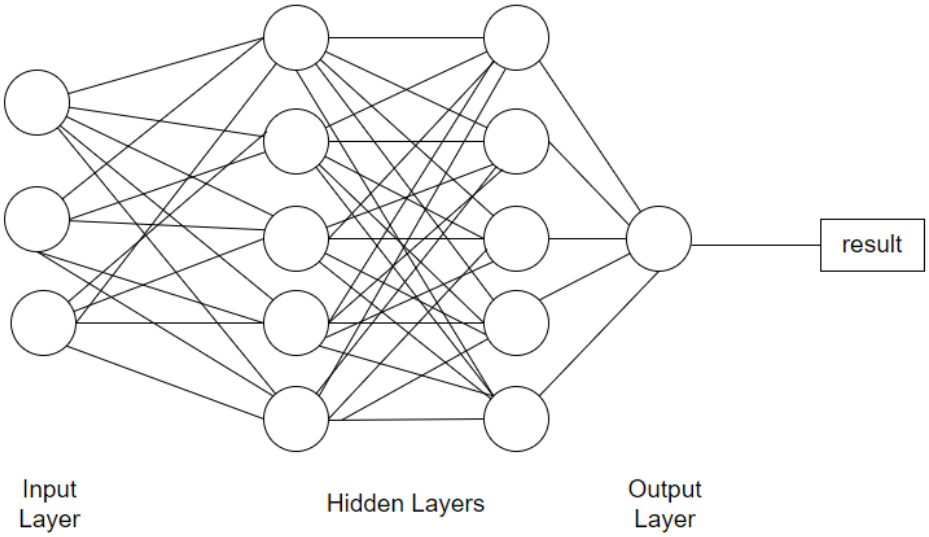

Figure 4.2 Basic Structure of MLP

BASIC STRUCTURE OF MLP

A multi-layered perceptron consists of interconnected neurons transferring information to each other, much like the human brain. Each neuron is assigned a value. The network can be divided into three main layers.

Input Layer

This is the initial layer of the network which takes in an input which will be used to produce an output.

Hidden Layer(S)

The network needs to have at least one hidden layer. The hidden layer(s) perform computations and operations on the input data to produce something meaningful.

Output Layer

The neurons in this layer display a meaningful output.

The MLP is a layered feedforward neural network in which the information flows unidirectionally from the input layer to the output layer, passing through the hidden layers. Each connection between neurons has its own weight. Perceptron's for the same layer have the same activation function. In general, it is a sigmoid for the hidden layers. Depending on the application, the output layer can also be a sigmoid or a linear function.

Among many other algorithms, the widely known MLP learning algorithm is a backpropagation, which is a generalization of the Least Mean Squared rule. Weights can be corrected by propagating the errors from layer to layer starting with the output layer and working backwards, hence the name backpropagation. The MLP model performance depends not only on the choice of the variables, the numbers of hidden layers, nodes, and training data but also on the training parameters such as learning rate, momentum controlling the weight change, and number of iterations. A MLP with one hidden layer identifies the nonlinear function with lower accuracies. Networks with more hidden layers are likely to overfit the training data. The learning rate and the momentum control the speed and effectiveness of the learning process. In land change modeling, the analysis of the complex relationships between land transition and the large number of variables acting as drivers needs advanced empirical techniques to find a nonlinear function that describes such a complex relationship. Variables such as distance, slope, type of soil, land tenure, etc. are presented at the input node of the network. Each output node represents a different land transition for which explanatory variable values are known, as well as the land transition observed in the past. After the training step, the MLP is able to predict the potential change of each transition when new input data is presented to the network.

4.6 VALIDATION AND EVALUATION

Validation and evaluation are related processes that can be used in a variety of contexts, including machine learning, training, and software testing. Validation in machine learning is like an authorization or authentication of the prediction done by a trained model. A validation data set is used in deep learning to compare the performance of different trained models. Data validation means checking the accuracy and quality of source data before using, importing or otherwise processing data. While on the

other hand, evaluation in machine learning refers to assessment or test of entire machine learning model and its performance in various circumstances. The method of evaluation focuses on the accuracy of the model in predicting the end outcomes. Even though there are several stages, the stage of evaluating a DL model is the most crucial because it gives us an idea of the accuracy of model prediction.

CHAPTER 5
RESULT AND DISCUSSION

5.1 INTRODUCTION

The chosen language for this project was Python. Python is a general-purpose, high-level programming language. Python is a popular high-level, general-purpose programming language. In 1991, Guido van Rossum invented it, and the Python Software Foundation continued to develop it. Programmers may communicate their ideas in less lines of code because of its syntax, which was developed with readability of code as a primary focus. For many reasons, this was an easy decision. The Python language is supported by a large community. A visit to Stack Overflow can easily fix any issues that may arise. Since Python is one of the most often used languages on the website, it's probable that any query will get a straightforward response. Python offers several strong tools that are ready for scientific computing. Packages with extensive documentation and open availability include NumPy, Pandas, and SciPy. These kinds of packages will drastically reduce and alter the amount of code needed to develop a certain program. Iteration is quick as a result. The language Python is tolerant and allows programs to seem like pseudo code. When the pseudo code provided in instructional papers has to be checked and enforced, this might be useful. With a strong emphasis on indentation, its design philosophy prioritizes code readability. Python uses garbage collection and dynamic typing. It is compatible with several programming paradigms, such as object-oriented, functional, and structured (especially procedural). Because of its extensive standard library, it is frequently referred to as a "batteries included" language. Python was developed by Guido van Rossum in the late 1980s as a replacement for the ABC programming language. Python 0.9.0 was initially made available in 1991. 2000 saw the introduction of Python 2.0. 2008 saw the introduction of Python 3.0, a significant update that was not

entirely backwards compatible with previous iterations. The final Python 2 release was Python 2.7.18, which was made available in 2020. Programming language Python routinely ranks among the most popular ones. Included with the Anaconda distribution is a graphical user interface (GUI) called Anaconda Navigator, which facilitates the installation, configuration, and use of tools like Jupyter Notebook. An isolated environment is what a Conda Python environment is. It lets you install packages without changing how Python is installed on your system. The Anaconda program facilitates the creation of environments for several Python and package versions. In your project environments, Anaconda is also used for package installation, removal, and upgrades. With the Jupyter Notebook, an open-source online tool, you can create and distribute documents with narrative text, equations, live code, and visualizations. Python is a programming language that lets you work quickly and integrate systems more efficiently.

5.2 INPUT DATASET

This section presents results obtained from analyzing several banking samples to ascertain whether the byte sequences extracted by the proposed method provide useful information for manual analysis. Figures 5.1 display the input dataset of the fraud detection.

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9
0	0.0	-1.359807	-0.072781	2.536347	1.378155	-0.338321	0.462388	0.239599	0.098698	0.363787
1	0.0	1.191857	0.266151	0.166480	0.448154	0.060018	-0.082361	-0.078803	0.085102	-0.255425
2	1.0	-1.358354	-1.340163	1.773209	0.379780	-0.503198	1.800499	0.791461	0.247676	-1.514654
3	1.0	-0.966272	-0.185226	1.792993	-0.863291	-0.010309	1.247203	0.237609	0.377436	-1.387024
4	2.0	-1.158233	0.877737	1.548718	0.403034	-0.407193	0.095921	0.592941	-0.270533	0.817739
5	2.0	-0.425966	0.960523	1.141109	-0.168252	0.420987	-0.029728	0.476201	0.260314	-0.568671
6	4.0	1.229658	0.141004	0.045371	1.202613	0.191881	0.272708	-0.005159	0.081213	0.464960
7	7.0	-0.644269	1.417964	1.074380	-0.492199	0.948934	0.428118	1.120631	-3.807864	0.615375
8	7.0	-0.894286	0.286157	-0.113192	-0.271526	2.669599	3.721818	0.370145	0.851084	-0.392048
9	9.0	-0.338262	1.119593	1.044367	-0.222187	0.499361	-0.246761	0.651583	0.069539	-0.736727

V21	V22	V23	V24	V25	V26	V27	V28	Amount	Class
-0.018307	0.277838	-0.110474	0.066928	0.128539	-0.189115	0.133558	-0.021053	149.62	0
-0.225775	-0.638672	0.101288	-0.339846	0.167170	0.125895	-0.008983	0.014724	2.69	0
0.247998	0.771679	0.909412	-0.689281	-0.327642	-0.139097	-0.055353	-0.059752	378.66	0
-0.108300	0.005274	-0.190321	-1.175575	0.647376	-0.221929	0.062723	0.061458	123.50	0
-0.009431	0.798278	-0.137458	0.141267	-0.206010	0.502292	0.219422	0.215153	69.99	0
-0.208254	-0.559825	-0.026398	-0.371427	-0.232794	0.105915	0.253844	0.081080	3.67	0
-0.167716	-0.270710	-0.154104	-0.780055	0.750137	-0.257237	0.034507	0.005168	4.99	0
1.943465	-1.015455	0.057504	-0.649709	-0.415267	-0.051634	-1.206921	-1.085339	40.80	0
-0.073425	-0.268092	-0.204233	1.011592	0.373205	-0.384157	0.011747	0.142404	93.20	0
-0.246914	-0.633753	-0.120794	-0.385050	-0.069733	0.094199	0.246219	0.083076	3.68	0

Figure 5.1 Input Dataset

5.3 DATA DISTRIBUTIONS

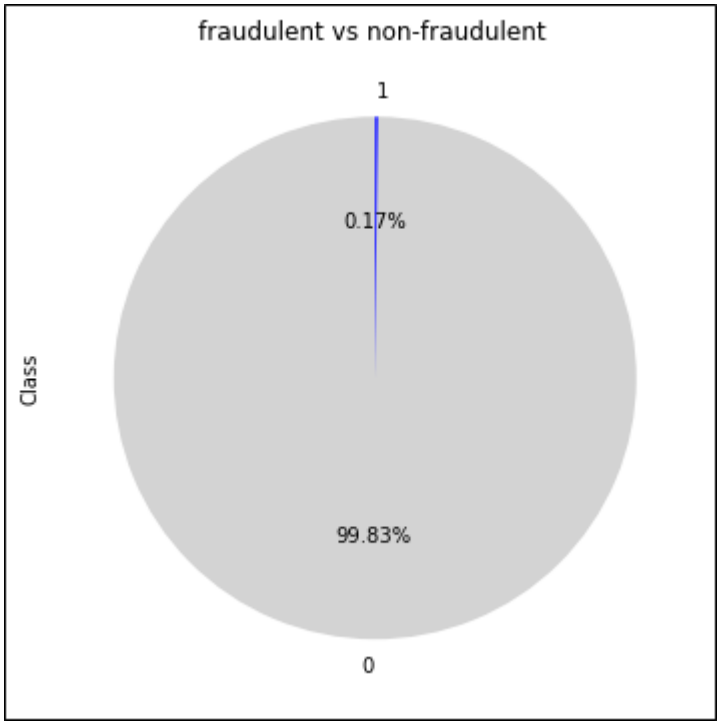

Figure 5.2 Data distributions for each class

Figure 5.2 displays the data distribution of the fraudulent data and non-fraudulent.

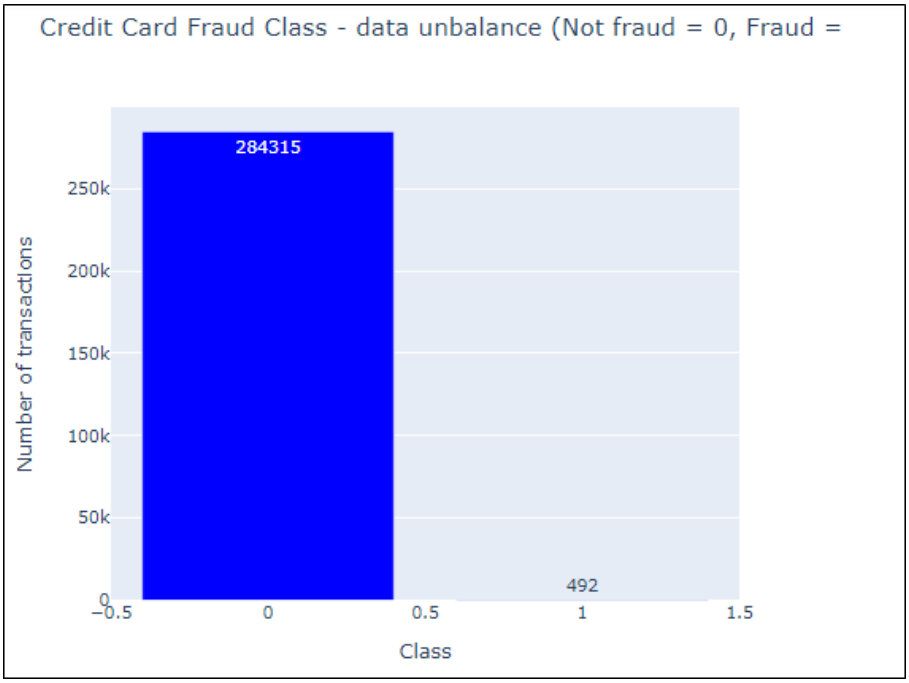

Figure 5.3 Class distributions

Data distributions for each class between normal transaction and fraud transaction image is displayed in Figure 5.3.

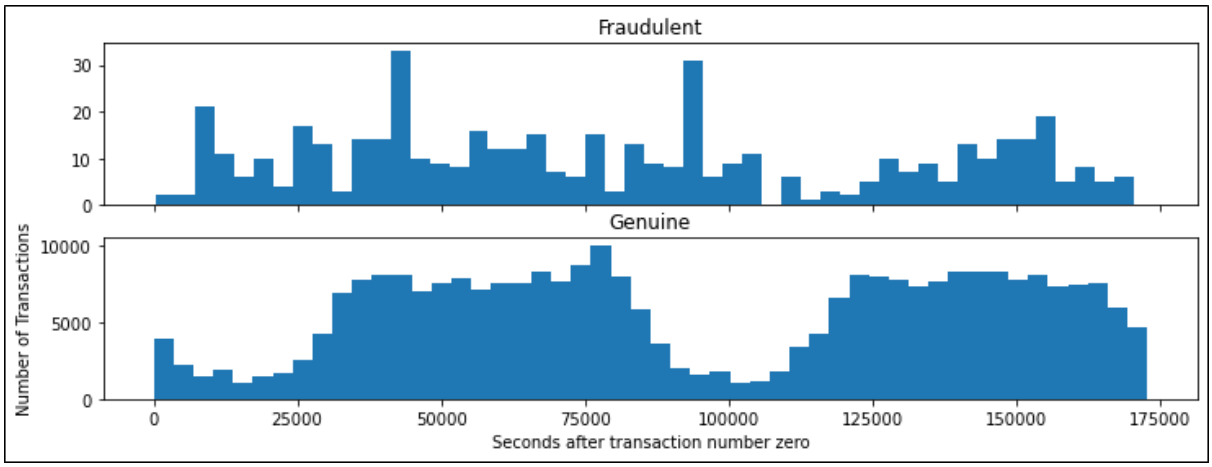

Figure 5.4 Exploration of transaction

Figure 5.4 shows illustrations of the exploration of transactions. These histogram images are deep exploration of the dataset based on the number of transactions.

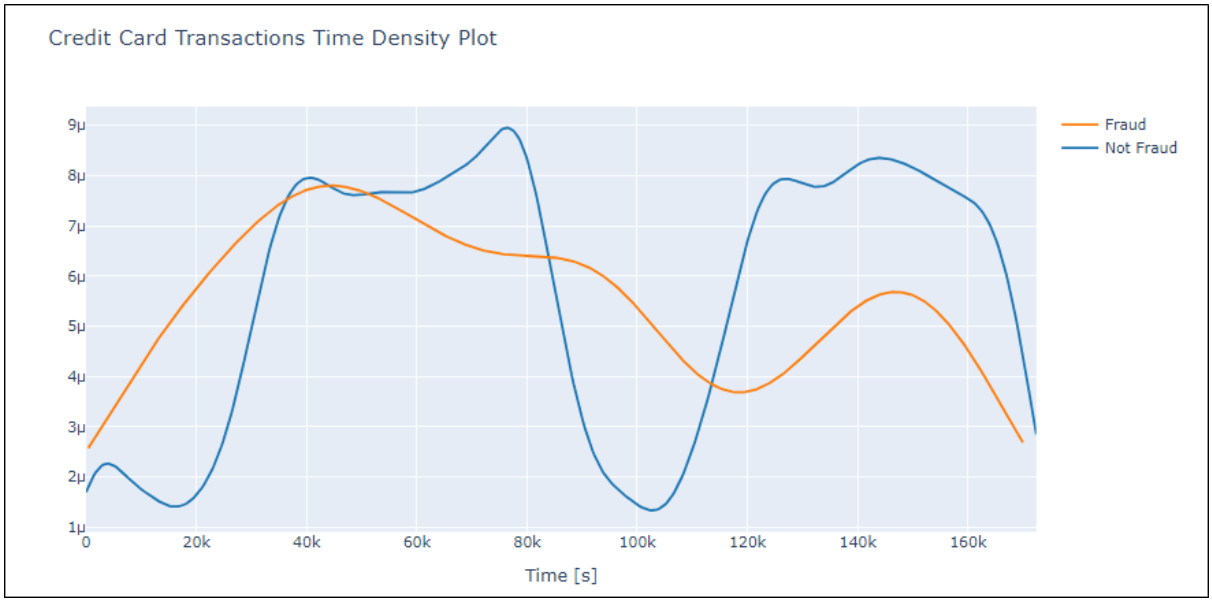

Figure 5.5 Time density plot

The time density plot is shown in Figure 5.5. In this figure, the blue colour line represents the normal transaction, and the orange colour line represents the fraud transaction based on time.

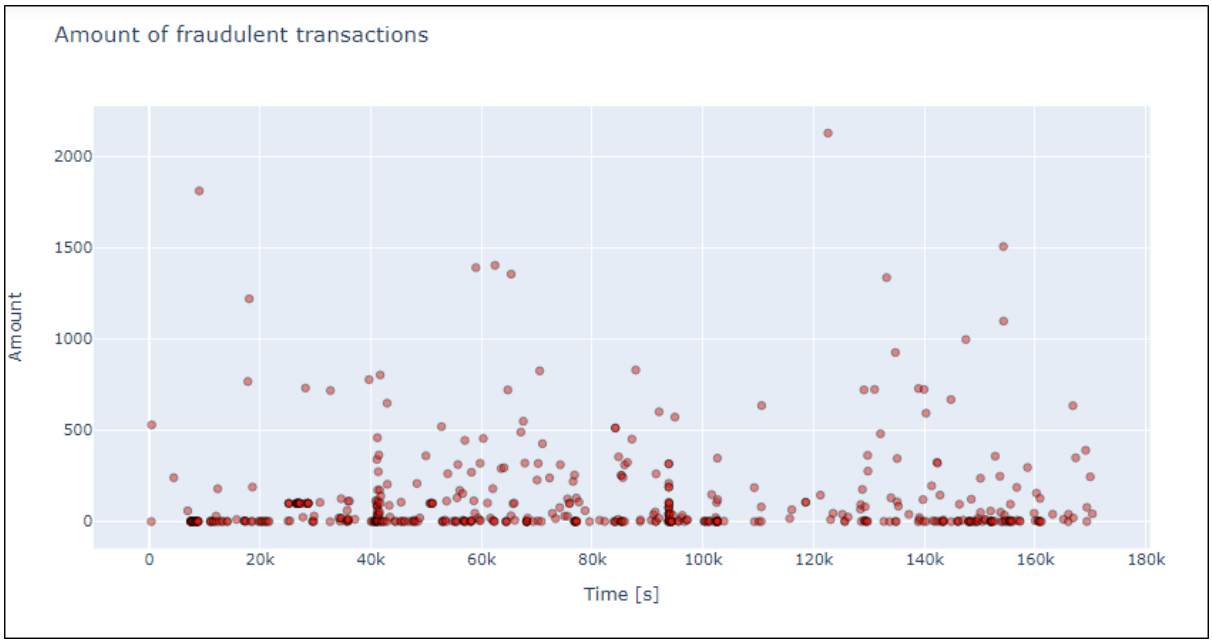

Figure 5.6 Fraudulent Transaction

Figure 5.6 establishes the fraudulent transaction.

5.4 MULTILAYER PERCEPTRON

A multilayer perceptron is a neural network connecting multiple layers in a directed graph, which means that the signal path through the nodes only goes one way. Each node, apart from the input nodes, has a nonlinear activation function. This figure represents that performance of MLP. Figure 5.7 shows that performance of MLP.

Performance of MLP

```
precision: [0.99953583 0.8411215 ]
recall: [0.99976083 0.73170732]
fscore: [0.99964832 0.7826087 ]
support: [71079   123]
```

Figure 5.7 Performance of MLP

Confusion Matrix of MLP

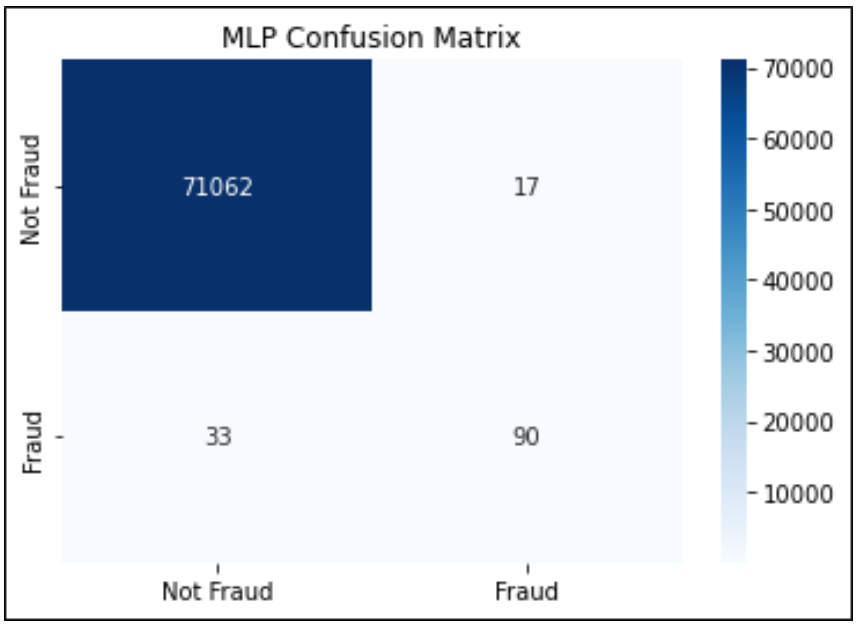

Figure 5.8 Confusion matrix for MLP

Figure 5.8 represents that confusion matrix for MLP. In this figure shows that value for malign and benign. It is often used to measure the performance of classification models, which aim to predict a categorical label for each input instance.

Roc Curve of MLP

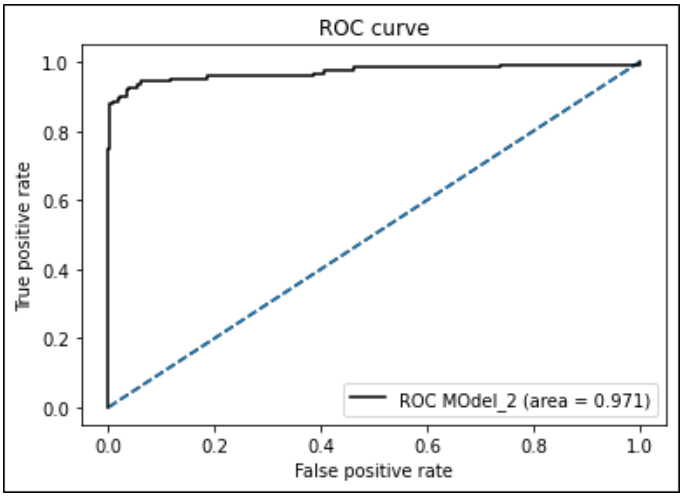

Figure 5.9 ROC Curve for MLP

Figure 5.9 represents a multilayer perceptron algorithm ROC Curve. An ROC curve (receiver operating characteristic curve) is a graph showing the performance of a classification model at all classification thresholds. This curve plots two parameters: True Positive Rate. False Positive Rate.

Precision Recall Curve for MLP

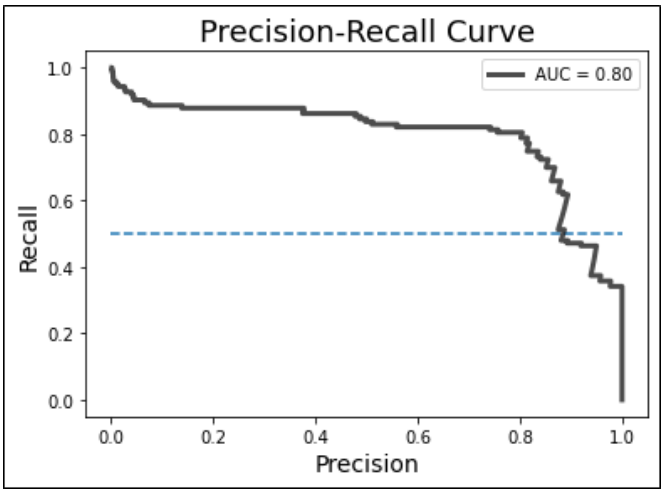

Figure 5.10 Precision Recall Curve for MLP

A multilayer perceptron algorithm (MLP) precision-recall curve is shown in Figure 5.10. A precision-recall curve is a plot of the precision (y-axis) and the recall (x-axis) for different thresholds, much like the ROC curve. A no-skill classifier is one that cannot discriminate between the classes and would predict a random class or a constant class in all cases.

CONCLUSION

In this project proposes, a Deep Learning Algorithm Implementation for Credit Card Fraud Analysis and Detection was implemented. The detection of credit card fraud is a vital research field. This is because of the increasing number of fraud cases in financial institutions. This project opens the door for employing machine learning to build systems that can detect fraud. Building an automated-based system to detect fraud requires a database to train the system (or classifier). This work demonstrates the advantages of applying deep learning techniques, including MLP classification techniques, to the credit card fraud detection problem for the purpose of reducing the bank's financial risks. Finally, the proposed classifier is evaluated based on its accuracy, and the MLP classifier generates the best results. Using these methods for the detection of credit cards yields better performance than traditional algorithms.

REFERENCES

1. F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan and M. Ahmed, "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms," in IEEE Access, vol. 10, pp. 39700-39715, 2022.
2. E. Ileberi, Y. Sun and Z. Wang, "Performance Evaluation of Machine Learning Methods for Credit Card Fraud Detection Using SMOTE and AdaBoost," in IEEE Access, vol. 9, pp. 165286-165294, 2021.
3. F. A. Ghaleb, F. Saeed, M. Al-Sarem, S. N. Qasem and T. Al-Hadhrami, "Ensemble Synthesized Minority Oversampling-Based Generative Adversarial Networks and Random Forest Algorithm for Credit Card Fraud Detection," in IEEE Access, vol. 11, pp. 89694-89710, 2023.
4. I. D. Mienye and Y. Sun, "A Deep Learning Ensemble with Data Resampling for Credit Card Fraud Detection," in IEEE Access, vol. 11, pp. 30628-30638, 2023.
5. Y. Ding, W. Kang, J. Feng, B. Peng and A. Yang, "Credit Card Fraud Detection Based on Improved Variational Autoencoder Generative Adversarial Network," in IEEE Access, vol. 11, pp. 83680-83691, 2023.
6. W. Ning, S. Chen, S. Lei and X. Liao, "AMWSPLAdaboost Credit Card Fraud Detection Method Based on Enhanced Base Classifier Diversity," in IEEE Access, vol. 11, pp. 66488-66496, 2023.
7. H. Tingfei, C. Guangquan and H. Kuihua, "Using Variational Auto Encoding in Credit Card Fraud Detection," in IEEE Access, vol. 8, pp. 149841-149853, 2020.
8. B. Lebichot, T. Verhelst, Y. -A. Le Borgne, L. He-Guelton, F. Oblé and G. Bontempi, "Transfer Learning Strategies for Credit Card Fraud Detection," in IEEE Access, vol. 9, pp. 114754-114766, 2021.

9. R. San Miguel Carrasco and M. -Á. Sicilia-Urbán, "Evaluation of Deep Neural Networks for Reduction of Credit Card Fraud Alerts," in IEEE Access, vol. 8, pp. 186421-186432, 2020.
10. D. Lunghi, G. M. Paldino, O. Caelen and G. Bontempi, "An Adversary Model of Fraudsters' Behavior to Improve Oversampling in Credit Card Fraud Detection," in IEEE Access, vol. 11, pp. 136666-136679, 2023.
11. Trivedi, N.K., Simaiya, S., Lilhore, U.K. and Sharma, S.K., An efficient credit card fraud detection model based on machine learning methods. International Journal of Advanced Science and Technology, 29(5), pp.3414-3424, 2020.
12. Yee, O.S., Sagadevan, S. and Malim, N.H.A.H., Credit card fraud detection using machine learning as data mining technique. Journal of Telecommunication, Electronic and Computer Engineering (JTEC), 10(1-4), pp.23-27, 2018.
13. A. Roy, J. Sun, R. Mahoney, L. Alonzi, S. Adams and P. Beling, "Deep learning detecting fraud in credit card transactions," 2018 Systems and Information Engineering Design Symposium (SIEDS), pp. 129-134, 2018.
14. A. Thennakoon, C. Bhagyani, S. Premadasa, S. Mihiranga and N. Kuruwitaarachchi, "Real-time Credit Card Fraud Detection Using Machine Learning," 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 488-493, 2019.
15. Bhanusri, A., Valli, K.R.S., Jyothi, P., Sai, G.V. and Rohith, R., 2020. Credit card fraud detection using Machine learning algorithms. Journal of Research in Humanities and Social Science, 8(2), pp.04-11.
16. Patil, S., Nemade, V. and Soni, P.K., Predictive modelling for credit card fraud detection using data analytics. Procedia computer science, 132, pp.385-395, 2018.

17. Maniraj, S.P., Saini, A., Sarkar, S.D., Ahmed, S., Credit Card Fraud Detection using Machine Learning and Data Science. IJERT, 8(9), pp.110-115, 2019.
18. Varmedja, D., Karanovic, M., Sladojevic, S., Arsenovic, M., Anderla, A., Credit Card Fraud Detection - Machine Learning methods. 2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH), pp. 1-5, 2019.
19. Xuan, S., Liu, G., Li, Z., Zheng, L., Wang, S., Jiang, C., Random Forest for credit card fraud detection. 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC), 2018, pp. 1-6, 2018.
20. Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.E., He-Guelton, L., Caelen,O., Sequence classification for credit-card fraud detection. Expert Systems with Applications (2018), 100(15), pp- 234-245, 2018.