**DEVELOPING A CREDIT CARD FRAUD DETECTION SYSTEM USING DECISION TREE AND ARTIFICIAL NEURAL NETWORK**

Name: Hussain Syed

Student ID: 19050831

Supervisor: Renuga Jayakumar

**MSc Final Project Declaration**

This report is submitted in partial fulfilment of the requirement for the degree of Master of Science in 7COM1039-0109-2021 - Advanced Computer Science Masters Project at the University of Hertfordshire (UH).

It is my own work except where indicated in the report.

I did not use human participants in my MSc Project.

I hereby give permission for the report to be made available on the university website provided the source is acknowledged.

**Abstract**

***In Chapter 1,*** the different types of credit card fraud that lead to the detection of the fraud issue are discussed. The current issue, specification of the study, and ultimate aims and objectives are also defined well.

***In Chapter 2,*** the development of the credit card fraud detection system using the decision tree and the artificial neural network is discussed. This has been implemented using the python programming language which has been done in the notebook platform. The background of this research also discusses the techniques and tools that are available to detect fraud through the merchants employing a combination.

***In Chapter 3,*** area of the research method, the elements that are crucial for conducting the steps for this particular assessment are covered. This section discusses the data analytics method and research strategy for the chosen method. This chapter has also sparked a discussion about the research methods that can help to detect fraudulent activities.

The results of the model developed using the python language have been developed in the ***result section, Chapter 4***. The machine learning algorithm has helped to generate the prediction and accuracy of the model. The machine learning algorithm has been developed using the decision tree classification and ANN model.

A critical evaluation of the results gathered in the Jupyter notebook has been performed to determine the effectiveness of the model in ***Chapter 5***. The model has achieved high accuracy and performance with the help of a “decision tree classifier” and ANN. Recommendations have been made to further enhance the performance and usage of the model in terms of credit card fraud detection. All the objectives have been successfully achieved in this research and methods to improve the model for further works have been suggested. Different steps can be performed that can increase the viability and accuracy of the model to prevent credit card fraud.

**Acknowledgment**

I thank my supervisors for providing guidance and feedback throughout this project.

**Table of Contents**

[Chapter 1: Introduction 8](#_Toc109350505)

[1.1 Introduction 8](#_Toc109350506)

[1.2 Project Specifications 8](#_Toc109350507)

[1.3 Aims and objectives 9](#_Toc109350508)

[1.4 Current issues 9](#_Toc109350509)

[Chapter 2: Research 10](#_Toc109350510)

[2.1 Quality of background research 10](#_Toc109350511)

[2.2 Use of literature 11](#_Toc109350512)

[2.3 Critical assessment 13](#_Toc109350513)

[2.4 Linkage to aims 14](#_Toc109350514)

[Chapter 3: Methodology 18](#_Toc109350515)

[3.1 Choice of methods 18](#_Toc109350516)

[3.2 Justification and support of choices 19](#_Toc109350517)

[3.3 Data collection 19](#_Toc109350518)

[3.4 Validation 20](#_Toc109350519)

[3.5 Ethical issues 21](#_Toc109350520)

[3.6 Commercial risk and risk management 21](#_Toc109350521)

[Chapter 4: Result 23](#_Toc109350522)

[4.1 Critical analysis 23](#_Toc109350523)

[4.2 Evidence of practical work 23](#_Toc109350524)

[4.3 Awareness and solutions to the technical challenges 26](#_Toc109350525)

[4.4 Novelty 27](#_Toc109350526)

[4.5 Interpretation of results 27](#_Toc109350527)

[4.6 Use of tools and techniques 32](#_Toc109350528)

[4.7 Appropriate tools for analysis 32](#_Toc109350529)

[4.8 Linkage to objectives and literature 32](#_Toc109350530)

[Chapter 5: Evaluation and Conclusion 33](#_Toc109350531)

[5.1 Critical Evaluation 33](#_Toc109350532)

[5.2 Summary of achievements 34](#_Toc109350533)

[5.3 Reflection 34](#_Toc109350534)

[5.4 Research recommendation 35](#_Toc109350535)

[5.5 Achievement of objectives 36](#_Toc109350536)

[5.6 Further work 37](#_Toc109350537)

[References 38](#_Toc109350538)

**Table of Figures**

[Figure 1: Importing and Installing the Libraries in Jupyter Notebook 23](#_Toc109062702)

[Figure 2: Importing the dataset into Jupyter Notebook 24](#_Toc109062703)

[Figure 3: Visualization of the dataset in Jupyter Notebook 24](#_Toc109062704)

[Figure 4: Checking for null values in the dataset in Jupyter Notebook 25](#_Toc109062705)

[Figure 5: Bar plot of the Fraud column in the dataset in Jupyter Notebook 25](#_Toc109062706)

[Figure 6: Division of the dataset into X and Y in Jupyter Notebook 27](#_Toc109062707)

[Figure 7: Visualization of X division in Jupyter Notebook 27](#_Toc109062708)

[Figure 8: Visualization of Y division in Jupyter Notebook 28](#_Toc109062709)

[Figure 9: Division of Dataset into Train and Test data in Jupyter Notebook 28](#_Toc109062710)

[Figure 10: Implementation of decision tree classifier and the accuracy score 29](#_Toc109062711)

[Figure 11: Confusion Matrix 29](#_Toc109062712)

[Figure 12: Tree plot of the dataset](#_Toc109062713) 30

[Figure 13: Results after performing the ANN in Jupyter Notebook 30](#_Toc109062714)

[Figure 14: Development of the confusion matrix in Jupyter Notebook 31](#_Toc109062715)

[Figure 15: Final heat map after completing the machine learning in Jupyter Notebook 31](#_Toc109062716)

# Chapter 1: Introduction

## 1.1 Introduction

In the current scenario, the application of the credit card has evolved and is majorly used all over the world. This has created a significant problem requiring special attention for the fraud detection of the credit card which is a huge problem and the ethical problem of the scenario. It has been generally understood that the different types of credit card fraud which lead to the detection of the fraud issue with the credit card and major problems. A credit card has used different types of techniques for detecting fraud using the python language which is widely used in machine learning. The platform that has been used is the jupyter notebook. The technique that has been used for the detection of the fraud of the credit card uses the decision tree and ANN which is the artificial neural network. These techniques are used for the training of the model and testing the model which can detect suspicious activity and potential fraud.

## 1.2 Project Specifications

In this research study, the detection of the fraud credit card has been implemented using the python programming language which has been done in the notebook platform. As per the view of John and Naaz (2019), the dataset has been used to perform and understand detecting suspicious activity using machine learning techniques. It is specified for identifying the various suspicious activity and learning different detection techniques for the identification of fraud activity using machine learning. However, it enhances the process of payment for analyzing and recognizing the high authentication level and high level of accuracy for the better prediction of the accurate result. This technique is used to reduce the issue of fraud for the credit card holder or owner which can increase and enhance the opportunities for credit card payments. Where machine learning uses the decision tree for the identification of suspicious activity while model training is required to understand and prepare different techniques for analyzing and tracking the different activities of the detection of the fraud for the credit which can obtain a high level of accuracy using the model. The model has used different detection techniques which are the neural network, back propagation, and decision tree which help in the detection of the fraud potential. The artificial neural network is generally used to identify fraud activity where the performance is measured by obtaining the level of accuracy for the prediction. The classification of the process is used where the k nearest neighbor and support vector machine has been used for the detection of credit card fraud activity. These three algorithms helped in the classification process and detection which supports the artificial neural network in detecting fraud of the credit card. This system will show and give 0 or 1 of the attribute where the model has analyzed the fraud detection. As per the view of Arya and Sastry (2020), the artificial neural network generally works like a human brain where the different systems are used for connecting the different nodes of the neurons similar to the artificial network.

## 1.3 Aims and objectives

**Aims**

The aim of the research study is to detect the fraud activity of credit cards using the artificial neural network and decision tree.

**Objectives**

* To authenticate and recognize the details of the credit card holder or owner and analyze the pattern of credit card usage using machine learning techniques.
* To identify the distrustful activities and analyzed the fraud activities and ensure along with the report or alert to the credit card holder; detecting fraud and suspicious activity.
* To reduce the fraud potential and loss of payment of the credit card system using the decision tree and the ANN model.
* To obtain a high level of accuracy that can detect the high potential of fraud activity using the supervised algorithm in machine learning like a decision tree and artificial neural network.
* To create or develop a model for training that can predict the different classes of the target variable by using the learning technique in the decision tree algorithm.

## 1.4 Current issues

In the research study the issue that has been seen in the artificial neural network where the major advanced techniques and extreme part has been contributed in the field of technology which has also resulted in the increasing the threat and a high potential for data exploitation. This issue can create major complications and affect the expected result which can bring challenges to the survival of effective technologies. As per the view of Safa and Ganga (2019), the artificial neural network plays a dual role as defensive and as well as offensive which can be attacked by different cyber threat activities. This cyber threat generally takes the benefit of the weakness of the credit card which can be identified and detected as an effective way for exploitation which can hamper and increase the threat risk. Credit card fraud is generally increasing due to an increase in data leakage which attacks take the advantage of exploitation. In the present situation, credit card fraud detection can have low availability of the data which leads to a decrease in the efficiency of the model which cannot predict the desired result. The mode generally is used to detect current issues for the specified factor on fraud detection.

The model is required to prepare with the effective technique for training and testing the data which can make it better for reducing the risk and increase the utilization of the credit card and enhance credit card fraud detection. The low dataset cannot show the effective result using the model. The training of the low dataset model could only reflect the lower potential and lead to a decrease in the potential of the model. Credit card fraud detection could result in facing major problems where the detection of the potential fraud is generally due to mismanagement while preparing and data training process for the credit card fraud detection.

.

# Chapter 2: Research

## 2.1 Quality of background research

Credit card fraud is a process that identifies the attempts of purchase that can be “fraudulent " by rejecting the fraud instead of carrying on with the order (Varmedja*et al*. 2019). There are many techniques and tools that are available to detect fraud through the merchant’s employ a combination of many of them. Fraud detection through machine learning from a model of machine learning and a data set of the transaction of the credit cards in order to train the model so that it can recognize the patterns of fraud. It is a self-learning model that enables one to adapt to new as well as unknown patterns of fraud. The main aim of this study is to predict if the transaction of a credit card is fraudulent or not, that is based on the transaction data of the credit card, and it can be detected through anomalies in the transaction of the data. ***“Payment fraud detection”*** is one of the most common types of fraud that is tackled by ***“Artificial Intelligence”.*** The variation is diverse as the imaginations of the fraudsters, these are some of the common types of payment fraud such as stolen cards, lost cards, card ID theft, and others (Yousefi*et al*. 2019,p.907). The technology is developing daily and the increase of credit card fraud is also increasing, it cost both consumers and businesses through these frauds it causing a loss of about billions of dollars a year. The fraudsters are looking for new ways to commit these crimes and activities. But it has become necessary for the business and financial institutes to build an advanced system to detect any kind of fraud and to encounter such threat of credit card fraudulent transactions and to identify the theft.

## 2.2 Use of literature

Credit card fraud is very common among thieves because the culprit can steal in a short time a large amount of money (Dornadula and Geetha, p.201). This kind of fraud takes time to be discovered and the risk to catch the theft is minimal as compared to other crimes. The fraud of credit cards occurs online and offline and in many ways. The fraudsters commit fraud online via the internet or phone with the help of the card details and offline frauds are committed when a card is stolen and is physically used by the fraudster. There are different types of ***“credit card fraud”.*** False information that is given in the application is known as application fraud, it is classed as a duplicate application when such applications come from the same user and if it is from a different user, it is identified as fraud (Tennakoon*et al*. 2019, p.176). If a stolen Credit card or lost and lands over a fraudster then Offline fraud can be committed. There are various methods that are used by thieves to pick-pocket customers in very busy urban areas. The thieves distract the commuters on trains, and buses and steal their credit cards.

A fraudster when obtaining the personal information of a customer is known as an account takeover, by it takes over the account by providing the details of the account. The thieves target the trash can of households and retrieve the information of the account. Most of the thieves also target the mailboxes in the building when the pins of the cards are posted to the customer. The cards that are scanned without the permission of the cardholder is known as ***“counterfeit card”.***  The fraudsters used sophisticated machines to make cards, but it is difficult nowadays since more features of security have been developed (Maniraj*et al*. 2019). Techniques used by the fraudsters like changing the details of the card, there are cases where credit card fraud uses methods of skimming that copies the data that is extracted from the card electronic strip.

The bills for such activities of fraud are not paid by the consumer but by the issuers of the banks and cards where most of the cost of the fraudulent activity is covered that is in the terms and service.  A huge amount of loss where the card issuers have developed a sophisticated system in order to protect themselves so that it detects the fraud at the earliest as possible (Sailusha*et al*. 2020). These kinds of system help the customer conduct their transaction of “day-to-day business” and also flag transaction that is unusual. There are two categories where the technique of Statistical fraud detection can be performed: "***supervised and unsupervised”.***

The supervised method is the model that can classify the transactions that are new if it is legitimate or fraud that is based on the previous transaction. Whereas in the supervised method the unusual transactions can be identified as fraudulent transactions. These two methods are used to calculate the probability of fraud in any transaction (Khatri *et al*. 2020, p.144).

There are many challenges that are needed in credit card detection, it completely relies on the study of the data to detect these types of fraud and most of these data are unavailable even in the financial institution and banks because of the personal nature and also because it is sensitive. There are a number of transactions that happen every day, and the analysis has issues when it comes to information technology. The technique for fraud detection is developed and has become very sophisticated which makes the fraudster change new methods to achieve the goal and perform the crime (Patil and Lilhore, 2018). The techniques that have been used to detect credit card fraud in this research are the “Decision Tree and Artificial Neural Network”

**Decision Tree:** A decision tree is a common tree that helps for prediction and classification; it analyzes the decision that can be used to represent the decision-making. In classification models that are based on a decision tree that is applied to the problem of detection of credit cards. Thus if these approaches are implemented in fraud detection systems it helps to develop the data understanding. Lass in financial because of fraudulent transactions can be drastically decreased (Ryman-Tubb *et al*. 2018, p.123). The method of the Decision Tree works when it uses a similar kind of tree that is created by using the logic of a decision tree. Nodes are outlined on the similarity tree and the leaves have the factors and attributes. Decision Trees are supervised learning methods that are used for regression and classification, that are used to create a model that allows predicting the target with the help of simple learning rules of decision that is inferred from the features of data. The detection of credit cards is a problem for classification, in this study the credit card detection is done with the classification algorithms in python.

**Artificial Neural Network:** it reflects the behavior of a human that allows the programs of a computer so that it can recognize patterns and can solve problems that are in the field of machine learning, deep learning, and AI (Yee *et al*. 2018). ANN is a technique of machine learning that is used in classification problems, it is also a set of connected networks of input-output, where the weight is connected to each of the connections. It has “***one output layer, one input layer, and one or more intermediate”.***

The artificial neural network is used to find fraud in the transaction of a credit card. The performance and as well as the calculation of the accuracy are based on the prediction. The data for processing the toll with the help of ***“high classification accuracy”*** and also can process the data in parallel form. The data are stored in a distributed manner that can learn the data. Using the ***“artificial neural network”*** gives an approx. of 100 percent accuracy for the detection of credit cards (Itoo and Singh, 2021). The artificial neural network is a deep learning concept that is implemented where three layers are used, the output layer is considered as the final slayer where the output is classified, it is either 1 or 0 where 1 helps to indicate the case of fraud, and 0 indicates the normal. In the data process where the libraries are used for processing the data, the operation is applied through featuring scaling. The libraries that are used for the visualization are Matplotlib and the plot of the histogram of the class columns (Raghavan and El Gayar, 2019). The framework is used for the neural network where it gives the model parameters of the study as well as the model summary and compilation. The libraries that are used for the model evaluation are the confusion matrix and Classification Report, the accuracy score is also retrieved along with the infusion matrix and Classification Report.

## 2.3 Critical assessment

The study is to develop a credit card detection system with python using the decision tree and the artificial neural system. That has been implemented using the python programming language which has been done in the notebook platform. Using the machine learning technique, the dataset has been used to perform and understand to detect suspicious activity. The identification of suspicious activity and performing various detection techniques is to identify the activity of fraud using machine learning. Enhancing the process of payment and analyzing and recognizing the high authentication level and high level of accuracy for the better prediction of the accurate result of both the decision tree and ANN.  Credit card fraud detection is the most important for all banks or any financial organization, through machine learning the problem of detecting credit card fraud is solved using Sklearn for the decision tree and Keras for implementing the neural networks and the python libraries. The use of the literature in this study helps to develop the understanding and awareness of the fraud of credit cards that occurs both online and offline and in many ways. The critical assessment help to study systematic the risk and result that has been discussed in this study.

The problem can be overcome with a binary classifier created and thus by experimenting with the machine language technique in order to see better results. The data set also consists of parameters and the libraries and frameworks in the detection of credit card are developed with python. A credit card fraud classifier is developed in machine learning by performing. Data Analysis is in the dataset by applying the different “Machine learning algorithms” in the dataset. The models are trained and evaluated on the dataset, where the best one is picked. Using the matplotlib the number of all the class labels is checked and the information is plotted. The algorithms of machine learning are applied to the dataset of the credit card, in this machine learning research the decision tree classifier and artificial neural networks are built to see the best performance model. The models are then trained and evaluated on the dataset and the best one is picked.  The confusion matrix is visualized and the evaluation matrix of both the two methods such as the decision tree and ANN. The confusion matrix of this study in python is displayed by importing the method of plot\_confusion matrix and the array of the parameter is done by passing the confusion matrix. The study of python machine learning is performed to build the binary classifier with the help of a decision tree and ANN algorithms to detect the transactions of credit cards detection. The accuracy that has been retrieved in this study is for both the decision tree and the artificial neural method. The different objectives that are linked to aims have also been analyzed and discussed in this study.

## 2.4 Linkage to aims

**Authenticating and recognizing the details of the credit card holder or owner and analyzing the pattern of credit card usage using machine learning techniques.**

Machine learning is an AI that allows the application of the software to predict an accurate outcome that is not explicitly programmed. The historical data as the input is used by the machine language so that it can predict the values of the new output which can obtain authentication and recognition from the fraud activities. Machine learning focuses on the data that are used and algorithms that are used to imitate how the human learns and slowly it improves the accuracy.

The world is based on the payment methods of card payments and online and the digital market is rising worldwide (Trivedi *et al*. 2020). The growing problems due to cybercrime are also increasing at the same time at a faster rate. Credit card detection is a priority for all financial organizations and banks, and it is possible with the help of machine learning, that it becomes much more efficient and easier to track the patterns and the preventions of abnormal transactions. There are various techniques of credit card fraud the following are some of it

* Using the legitimate information of the credit card creates a credit card counterfeit.
* Using a new credit card and creating an account by using the name of someone else
* Stealing the existing holder's account or a credit card.
* Skimming the shopper's credit card at the time of swiping thus makes a copy fraudulent.
* The payer is manipulated by the fraudsters to make multiple transactions using credit cards against the same purchase
* Via email phishing, the fraudsters take over the details of the cardholder's account

**Identifying the distrustful activities and analyzing the fraud activities and ensuring along with the report or alert to the credit card holder for detecting fraud and suspicious activity**

The fraud is detected by the victims when they discover the fraudulent activity following are the signs and warnings of such activity to identify such suspicious activity

Monitoring the account frequently is very necessary, online accounts should be viewed so that they can detect the fraud earlier and the financial institution should be contacted immediately if there is any suspicious activity (Jain *et al*. 2020). Using online alert tools and services, by signing up for text and email alerts so that it can notify whenever such suspicious event occurs. It also helps to notify whenever there are low balances in the account and if there is an unauthorized transaction of a huge amount made. Thus, if the alerts are set it helps to signal the user and it can help to stop such activity. Using a monitoring service of credit cards, signing up to the service notifies the users when there are any changes in the credit card report that are posted.

**Using the decision tree and the ANN model in order to reduce the fraud potential and loss of payment of the credit card system**

The decision tree machine learning models are used in this research for financial  Fraud Detection, along with development in information Technology where the spreading of fraud is also rising all around the world causing huge losses in the financial. There are difficulties to detect the fraud and detecting such fraud requires studying the data. There are a number of transactions every day, thus the detection of credit card fraud can be done by using the method of the decision tree. This defines the ratio of the transaction that helps to satisfy the conditions, this method can be easily displayed and comprehended (Jain *et al*. 2020, p.188).

ANN is a machine language that is supervised that is inspired by the human brain, the structure of ANN has ***“one input layer, one output layer, and one hidden layer”.*** The size of the input layer is based on the features of the number of the dataset that is given in this study. The size of the hidden layer varied that is based on the complexity of the task. The size of the output layer generally depends on the problem type that is about to solve. The neuron is considered the basic component of ANN which is a node.

**Obtaining a high level of accuracy that can detect the high potential of fraud activity using the supervised algorithm in machine learning like a decision tree and artificial neural network**

The high level of accuracy that has been detected in this study is the high potential of fraud activity using the supervised algorithm in machine learning in the decision tree (Lucas and Jurgovsky, 2020). The ***“Supervised learning technique”*** is used in this research in the detection of credit card fraud where the patterns of fraud can detect by analyzing the transactions of past details. The data set has been properly cleaned and all the null and redundant values from the data set. The data set has been properly visualized to detect all the patterns in the data set.

The accuracy of the decision tree is found with the help of the number that is predicted correctly and that is divided by the predictions of the total number that are made. The majority of the class is to predict the association with the help of a particular node that is true and by using the value that is a larger attribute from each of the nodes.

The accuracy of the average ANN classification model indicates the rate of the highest accuracy of the classification results. The neural network is the accuracy is equal to the ***“number of predictions that are correct”*** of the predictions total number. The accuracy of the binary classification is calculated in terms of the negative and positive. The accuracy of the ANN model can check the value of the regression, the error of the histogram, and others. The accuracy that can be detected with the high potential of fraud activity in the supervised algorithm of the artificial neural networks

**Creating or developing a model for training that can predict the different classes of the target variable by using the learning technique in the decision tree algorithm**

The model is created and developed for predicting different classes of the target variable by using the learning technique in the decision tree algorithm such as the implementation of SKlearn which is termed the decision tree used for the regression and classification task. This research is to create a decision tree classifier with the help of ***“Sklearn and Python”.*** The decision tree is the “supervised machine learning algorithm” that helps to classify the data and the accuracy that is retrieved in this study for the potential of fraud activity using the supervised algorithm in machine learning. The Sklearn matrix is imported to obtain the accuracy score and the confusion matrix of the decision tree. The confusion matrix is to predict and visualize the importance of analytics such as the ***“accuracy, specificity, and precision”*** of the confusion matrix. Credit card fraud detection is the most important for all banks or any financial organization, through machine learning the problem of detecting credit card fraud is solved using the Sklearn for the decision tree and Keras for implementing the neural networks and the python libraries. The use of the literature in this study helps to develop the understanding and awareness of the fraud of credit cards that occurs both online and offline and in many ways.

# Chapter 3: Methodology

## 3.1 Choice of methods

In the research methodology, the model has been prepared using the python programming language where the model uses a classification process with the help of an artificial neural network and decision tree. As per the view of Vijayakumar *et al.* (2020), the fraud detection of the credit card leads to preventing the credit card holder from fraud activity and enhances privacy for it. The dataset has been used for the training of the model using the dataset where the deep learning and artificial neural network, decisions tree. The supervised and unsupervised algorithms and techniques have been implemented in machine learning and deep learning. Supervised learning has been used due to increase the detection level for the fraud of credit cards where the input is taken and generates or produces a result as an output. Unsupervised learning has been used due to training the model for the identification of the model pattern of the data where the group of data is observed for the identification of the problem and classification process of the similar pattern or data for the credit card detection.

Initially, the input is taken as a piece of information or the details of the credit card holder along with the transaction ID details which can help in the better prediction of the fraud activity using machine learning. The classification of the model is done using the perceptron algorithm, k nearest neighbor algorithm, and SVM. The perception uses the binary perceptron where the selection of the feature is done for the classification of the problem. The artificial neural network has used a three-layer for the detection of the fraud activities where the initial stage is the input layer where the input is given for the model training process which uses the information of the credit card holder and transaction detail for processing the model for training. The second stage or layer is the hidden layer where the process and the combination are established and different neurons are created by the combination using the input function. Finally, in the third stage, the output is analyzed and predicts the desired combination results. As per the view of Varmedja*et al.* (2019), the model required analyzing and reading the data at the beginning of the preparation model which can highly reflect the result for the credit card fraud detection. Initially, the importation of the dataset and libraries is required for the better detection of the problem.

The detection technique that has been used is a neural network, back propagation, and decision tree. The common technique which has been used in the research study project is naive Bayes (NB), k nearest neighbor technique, and support vector machine (SVM).

## 3.2 Justification and support of choices

The technique has been used for the detection of credit card fraud detection. There are several reasons for the detection using neural networks, back propagation, and decision tree.

**Neural network:** ANN or artificial neural network generally delivers effective strategies for learning the real values for the credit card data which generally uses the functional value of the vector and discrete. As per the view of Awoyemi*et al.* (2017), the different methods has been used for the learning and training of the model which passes from the different layers of the neural network where the different layer generally consist of the effective method and training of the data model into the different level of the layers where the input, output, and hidden layer is effective for the model. In the input layer, the information of the data is collected where the dataset is analyzed for the detection process of the fraud activity. The hidden layer generally includes the calculation section where a certain parameter is set which is required to be fulfilled for the proceeding or pass the parameter and reach for the output stage which can result in the output and obtain a higher level of accuracy. The propagation of the takes place or flows backward due to errors in the compilation and errors in the data where it reflects back as the error output after the compilation of the code.

**Back propagation:** This algorithm helps in the understanding and learning of the different consequences of the multi-layers of the network which generally fixes the set for the interconnection and the units. The process is generally propagated and the supervised learning is a part of the neural network that uses the gradient fall technique. The function of back tracking and error of the neural network shows the error and different problems in the output layer. The back propagation checks and compared with the outcome of the output with an error, if an error is detected then it propagates for the correction and else has to pass the parameters to obtain a higher level of accuracy for the model.

**Decision tree:** the algorithm has been used for the classification of the using the library of sci-kit learn. The method will be used for the mathematical operation for the detection of the credit card fraud where the data is used in the testing process for the decision-making process and classification of the different prevention of the fraud activity.

## 3.3 Data collection

In the research study, the collection of the data has used a primary understanding of the problem where the detection occurs the problem like the low data analysis system capability real values analysis, dataset efficiency, and many other problems have been analyzed for the credit card fraud detection. In the primary problem, it has been seen that there are major issues that could not enhance the performance of the model, a decrease in the productivity of the model which is not able to adopt new technologies and change according to the trends and can get exploited by the various attack. As per the view of Dornadula and Geetha (2019), the model can affect the different potentials using the different algorithms which can help the model in adapting the technique for the detection of the fraud and prevention of the risk associated.

In the secondary collection of the data, different research papers and data analysis of credit card fraud have helped in understanding the risk involvement required attention for understanding credit card detection. The different research papers are used for understanding the process for the detection technique which can help in the development of different methods for increasing the performance of the detection of the fraud of credit cards using python language in machine learning and deep learning. The data collection has been done which will be helping in understanding and enhancing the performance of the research.

The collection of the data can help in the better prediction of the model which can help in the development of the better prediction model using the python programming language where python is majorly used in the field of machine learning. The data collection has been collected from accessing the public library where huge data was used for understanding and clarification of the data. The major information is collected from the available resource on the internet. The available information helps in developing the system for credit card fraud detection. The model is generally prepared with the help of data resources and collected data which can help in understanding and increasing the performance along with obtaining a high level of accuracy for the model.

## 3.4 Validation

This research part, this section describes the implementation, including the algorithm used to put the suggested system into practice. As per the view of Trivedi *et al.* (2020), the loading of the dataset is where Implementation in this paper begins. Then, data pre-processing, which includes data cleaning and normalization, is carried out. The dataset is divided into two datasets for the training and testing of the model. In the end, the system determines if a transaction is fraudulent or not.

In this area of the research study, the elements that are crucial for conducting the steps for this particular work are covered. This section discusses data analytics, research plan, research strategy, and research philosophy. This chapter has also sparked a discussion about the research design. This chapter also covers data collecting from the appropriate sources and data analysis. To conclude the research project in a given amount of time, this component also covers several research methods, research ethics, and an appropriate plan.

The model validation is been verified with the capability of the model which can be understood for increasing the performance. The model generally uses the two processes for checking the validation for training the dataset and testing the dataset of the model. The preparation of the dataset is required for the training process model which can be helping to obtain high accuracy within the model. It required multiple time training using the dataset which can help in increasing the performance of the model. The testing is generally done for the model which can check the capability of the model which can obtain better performance for the prediction of the result where the credit card can detect suspicious activity where the data is understood and analysis for before predicting the result which can enhance the different objective for analyzing the suspicious activities. The model generally uses the prediction using the dataset which can use the predictive method and technique in the model for analyzing the fraud of the credit card where the card is used in different places like shopping, and other various activities in the market.

## 3.5 Ethical issues

The ethical consideration of a research study is important since it necessitates the application of numerous ethical norms and principles to complete the research study. when acquiring data from the appropriate information, the researcher followed the appropriate ethical guidelines. Any illegal method was not used to obtain any information. Additionally, the researcher put in place the required security measures to protect that no unauthorized user utilizes the data for any illegal purposes since the dataset includes a variety of personal and sensitive data about individuals. The data was utilized for the creation of a system that will assist in identifying the sensitivity analysis of the used model. As per the view of Husejinovic (2020), the ethical issue is majorly concerned with the model which should provide the basic security within the model and ensure better prediction.

## 3.6 Commercial risk and risk management

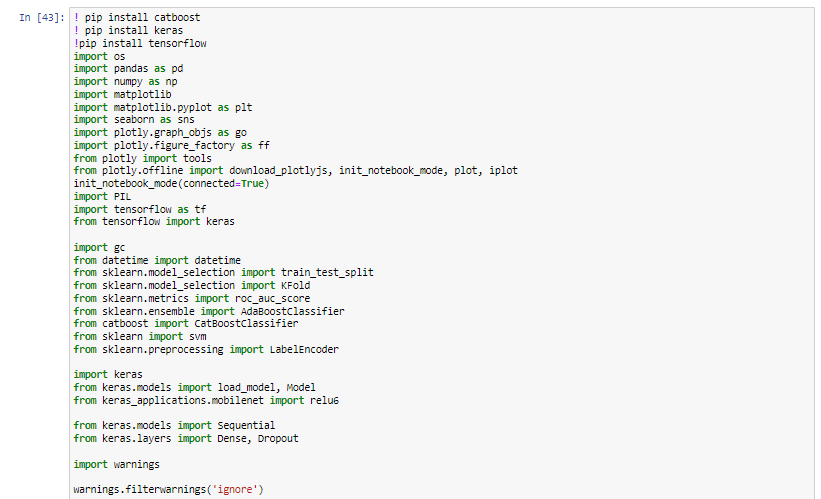
In the banking world, most fraud happens in credit cards transaction and so it becomes a risk for the bank as well as the customers. There are various ways that have been already implemented for credit card fraud detection. Using the Python programming language in this project the data analysis of the credit card fraud detection has been done to predict the type of fraud. This data analysis can provide bank systems to get an idea of fraud and machines can be able to detect it in a faster way. As per the view of Lenka *et al.* (2021), the commercial risk of credit card fraud detection is that it requires all the personal details of the customers of the bank to proceed with the analysis. There is a risk for the banks to proceed with this kind of job since during the data migration the personal details of the customers can exploit.

# Chapter 4: Result

## 4.1 Critical analysis

The development of the fraud detection system has been done in the software language python using the Jupyter notebook as the platform. The development includes the implementation of the decision tree algorithm and artificial neural network ANN for the analysis of the system. The use of a decision tree helps in the structuring of the used dataset that has been used for the analysis. The decision tree algorithm helps in the splitting of the features of the used dataset of credit card fraud. The decision tree has been used for the study of irrelevant features and removing them with the help of the process known as pruning (Patil and Kulkarni, 2019). The use of a decision tree helps in the evaluation of options that are related to the several courses of action that can be chosen. The decision tree provides a structure that is highly effective with the help of which the dataset can be investigated and possible outcomes have been chosen. The implementation of ANN in machine learning has been implemented on the dataset so that the data can be analyzed with close similarity to the thinking of the human brain. The human brain replication and its way of thinking and analysis have been performed on the dataset so that in-depth analysis of the data can be done.

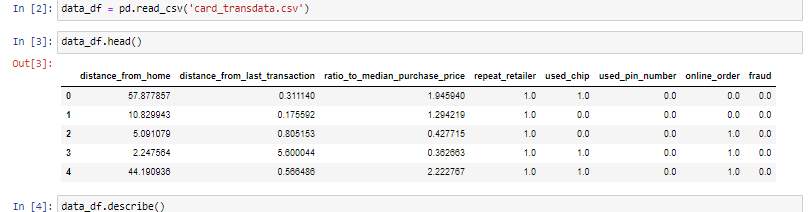
## 4.2 Evidence of practical work

****

###### Figure 1: Importing and Installing the Libraries in Jupyter Notebook

(Source: Self-Created)

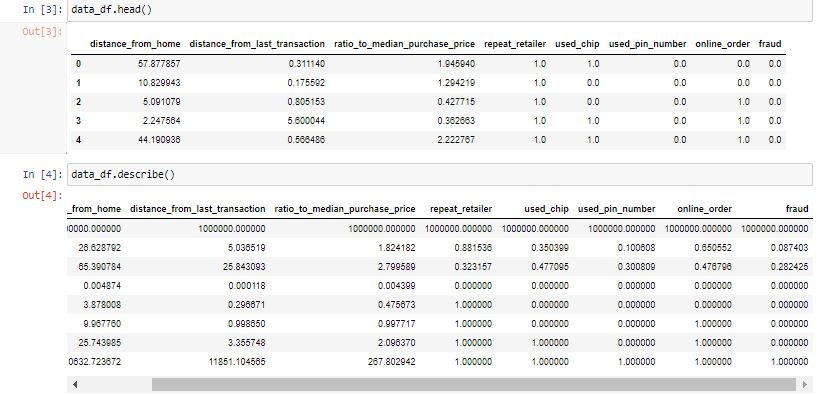
The image above shows the installed and imported library that has been used while performing the analysis of the dataset that deals with fraud detection data of credit cards.



###### Figure 2: Importing the dataset into Jupyter Notebook

(Source: Self-Created)

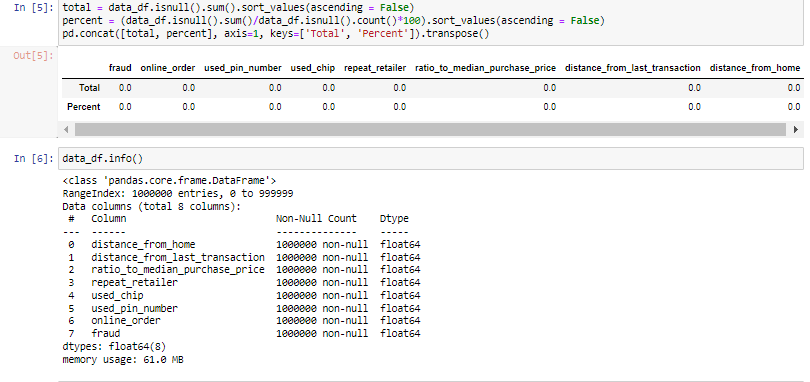
The import of the. CSV file that contains the data of the credit card fraud has been imported and the heads of the data have been visualized in a jupyter notebook.



###### Figure 3: Visualization of the dataset in Jupyter Notebook

(Source: Self-Created)

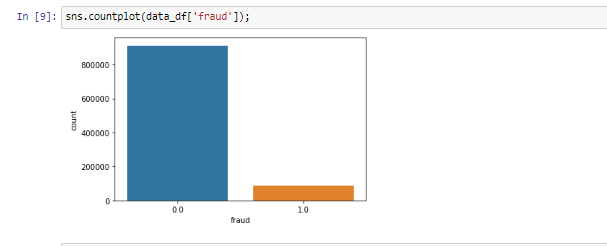
The visualization of the imported dataset in the jupyter notebook has been done with the help of the command data\_df.describe() which helps in the data visualization process.



###### Figure 4: Checking for null values in the dataset in Jupyter Notebook

(Source: Self-Created)

The above image shows the information that helps in the detection and presence of null values in the dataset and helps in the removal of the null values. The dataset used has been found to have no traces of null values, and no removal of null values has been required.



###### Figure 5: Bar plot of the Fraud column in the dataset in Jupyter Notebook

(Source: Self-Created)

The graphical representation above shows the presence of fraud transactions and the non-fraud transactions that are present in the dataset. The 0 column represents the non-fraud plot of the data, whereas the 1 shows the fraudulent transactions that have been done from the dataset.

## 4.3 Awareness and solutions to the technical challenges

The current performing work is developed on the topic of machine learning and its execution in the industrial field as per the demands. The current world is looking at the enhancement or improvement in the existing technology to improve their respective lifestyle as per the demands (Dornadula and Geetha, 2019, p.631). The execution of the machine learning technologies into an existing system is helpful to enhance the performance of those old technologies in the industrial field.

The old technologies have some issues while implementing that but the execution of the machine learning in those technologies is helping to improve the accuracy level to an acceptable limit as per the demands of this current work. The current performing work is developed to detect fraudulent activities in a credit-card system as per the demands. The current work is developing a prediction model to detect frauds from a dataset of credit-card by using the concept of machine learning. The execution of machine learning is helpful to predict those available fraudulent activities with an acceptable accuracy level.

An acceptable accuracy level is stating the success of executing machine learning as per the demands. In this current work, the technology regarding machine learning is executed through the execution of the decision tree and ANN. The concept of ANN or “Artificial Neural Network” comes from the structure or the working procedure within a human brain as per the demands. The execution of the ANN is helpful to train a system to learn like a human brain, think like a human brain, and make a decision like a human brain as per the demands of this current work (Najadat *et al.* 2020, p.204). A developed prediction model is capable enough to make a decision with an accuracy of above ninety percent but sometimes failed to deliver accurate results.

The working progress of a system with machine learning is completely based on the training datasets, in case of any unknown data type, the system cannot able to provide accurate results and this is the biggest issue while executing machine learning in a system. The execution of machine learning is also demanded an upgraded system to deliver the results as per the demands. Sometimes machine learning is also demanded upgraded knowledge to execute this in the industrial platform. Improper knowledge of the technology might create a big issue for the users. The use of machine learning is only helpful in understanding the predictable outcome only, and it does not apply to making a decision on a crucial topic.

## 4.4 Novelty

In this current work, the technology of machine learning is executed to develop a prediction model to detect fraudulent activities in a credit card system. In this work, the execution of the machine learning is performed to train a system capable to make a decision in certain circumstances as per the demands. The current work is working on the concept of ANN to train a system to perform like a human brain as per the demands of this current work (Varmedja*et al.* 2019, p.1). In this work, the system is first trained regarding some similar scenarios and after that, the testing part is performed to display the accuracy level after executing the machine learning as per the demands.

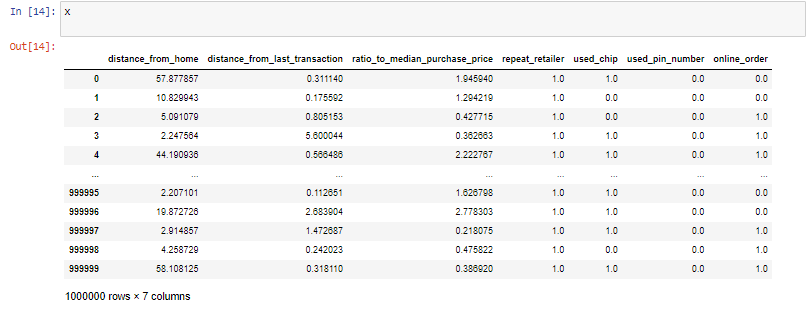
## 4.5 Interpretation of results



###### Figure 6: Division of the dataset into X and Y in Jupyter Notebook

(Source: Self-Created)

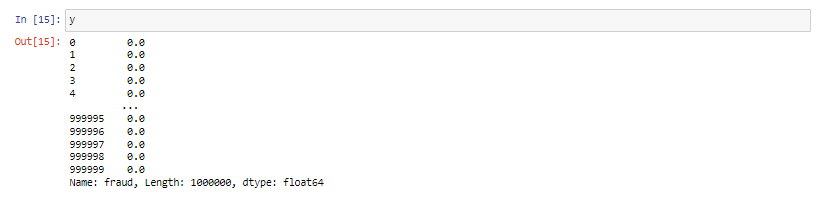
The dataset here has been divided into X and Y variables so that the analysis of the used dataset can be facilitated.



###### Figure 7: Visualization of X division in Jupyter Notebook

(Source: Self-Created)

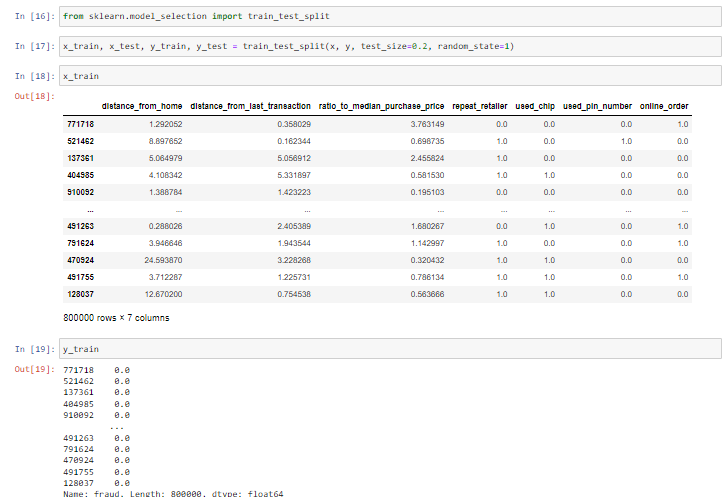
The above image shows the X division of the dataset in a tabular format that helps in the visualization of the X division of the dataset.



###### Figure 8: Visualization of Y division in Jupyter Notebook

(Source: Self-Created)

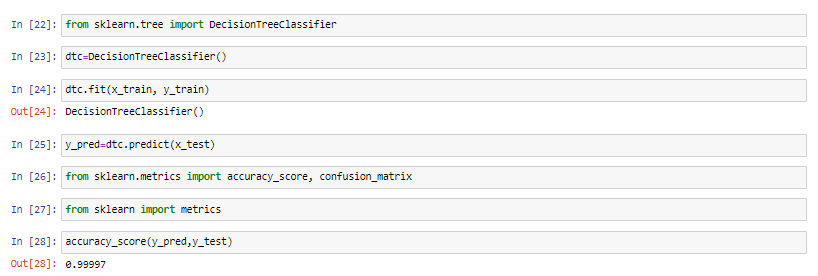
The above image shows the y division of the dataset in a tabular format that helps in the visualization of the Y division of the dataset. The Y division contains only the fraud column from the dataset that has been dropped from the X division.



###### Figure 9: Division of Dataset into Train and Test data in Jupyter Notebook

(Source: Self-Created)

The above image is the representation of the division of the dataset of credit card fraud into train and test datasets. The train\_test\_split function that has been imported helps in the splitting of data into training data and testing data that have been used for the analysis of the dataset.



###### Figure 10: Implementation of decision tree classifier and the accuracy score

(Source: Self-Created)

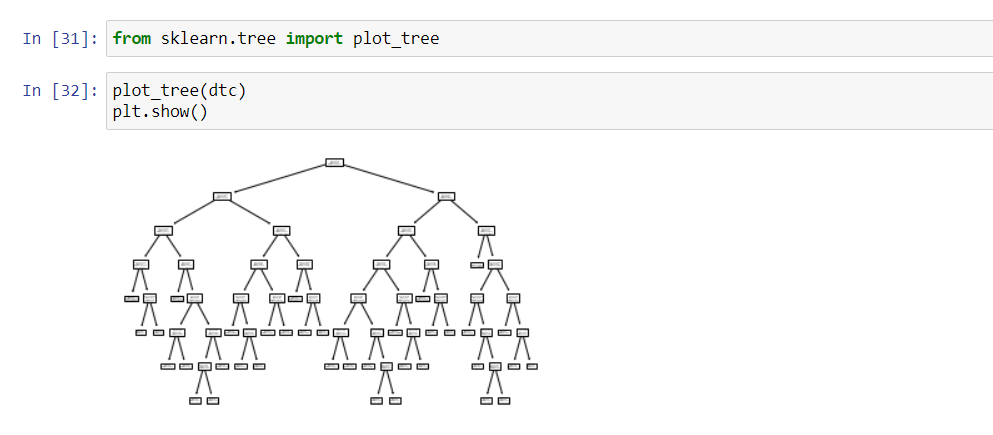
The above image shows the implementation of the decision tree classifier on the training dataset, and the prediction model has been created based on the predictor data. The decision tree algorithm helps in making decisions based on the analysis of the dataset. The accuracy score of the decision tree analysis has been found to be 99% which makes the created model very accurate.



###### Figure 11: Confusion Matrix

(Source: Self-Created)

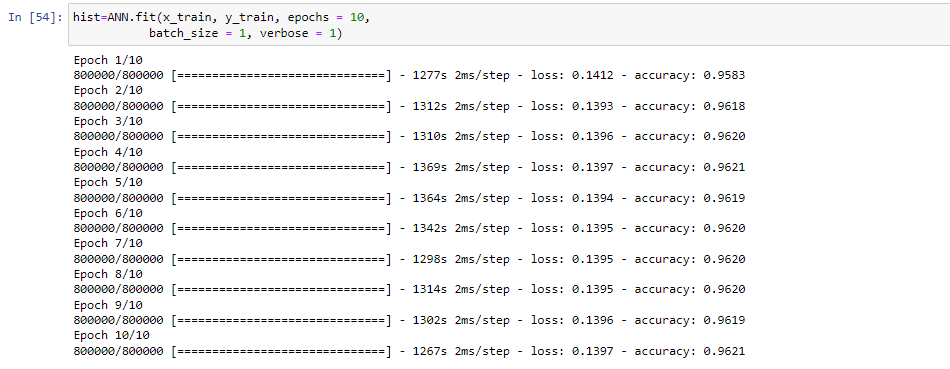
The above image shows the confusion matrix of the decision tree classifier helps in the determination of the performance of the test dataset of the analysis. The accuracy and precision, specificity of the created model have been determined with the help of the confusion matrix.



###### Figure 12: Tree plot of the dataset

(Source: Self-Created)

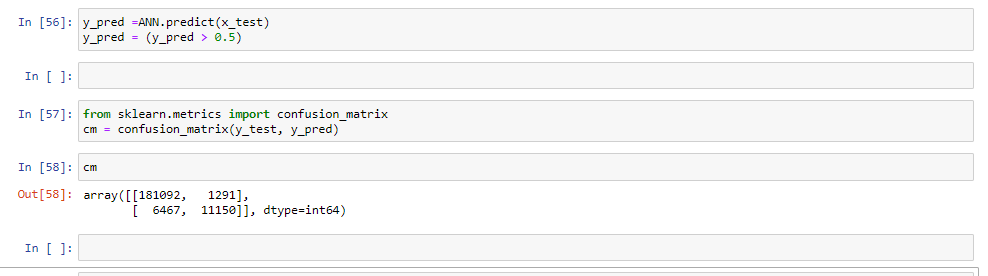
The above image is the output result of the tree plot of the decision tree classifier that has been implemented on the dataset. The model helps in the determination of important branches and helps in the segregation of the non-predictive and non-useful branches of the decision tree model.



###### Figure 13: Results after performing the ANN in Jupyter Notebook

(Source: Self-Created)

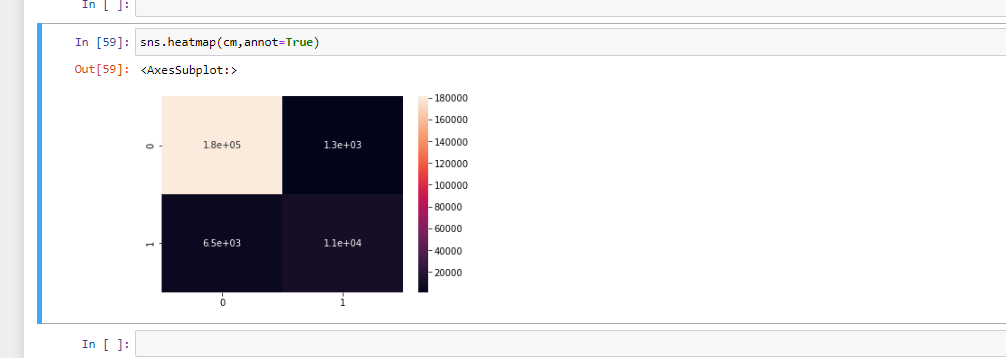
In the above diagram, the ANN is performed for 10 epochs, and for all the 10 individual cases, the developed model is capable enough to predict the required model with an accuracy of above 95% and this is acceptable in the industrial field. The high accuracy level is indicating the success of this research work as per the demands.



###### Figure 14: Development of the confusion matrix in Jupyter Notebook

(Source: Self-Created)

The above image shows the representation of the confusion matrix of the ANN algorithm that has been implemented in the credit card fraud dataset. The confusion matrix helps in the prediction of accuracy and precision factors of the ANN algorithms.



###### Figure 15: Final heat map after completing the machine learning in Jupyter Notebook

(Source: Self-Created)

The above diagram shows the final heat map after completing the machine learning as per the demands.

## 4.6 Use of tools and techniques

The use of the Jupyter notebook which works in python language has been used in this machine learning on the credit card fraud detection system. The use of two different algorithms has been used on the procured dataset so that pattern detection of the dataset can be done (Otchere *et al.* 2021). The decision tree classifier and the artificial neural network has been used in the analysis of the dataset for the detection of fraud and the study of these fraud patterns. The use of Microsoft Excel has been done so that the raw data can be visualized and checked for corrupt data.

## 4.7 Appropriate tools for analysis

The tools that have been used in performing the analysis in python software are Pandas which are structured for operations on data and manipulation of data. The use of matplotlib has been done for plotting and visualization of various types of graphs and plots that help in the result analysis of the data. The use of seaborn has been done for the facilitation of statistical data visualization that has been obtained from the analysis of the dataset as result. The use of Numpy has been done on the dataset that enables the features of fundamental computing in a scientific manner and helps in the implementation of an N-dimensional array.

## 4.8 Linkage to objectives and literature

The use of machine learning in credit card fraud detection has been facilitated by the implementation of machine learning on the historical data of credit card fraud data. The detection of fraudulent transactions based on the data and the analysis helps in the detection of similar patterns (Charbuty and Abdulazeez, 2021). The pattern analysis helps in the predictions that are based on the implemented algorithms on the used dataset of credit card fraud. The dataset has been implemented to study so that predictions can be made on whether a credit card transaction is fraudulent or not. The other parameters that are related to the transactions have also been used in the study so that all the parameters can be used to make decisions. Pattern detection helps in understanding and making future predictions related to fraudulent transactions using credit cards.

# Chapter 5: Evaluation and Conclusion

## 5.1 Critical Evaluation

A credit card fraud detection system has been developed using the python notebook with the help of “Deep learning algorithms”. These algorithms help to generate an efficient Neural Network model using python language that can help to predict the outputs easily. The libraries that have been imported into the python notebook contain all the required packages to develop the fraud detection system. The libraries are required to build an effective neural network model for4 the prediction of fraud. These libraries are provided by python and have been used in the research topic to increase the performance of the program. Pandas and Numpy have been imported into the notebook to provide the tools required to operate the data sets. The toll provided by these libraries helps to perform multiple numeric and statistical operations on the data set. The Numpy library helps to calculate mathematical and statistical equations present in the data set.

The data set has been imported into the notebook using the Pandas library to perform all the required operations on it. The null values present inside the data set have been calculated using the IsNull () function. This function calculates all the null values present inside the data set according to their respective columns. The null values have been removed from the data set to prevent any errors in accuracy. The summary of all the data present inside the data set has been calculated to understand the nature of values present inside the data set. The data set has been divided into the train and test data set to develop the “deep learning algorithm”. The x\_train and y\_train data sets have been formed to develop the classifier model and predict the accuracy. The “decision tree classifier” model has been used in the program to train the model. The model has been fitted into the classifier algorithm with the help of the sklearn package used in the notebook.

The accuracy and confusion matrix of the model has been calculated to understand the performance and efficiency of the model developed. The accuracy of the classifier model is 99% and a heat plot has been developed for the confusion matrix. The sklearn library package has been implemented in the program to build the classifier models. The tree has been plotted in the notebook with the help of the plt () function imported from the matplotlib library. Standard scalar and Labelencoder packages have been implemented using the sklearn library to transform the model for the implementation of the neural network. “Artificial Neural Network” or ANN has been chosen as the model for the neural network in this research. Adam optimizer has been preferred to be used in the model for generating the epochs. The neural network model has been optimized for better performance and efficiency. The accuracy score of the model has resulted at 96% along with the confusion matrix.

## 5.2 Summary of achievements

All the achievements that have been achieved by the researchers after developing the algorithm have been described in this section. The algorithm developed in the python notebook has achieved all the objectives that are required from the program. The program has achieved to perfectly remove all null and redundant values present in the data set and remove all unnecessary columns. The algorithm has effectively imported all the required libraries required to generate an efficient “deep learning algorithm” in python. The model has successfully achieved to split the data set into train and test data sets for developing the machine learning technique. The program has achieved to develop a perfect “deep learning algorithm” that can be used to detect credit card frauds. Implementation of the “decision tree classifier” into the model has been one of the greatest achievements for researchers. The search that has been performed to successfully implement the decision tree classifier model has paid off. Fitting the data sets into the model has been successfully implemented in the program using all the necessary libraries and packages of python.

The researchers have successfully achieved to develop a model of high accuracy for detecting frauds in credit cards. The performance of the model has been the most significant achievement for the researchers despite all the difficulties faced in the development of the model. The ANN model has been developed using all the research materials gathered during the data collection stages. The model has been successfully implemented in the program and to implement the concepts of the neural networks into fraud detection. The accuracy of the ANN model has been an achievement for the researchers after struggling to fit the data into the model.

## 5.3 Reflection

I have performed all the background research regarding the development of the classifier model and neural network for the program. The research and data collected have helped guide me through the development stages. I have analyzed different models that can be used to perform the credit card fraud detection system using machine learning. All the models that have been analyzed have shown some complexities in the scenario of the research. I have to perform various research regarding the best model and have chosen the “decision tree classifier” and ANN for the program. I have found all the libraries that are required to build the “deep learning algorithm” with maximum efficiency. All the libraries are helpful in the development of the algorithm. I have analyzed all the different functions that can be used in the program to successfully implement all the algorithms of machine learning. The “deep learning algorithm” has been developed after a lot of research regarding the proper functions and models best for the research.

I have performed all the research regarding the “decision tree classifier” model and ANN to increase the performance and efficiency of the model. I have analyzed all the different aspects of the model that can be used to improve the performance of those models. The accuracy of the model has been increased after implementing all the necessary modifications to the models. I have studied all the implementations of the “deep learning algorithms” in terms of fraud detection to successfully build the algorithm. I have found all the necessary optimizers and encoders that can be useful for increasing the efficiency and accuracy of the model. The necessary scalars have been used to scale all the features in the data set. I have analyzed all the uses of the Standardscalar package present inside the sklearn library module.

## 5.4 Research recommendation

Proper recommendations have been provided in this section that can be used to further enhance the implementation of the model. The model can be used to detect all sorts of credit card frauds that have occurred in the past years. The model has been properly trained to predict all types of credit card frauds that can occur using the data provided. The “machine learning algorithm” can be used to train multiple machines to predict the outcomes of the data imported into the model the model has an efficient classifier that can classify the data with maximum accuracy and efficiency. The algorithm has been made less complex so that it can be sued in several other places. The data used in the research can be implemented in other models for detecting all the patterns of fraud. Proper visualizations have been performed in the algorithm that can help understand the relationships of all the variables present inside the data set.

The visualizations can help detect patterns in the data set and generate effective graphs and plots. The efficiency of the model is the greatest achievement and can be used to detect fraud in other systems with maximum efficiency. The libraries used in this program can help to develop different types of classification models for machine learning purposes. The program can effectively generate all the predictions that are required for detecting fraud in transactions.

However, the model can be improved by adopting multiple classifier models and increasing the epochs for the neural network algorithm. These can help to create an efficient class classification for the data present in the data set. The model can be changed in case of unlabeled data present in the data set. The classifier model used in the algorithm cannot be used for unlabeled data sets and needs to be changed. The regression model must be implemented in the program to deal with unlabeled data presented in the data set. The model of the neural network can be changed according to the data present inside the data set. This can help to generate efficient predictions of the data required in the model. The epochs in the model can be adjusted to achieve more accuracy and efficiency in the predictions. The model can accommodate more visualizations of the data present in the data set to get a better idea regarding the patterns of the data in the data set.

## 5.5 Achievement of objectives

The objectives that have been mentioned at the beginning of the dissertation have all been successfully achieved. The program has been successfully developed in the python notebook while achieving all the objectives and aims of the research topic.

The model can accurately detect and analyze the owners of the credit cards that are present in the data set. The model prediction system can successfully predict the owners of credit cards with the help of the data and pieces of information provided in the data set. The classification model developed in the program classifies all the credit cards concerning their owners at an accuracy of 99%.

The machine learning algorithm can detect any anomalies or unnatural patterns in the transaction data of those people. These anomalies are properly analyzed by the model to detect any fraud occurring in those transactions. The model can then perform all the necessary methods to predict the fraud in that transaction.

The model can be used to prevent fraud in credit cards by predicting the anomalies present in the transaction details. The model can use the classification algorithm to classify the transaction details and prevent any forms of fraud during credit card transactions.

The algorithms used for supervised learning can help to increase the accuracy of the model generated in the python notebook. The algorithms of supervised learning help to classify the data efficiently and generate high accuracy with better confusion matrices.

The decision tree algorithms and ANN have been used in the program to predict the outcomes more accurately and efficiently. Proper modifications have been made to those algorithms to increase their accuracy and reduce code complexity. The neural network provided by ANN has been used to generate the nodes of the system that can effectively perform all the operations of the program.

The decision tree algorithm has been applied to the model with an accuracy score of 99%. The application of the neural network has been successfully integrated into the program for detecting credit card fraud. The model can prevent users from stepping into fraudulent transactions using their credit cards.

## 5.6 Further work

The program developed in the research can be further enhanced with the methods mentioned in this section. The neural network can be enhanced to detect fraud in all forms of transactions and prevent the user from performing anonymous transactions. The algorithm can be made to detect suspicious websites that can steal data of the user for credit card frauds. Multiple classification models can be incorporated into the program to classify the data more precisely and generate better results in the prediction process. The choice of the classification model can be changed according to the necessity of the data prediction. The data set can be visualized with the help of more graphs and plots to identify all patterns in the data. The model can be used in multiple systems to provide prevention and detection of credit card frauds. The model can be enhanced to reduce the time required to perform the “deep learning algorithm” and generate faster outputs.

# References

Arya, M. and Sastry G, H., 2020. DEAL–‘Deep Ensemble ALgorithm’framework for credit card fraud detection in real-time data stream with Google TensorFlow. *Smart Science*, *8*(2), pp.71-83.

Awoyemi, J.O., Adetunmbi, A.O. and Oluwadare, S.A., 2017, October. Credit card fraud detection using machine learning techniques: A comparative analysis. In *2017 international conference on computing networking and informatics (ICCNI)* (pp. 1-9). IEEE.

Charbuty, B. and Abdulazeez, A., 2021. Classification based on decision tree algorithm for machine learning. Journal of Applied Science and Technology Trends, 2(01), pp.20-28.

Dornadula, V.N. and Geetha, S., 2019. Credit card fraud detection using machine learning algorithms. Procedia computer science, 165, pp.631-641.

Dornadula, V.N. and Geetha, S., 2019. Credit card fraud detection using machine learning algorithms. *Procedia computer science*, *165*, pp.631-641.

Gao, Y., Jausseme, C., Huang, Z. and Yang, T., 2022. Hydrogen-Powered Aircraft: Hydrogen–electric hybrid propulsion for aviation. IEEE Electrification Magazine, 10(2), pp.17-26.

Hoelzen, J., Silberhorn, D., Zill, T., Bensmann, B. and Hanke-Rauschenbach, R., 2021. Hydrogen-powered aviation and its reliance on green hydrogen infrastructure–review and research gaps. International Journal of Hydrogen Energy.

Husejinovic, A., 2020. Credit card fraud detection using naive Bayesian and c4. 5 decision tree classifiers. *Husejinovic, A.(2020). Credit card fraud detection using naive Bayesian and C*, *4*, pp.1-5.

Inacio, G., Mourao, C., Castro, A.L. and Lacava, P., 2022. Feasibility Study of Using Liquid Hydrogen Tanks as Energy Carriers and Cooling Agents for a Small Aircraft Powered by PEMFCs (No. 2022-01-0009). SAE Technical Paper.

John, H. and Naaz, S., 2019. Credit card fraud detection using local outlier factor and isolation forest. *Int. J. Comput. Sci. Eng*, *7*(4), pp.1060-1064.

Kiely, M., Agarwal, R.K. and Rice, C., 2022. Automated Vortex Lattice Method Based Design Optimization of a Hydrogen Powered Aircraft. In AIAA AVIATION 2022 Forum (p. 3515).

Lenka, S.R., Barik, R.K., Patra, S.S. and Singh, V.P., 2021. Modified decision tree learning for cost-sensitive credit card fraud detection model. In *Advances in Communication and Computational Technology* (pp. 1479-1493). Springer, Singapore.

Najadat, H., Altiti, O., Aqouleh, A.A. and Younes, M., 2020, April. Credit card fraud detection based on machine and deep learning. In 2020 11th International Conference on Information and Communication Systems (ICICS) (pp. 204-208). IEEE.

Otchere, D.A., Ganat, T.O.A., Gholami, R. and Ridha, S., 2021. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. Journal of Petroleum Science and Engineering, 200, p.108182.

Patil, S. and Kulkarni, U., 2019, April. Accuracy prediction for distributed decision tree using machine learning approach. In 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 1365-1371). IEEE.

Safa, M.U. and Ganga, R.M., 2019. Credit Card Fraud Detection Using Machine Learning. *International Journal of Research in Engineering, Science and Management*, *2*(11), pp.372-374.

Silberhorn, D., Dahlmann, K., Görtz, A., Linke, F., Zanger, J., Rauch, B., Methling, T., Janzer, C. and Hartmann, J., 2022. Climate Impact Reduction Potentials of Synthetic Kerosene and Green Hydrogen Powered Mid-Range Aircraft Concepts. Applied Sciences, 12(12), p.5950.

Trivedi, N.K., Simaiya, S., Lilhore, U.K. and Sharma, S.K., 2020. An efficient credit card fraud detection model based on machine learning methods. *International Journal of Advanced Science and Technology*, *29*(5), pp.3414-3424.

Troeltsch, F.M., Engelmann, M., Scholz, A.E., Peter, F., Kaiser, J. and Hornung, M., 2020. Hydrogen powered long haul aircraft with minimized climate impact. In AIAA Aviation 2020 Forum (p. 2660).

Varmedja, D., Karanovic, M., Sladojevic, S., Arsenovic, M. and Anderla, A., 2019, March. Credit card fraud detection-machine learning methods. In 2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH) (pp. 1-5). IEEE.

Varmedja, D., Karanovic, M., Sladojevic, S., Arsenovic, M. and Anderla, A., 2019, March. Credit card fraud detection-machine learning methods. In *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH)* (pp. 1-5). IEEE.

Vijayakumar, V., Divya, N.S., Sarojini, P. and Sonika, K., 2020. Isolation forest and local outlier factor for credit card fraud detection system. *International Journal of Engineering and Advanced Technology (IJEAT)*, *9*, pp.261-265.