ASHESI UNIVERSITY

Adaptive Credit Card Fraud Prediction Using Artificial Neural Network

CAPSTONE PROJECT

B.Sc. Computer Engineering

Juliet Fatima Abdul-Aziz

2020

ASHESI UNIVERSITY

Adaptive Credit Card Fraud Prediction Using Artificial Neural Network

CAPSTONE PROJECT

Capstone Project submitted to the Department of Engineering, Ashesi University in partial fulfilment of the requirements for the award of Bachelor of Science degree in Computer Engineering.

Juliet Fatima Abdul-Aziz

2020

DECLARATION

I hereby declare that this capstone is the result of my own original work and that no part of it has been presented for another degree in this university or elsewhere.

Candidate's Signature:

……Juliet Fatima Abdul-Aziz…….

Candidate's Name:

……Juliet Fatima Abdul-Aziza……

Date:

……29/ 05/ 2020…….

I hereby declare that preparation and presentation of this capstone were supervised in accordance with the guidelines on supervision of capstone laid down by Ashesi University College.

Supervisor's Signature:

……

Supervisor's Name:

……

Date:

……

Acknowledgements

I truly appreciate God's grace in seeing me through my studies at Ashesi and for giving me the strength to complete this project. To my supervisor, Abdul Wasay whose encouragement and academic advice helped me undertake this project, I sincerely give him my gratitude for taking an interest in my ideas. Thanks to his patience and understanding, I was able to complete this project. I also want to specially thank Isaac Coffie for listening to me and for advising me whenever I got stuck in the process. Lastly, to all my friends and colleagues who were there for me when the journey become difficult and for assisting me in your own way, I appreciate you all.

Abstract

Currently, there is a growth in online transactions which has led to the immerse growth of the number of credit card fraud. A lot more people are opting to shop online due to convenience and therefore they make online payments to make a purchase that would be delivered to them and in some cases, they make payments online for a service rendered to them. With such an opportunity, fraudsters are also increasing their fraud activities online. Therefore, this study seeks to detect credit card fraud using an adaptive tool and also attempts to reduce the number of wrongly predicted valid transactions made by the model. Researchers have used tools such as K-nearest neighbour, logistic regression, random forest, decision trees and others however, this study uses an autoencoder neural network to detect credit card fraud. The study then evaluates the model using an appropriate evaluation metric.

Keywords: Fraud detection, adaptable, autoencoder neural network, credit card, online transactions

Table of Contents

Chapter 1: Introduction

1.1 Background of the study

With the convenience that comes with buying and selling items online, e-commerce businesses are growing exponentially. Items such as cloths, furniture, and even food are now bought online instead of the traditional way of buying items in a shop. Hence the number of online transactions has grown immensely leading to an increase in credit card transactions. However, as the internet paved way for such progress, people with bad intentions have also taken advantage of the opportunity to rob others of their money causing a huge security concern among e-commerce users hence affecting the economy drastically.

The Nilson Report in November 2019 estimated the revenue obtained from online transactions, the growth of online payment systems and also estimated the losses caused by credit card fraud. The credit card market, according to the report, generated $40.582 trillion in 2018 which is an increase of 17.7% from 2017. The report also showed that global credit card brands and users faced gross fraud losses of about $24.86 billion in 2018 which is an increase of 16.9% from 2017's gross fraud loss of $21.27 billion [15]. Due to this, credit card fraud places a huge cost on financial institutions and card issuers leading these institutions to place high demand on sophisticated fraud detection application in order to flag any suspicious transaction [10]. Credit card fraud keeps making the headline hence suggesting that current fraud detection applications have loopholes that is why e-commerce and credit card users still face so much inconvenience and makes the field highly researched [14].

Fraudulent transactions are illegitimate credit card transactions done without the awareness of the actual cardholder [14]. There are two categories of credit card payment namely; card-not-present (CNP) and card-present (CP) [16]. CNP is when details of a card

are used to make purchases online, phone or by mail while CP is when purchases are done using a physical credit card [17]. CNP fraud according to [18], also occurs when the transaction is made remotely. Each of these categories have specific types of fraud that occur within them. The nature of CNP payment makes it highly susceptible to fraud. This study is hence going to focus on CNP payment fraud. The types of fraud that can occur using CNP payment are identity theft fraud, behavioural fraud, application fraud and others [16]. CNP payment has been proven to have the fastest growth rate on an average of 15-20% per year as compared CP transactions at POS devices and ATM devices that has an average growth rate of about 4% per year [17]. This is caused by the rate at which e-commerce is growing rapidly hence making CNP have more fraud cases if measures are not taken [17].

Fraud can be avoided either through detection or prevention where prevention has to do with acting as a layer of protection to avoid any attacks and detection helps in identifying and signalling a fraudulent transaction as soon as it is triggered [14]. Detection is usually added to prevention so that once prevention fails, fraud is still detected as soon as possible [14]. Criminals are becoming more intelligent as technology progresses. They learn to adapt their fraud strategies whenever a new detection method is in place hence making them progress along with new detection systems [18]. The knowledge of the adaptability of criminals makes the fraud detection discipline a difficult field to solve. This is because researchers cannot publicly exchange ideas in detail about fraud detection techniques as criminals will evade the detection easily [18]. In light of this, as fraudsters adapt to new detection techniques, researchers need to be adaptable to new fraud strategies as well.

1.2 Problem Statement

Based on this background, the motivation of this project is to design an adaptive algorithm that detects new variations of credit card fraud transactions and predicts a current transaction as fraudulent or legitimate whiles reducing the number of wrongly predicted

legitimate transaction. There are many ways fraudulent transactions can occur and the specific type of online fraud this study would focus on is identity theft. Identity theft happens when a criminal steals details of a cardholder, impersonates the cardholder to make purchases and attributes the charges to the cardholder [9].

The people affected by this malicious action is the online seller and the customer because the seller lose their merchandise, pays a chargeback to the real cardholder and the cardholder deals with the aftermath. This problem would be addressed using a selected machine learning model based on studies done by other researchers and tested by feeding it with an already known class of a transaction for prediction. The implemented solution to the mentioned problem can be made as a plugin on e-commerce websites to detect fraudulent transactions as they happen.

During an incoming transaction, the system classifies the transaction as either valid or fraudulent with a rating which allows the user to either decline or permit the transaction. Hence solving the problem of huge financial losses due to credit card fraud. The benefit of this detection system is that it would provide confidence back to e-commerce users and would enable them to make any transaction especially huge ones comfortably without any concerns. This boosts online trading and allows the sector to grow.

Enabling people to gain confidence in online trading would also reduce cash services and make more people gear towards cashless services enabling a cashless economy in the long run. There are many benefits that comes with a cashless economy aside from it easing a person's life. According to research done in Nigeria [20], a cashless economy makes transactions faster, increases sales, reduces cash related robbery and corruption, and attracts foreign investors. Government expenditure especially would also be significantly reduced as there would be no cost incurred in printing currency notes.

1.3 Limitations of the Study

Despite the benefits that comes with creating an efficient credit card fraud detection system, there are a number of challenges that comes with building such systems. According to [10], there are many constraints that comes with creating the system. There is the issue of data unavailability since financial institutions prioritise their customer's data and hence protect it from being revealed to second parties [11]. This makes it barely possible to find real life data of credit card transactions.

Even if data is obtained, there will be another challenge of working with an unbalanced dataset. Fraudulent transactions are few as compared to legitimate transactions out of almost every dataset. Typically, there's about 99.9% legitimate transactions while only 0.1% or less are fraudulent [12]. Another major issue with the detection system is that there will always be a probability of misclassifying a transaction hence either the system does not identify a fraudulent transaction and misclassifies it as valid or the system misclassifies a valid transaction as fraudulent which is a huge prevalent problem with this system [12].

This brings the issue of finding an appropriate evaluation metric. Due to the imbalanced nature of credit card datasets, accuracy is not a good measure of performance for the model [13]. Most studies utilise the false positive and false negative rates as a performance metric and these two rates have opposite relationship with each other [10]. In this case, the false positives are the misclassified legitimate transactions and false negatives are the misclassified fraudulent transactions. According to [10], the error cost of misclassifying a legitimate transaction is less than the error cost of misclassifying a fraudulent transaction. Due to the opposing nature of the two rates, it can be hard to draw the line between them, that is, what percentage of each of the two is required to produce the optimum model. Another challenge in building the fraud detection system is that fraudsters

are constantly using intelligent adaptive techniques and changing their behaviours against any new detection system [10]. This is making fraud more unpredictable and making credit card fraud detection a hard problem to solve.

Chapter 2: Literature Review

2.1 Overview of Credit Card Fraud

According to [1], fraud is an illegal use of a system. Therefore, credit card fraud is the illegal use of credit cards. The extent to which fraud occurs cannot be quantified because companies do not make this information public else, they would frighten their customers [31]. There are various types of fraud identified by different researchers. According to [1], There are two types of credit card fraud, namely, illegal use of stolen or lost card and counterfeit fraud. Also, according to [4], there are four types of fraud namely; counterfeit fraud, behavioural fraud, application fraud and bankruptcy fraud. Several researchers have tried using various data mining techniques to solve the problem. However, there are two problems that usually cut across in the outcome of fraud detection systems; the lack of adaptability of the system and the fraud detection cost because the cost of the detection and the cost of the fraud itself must be taken into account [29].

There are also problems researchers constantly encounter when solving the problem as have been mentioned before, but researchers have found ways to deal with the problems they encounter during the detection process. One major problem is the unavailability of data due to data sensitivity hence there is barely any real-life data available for implementing the solution [2] [15]. This is because financial institutions cannot release the data of their customers to researchers due to data privacy and sensitivity. Hence most researchers use data found on Kaggle webpage which is a data science platform that has numerous data related problems and resources.

There is also a problem associated with the nature of the data since is it imbalance, that is, the data does not have a good balance between valid transactions and fraud transactions hence making the data skewed in nature [3]. To solve the problem of

imbalanced data, several researchers perform under sampling and oversampling on the dataset using synthetic minority oversampling technique (SMOTE). Mostly, there are relatively very few fraud transactions within a large dataset of valid transactions as mentioned earlier. There is also the problem of dealing with categorical data since most of the data generated from the transactions are categorical and most of the machine learning models do not support categorical data [29]. Hence researchers struggle to find the right model for a given dataset. Another similar challenge is selecting the right features especially because training takes more time than predicting.

Some of the fraud detection algorithms that have been implemented used the following machine learning models: Bayesian networks, Hidden Markov Model, neural networks, decision trees and data mining techniques [5]. However, [6] suggests that more emphasis have been on neural network and data mining techniques. Other models that have also been employed are outlier techniques [7], K-Nearest Neighbour and self-organizing maps [8].

2.1.1 Types of Credit Card Fraud

This section aims to expound on the most prominent types of credit card fraud using CNP and CP payment in order to provide a sense of how criminals perpetrate this crime of robbing others of their asset.

(a) Use of Stolen Cards

Using stolen cards is probably the easiest type of fraud. This happens when credit cardholders lose their cards or when their cards are stolen by criminals [18]. When this happens, criminals spend every single amount on the card as soon as possible to prevent the fraud transaction from being detected and to prevent themselves from being caught [18]. In this case, detecting the theft early is very crucial [18].

(b) Bankruptcy Fraud

This is a type of fraud where people use their credit cards but are unable to pay the depts they already owe or are insolvent [39]. Knowing very well they cannot pay for items; some people go ahead to make purchases with their credit cards. After banks realise these people are unable to pay, the banks have to cover for the losses unfortunately [39]. This is type of fraud according to [39] is one of the most challenging type to predict and the only way to stop this kind of fraud from happening is by performing a pre-check to be obtain information about the banking history of customers. Other research presented in [39] used regression techniques to detect bankruptcy fraud.

(c) Application Fraud

Application fraud is a type of fraud where criminals obtain new credit cards from companies that issues card [18]. What makes this a fraud is that they obtain these new cards with false personal information which was obtained from real personal and financial information [34]. Fraudsters also make the credit card companies mail the new cards to them through a specified mail drop [34]. Credit scoreboards are used to detect such default applications using statistical models that can observe behaviours over time and hence detects a card obtained through a false application [18]. Application fraud can also be detected when there are duplicate applications coming from the same person [39]. The duplicates are identified using techniques such as cross matching.

(d) Counterfeit or Theft Fraud

This is a type of fraud where people use credit cards that do not belong to them as many times as possible [39]. The perpetuators massively spend from the card before the original cardholder is able to block the card. This type of fraud can also be done

with just the card details and used remotely [39] hence resembling an identity theft. The merchants are at a risk because they have to always pay a charge back fee when the original cardholder files a complain.

2.1.2 How Criminals Find Credit Card Details

With CNP fraud, criminals only need the details of a card to make purchases that resembles the original cardholder and hence use various means of attaining such details. The main details acquired by fraudsters are merchant code, type of credit card, size of transaction, account number, type of purchase, date of transaction and client's name [18]. One of the ways people steal card information is through what is known as skimming, mainly perpetuated by employees where they illegally get a copy of the magnetic strip on credit cards [18]. They do this by swiping it though a card reader. There are others who also pose as legitimate employees of credit card companies to take details of credit card transactions over the phone [18]. Another way of obtaining credit card details is through the dark web.

2.2 Related Works

To solve the problem of credit card frauds, several data mining solutions in the field of supervised learning, unsupervised learning and hybrid approaches have been proposed to solve the problem. This notable method of using data mining is the process of using artificial intelligence, machine learning, mathematical and statistical techniques to gain insights, identify patterns and information from a large database [25]. Data mining plays a huge role in financial fraud detection application [26] and the techniques it presents have been proven to be successful in solving the problem [19]. In [24], the most prevalent techniques used in solving the problem are neural networks, Bayesian networks, logistic models and decision trees which are all applicable to classification problems. Studies in [24] identified various data mining categories applicable to the fraud detection problem namely; classification,

prediction, outlier detection, clustering, visualization and regression along with their algorithmic approaches as shown in the figure below.

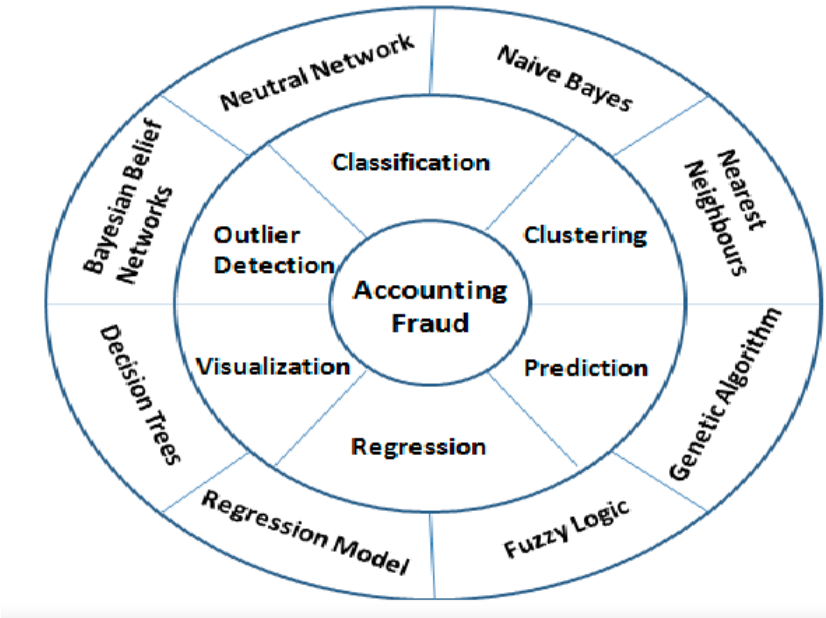

Figure 2.1: Conceptual data mining framework for fraud detection [24]

From the figure above, it can be seen that each of the algorithmic approaches can be used to implements multiple classes of problems. The review below are research done within several popular literatures that focus on credit card detection techniques using some of the data mining techniques mentioned.

2.2.1 Detection Using Neural Networks

Neural networks are widely used in the credit card fraud prediction [24]. It has been used by researchers such as [34][35]. They are modelling tools that imitates the functionality of the human brain using a number of interconnected nodes and it is widely used in solving classification problems [26]. According to [24] neural networks have three main advantages; it is adaptive, it is highly robust and can be easily modified.

2.2.2 Detection Using Support Vector Machines

Support Vector Machine (SVM) is a machine learning technique that is mostly suitable for binary classification problems and hence makes it a suitable technique to solve the credit card fraud detection problem since the problem has two possible outcomes [22]. The SVM has two strong properties namely; the kernel representation and the margin optimization. The SVM creates a hyperplane that separates the instances of the dataset into negative and positive [23]. It finds a maximum margin hyperplane which gives the greatest separation between classes and separates all instances of the dataset correctly [10]. When SVM is used in credit card fraud detection, if instances are found around the maximum margin hyperplane, the instances are normal else they are anomalous [10]. SVM does not work well with imbalanced dataset because the minority instance may rather be found around the maximum margin hyperplane [22] hence might be problematic in solving credit card fraud.

2.2.3 Detection Using Naive Bayes Classifiers

[10] Describes Naïve Bayes as a supervised learning algorithm that uses training dataset to predict future outcomes. [10]'s study mentioned that it has a good performance even with the least number of training dataset and can be used to solve both binary and multiclass classification problems. Studies have shown that Naive Bayes Performs well in credit card fraud detection. It classifies a given transaction by applying the Bayes rule which calculates the probability if the right class a transaction belongs to. It combines what is called the prior and likelihood to form the posterior probability which is the same as the Bayes rule given below.

$$posterior = \frac{Prior \times Likelihood}{Evidence}$$

Where prior and likelihood is given as

$$Prior\ Probability\ of\ Z = \frac{Number\ of\ Z\ instances}{Total\ number\ of\ instances}$$

$$likelihood\ of\ Y\ given\ Z = \frac{Number\ of\ Z\ in\ vicity\ of\ Y}{Total\ number\ of\ Z}$$

2.4.4 Detection Using K-Nearest Neighbour Algorithm

K-Nearest Neighbour (KNN) classifies based on similarities and is mostly used in pattern recognition [10]. When using KNN for credit card fraud detection, a similarity between two data set is used. An incoming transaction is classified based in the distance between the incoming transaction to the nearest point [22]. Depending on the class of the nearest point, the incoming transaction is assumed to be in that class. To calculate the distance between the two data set, Euclidean distance can be used if the data is continuous, and matching coefficients can be used if the data is categorical [10]. This can be problematic during the training process if the two data sets are unrelated hence making this technique quite inefficient [22].

2.2.5 Detection Using Logistic Regression

Logistic regression according to [21] is not exactly a regression algorithm. In this model, the prediction is done based on probability of the outcome instead of classes. The output is predicted through the combination of the input variable (x) with assigned weights. The general equation can be illustrated as $y = a0 + a1 \times x$. Where y is the output dependent variable, a0 is the biased term and x is the input independent variable multiplied with a weight a1. Since it estimates a likelihood which must fall in a range, it uses a sigmoid function, $z = \frac{1}{1+e^{-x}}$, to fit the output into the range of 0 to 1. A sigmoid is an activation function that limits outputs in the range of 0 to 1.

Chapter 3: Approach and Methodology

This chapter of the report highlights the user requirement and the system requirement as well as the methodology that was followed to develop the proposed solution.

3.1 User Requirement

(a) Users should be provided with safety and confidence when making purchases and payments online.

(b) Users should not lose their money or asset as a result of fraudulent activities and fraud detection systems.

(c) Users should be alerted or should know whenever there is an abnormal transaction.

(d) Users should be able to decline or accept a transaction based on a fraud likelihood score.

(e) Users should be provided with a faster and a more convenient way of purchasing items.

(f) More users should rather opt for online transactions as a channel of payment when it comes to buying and selling instead of the traditional way of purchasing.

(g) Users should interact with a GUI to able to accept or decline a transaction.

3.2 System Requirement

(a) The system should be built with thousands of data specifically, over 20,000.

(b) The system should have a fraud detection model.

(c) The system should have the most optimal fraud detection model.

(d) The system should be adaptable to new variations of fraud strategies.

(e) The system should block all transactions whose cardholder’s details are different from demographic details.

(f) The system should have accuracy above ninety percent and should be highly effective with over ninety percent fraud detection rate.

(g) The system should be scalable, that is, it should be able to be improved to a better version and also be able to be debugged.

(h) The system should allow the termination of a transactions as soon as fraud is detected.

(i) The system should provide a fraud rating in other to allow users to either accept or decline a transaction.

3.3 Methodology

Based on requirements elicitation gathered above, together with insights from the literature review, the system would be built using an artificial neural network since such a network would be useful in detecting new variations of fraud hence catering for adaptability. This is because the model can learn from both the past and current situations, making fraud detection done in a faster and more efficient way. Also, the review of the literature shows that neural networks can be easily modified making it useful for researchers to be able to adapt the system to handle new fraud strategies. This would cater for the hypothesis that an adaptable credit card detection system would help credit card users detect fraud and new variations of fraud as it occurs.

Moreover, there are variations of artificial neural networks. However, the proposed system would focus on the use of an artificial neural network called an autoencoder. The figure below gives an overview of the processes involved in implementing the solution.

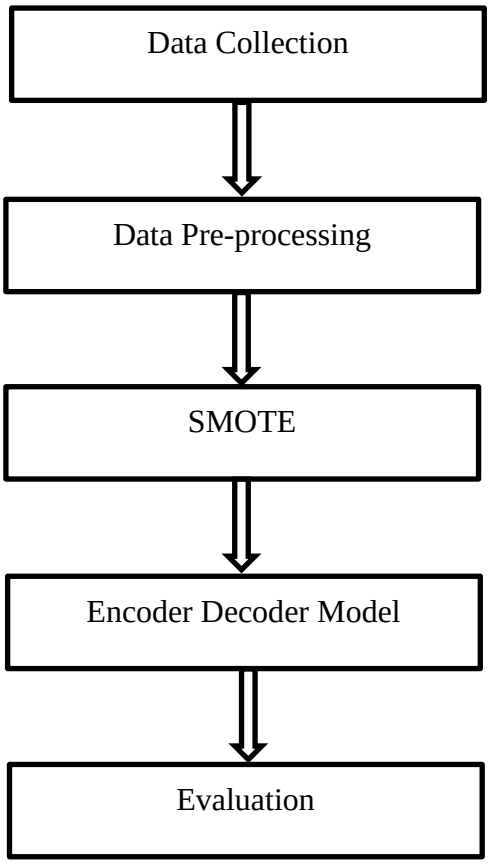

Figure 3.1: Credit card fraud detection process

3.3.1 Data Collection

In this study, the dataset used is a two days European transaction of credit card users downloaded from Kaggle, an online community that shares and allows the publishing of dataset. The dataset has 31 features of which 28 were transformed using Principal Component Analysis (PCA) due to data privacy and protection. Hence only three features out of 31 features are known and 23 features are unknown represented as V1, V2 to V28. The three features that were not transformed are time, amount and class which has the value of 1 representing fraudulent transaction and 0 representing legitimate transaction. The time is the elapsed seconds that occurred between each transaction with respect to the first transaction in the dataset. The dataset has a total of 284,807 transaction where 492 are

fraudulent and the rest are legitimate. Readers should know that a lot of researchers used this particular data to train their model hence some outputs during the process might be consistent with that of other research works including this one. This is as a result of data unavailability as extensively mentioned.

3.3.2 Data Pre-processing

The dataset was split into percentages where 60% was dedicated to training set, 20% was dedicated to test set and the other 20% was dedicated to validation set. The next step involved data cleaning by checking for missing values and also performing dimension reduction to remove unimportant aspects so as to prevent the model from learning wrong patterns. The nature of the data was also inspected to find the number of columns, the available features and the other features transformed by PCA. The ratio of valid and fraud transaction classes was also found and visualised and the shapes or dimensions of each of the two classes of transaction were identified as well. The amount of money used in each of the transaction class was inspected and visualised. Based on the assumption that more fraudulent transactions occur at certain times, the times of transaction in each of the transaction classes was also visualised. Hence determining the distribution of the two available features; amount and time.

3.3.3 SMOTE

Another task performed is to balance the two classes since there are relatively very less fraudulent transactions as compared to valid transactions. The is done by using a sampling technique to oversample the fraudulent transactions and under sample the valid transactions using the Synthetic Minority Oversampling Technique (SMOTE). This is because the amount of valid transaction is high hence under sampling it would cause it decrease and the amount of fraudulent transaction is low hence oversampling it would cause

it to increase. This creates a balance between the two classes. The new dimensions of the two transactions classes is also determined.

3.3.3 Modelling and Testing

As mentioned, the model used is an autoencoder neural network (AE). Motivated by the function of the brain, the AE can recognize similar patterns and make future predictions based on the pattern it had already learned. The idea of the AE is that it learns a compressed representation of an input given an input and predicts the exact input. A typical AE is synonymous to a multilayer perceptron model because it is a feedforward neural network [33]. As shown in figure 3.3, the AE has two main parts; the encoder and decoder which has an input layer, a hidden layer and an output layer. The only difference between the multilayer perceptron and the AE is that the AE has the same number of input layer as the output layer [33]. The AE learns to make an approximation of the identity function given by;

$$fW, b(x) \approx x$$

The cost function of the AE neural network is given by figure 3.2 below.

$$J_{A,E} = \frac{1}{m} \sum_{i=1}^{m} \left(\frac{1}{2} \| \hat{x}_i - x_i \|^2 \right)$$

Figure 3.2: Cost function of autoencoder [33]

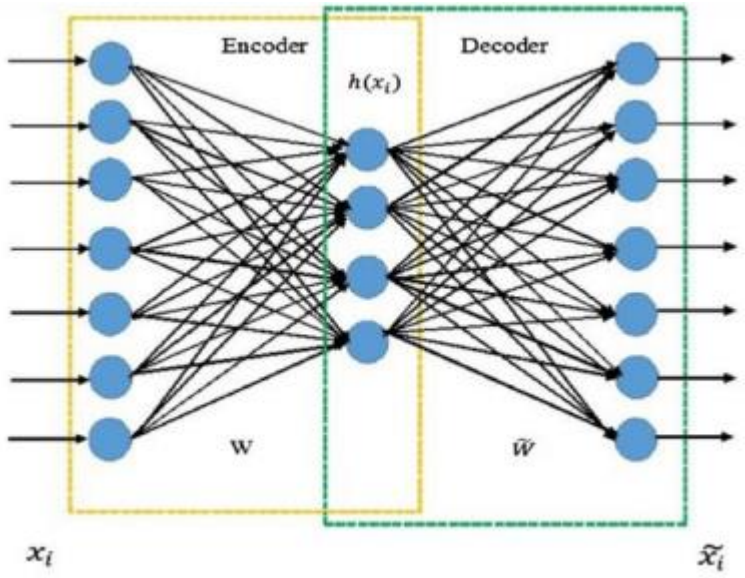

Figure 3.3: Autoencoder neural network architecture [33]

The AE used in this research has two hidden layers and uses the relu activation function. The hyperparameter, specifically the number of hidden layers, was chosen through an iterative means.

3.3.4 Model Evaluation

Finally, an appropriate performance metrics was used to see how well the model is performing. Commonly used metric such as accuracy cannot be used because it can be biased on an unbalanced dataset [10] [22]. In this case the confusion matrix was used to evaluate the model. The matrix is in the form of a tabular layout that typically creates a visualisation of the performance of an algorithm. The confusion matrix is made up of four parts namely; TP, TN, FP, FN as shown in figure 3.4.

Where:

TP = True positive; the number of fraudulent transactions predicted as fraud [22].

TN = True negative; the number of valid transactions predicted as valid [22].

FP = False positive; the number of valid transactions predicted as fraud [22].

FN = False negative; the number of fraudulent transactions predicted a fraud [22].

Classification	Actual Positive Sample	Actual Negative Sample
predict as positive	TP	FP
predict as negative	FN	TN

Figure 3.4: Confusion matrix

The rate of these four metrics is given by the false positive rate (false alarming rate), false negative rate, true positive rate (fraud catching rate) and true negative rate and their formulas are illustrated below.

$$True\ Positive\ Rate = \frac{TP}{TP + FN}$$

$$True\ Negative\ Rate = \frac{TN}{TN + FP}$$

$$False\ Positive\ Rate = \frac{FP}{FP + TN}$$

$$False\ Negative\ Rate = \frac{FN}{TP + FN}$$

Another rate that is mostly used for evaluating the model is the recall rate. The recall is the ability of the model to find all positive samples which are the percentage of legitimate transactions identified correctly by model. This is the same as finding the true positive rate since recall rate is given as;

$$Recall\ Rate = \frac{TP}{TP + FN}$$

Chapter 4: Methodology 2 – Implementation

This chapter discusses the implementation details of the methods highlighted in the previous chapter. The first implementation task in most machine learning project is to import necessary libraries. In this case, libraries such as NumPy, pandas, matplotlib, seaborn and other useful libraries were imported. The important APIs that were also utilised to make the implementation easier are Keras and Tensorflow which are open source libraries used in training deep learning models. The next thing was to load the data and inspect the data type.

Figure 4.1 below shows all the features in the dataset including the ones transformed by PCA. Figure 4.2 also describes the features in terms of the count, mean, minimum and maximum values as shown below.

```
Index(['Time', 'V1', 'V2', 'V3', 'V4', 'V5', 'V6', 'V7', 'V8', 'V9', 'V10',
       'V11', 'V12', 'V13', 'V14', 'V15', 'V16', 'V17', 'V18', 'V19', 'V20',
       'V21', 'V22', 'V23', 'V24', 'V25', 'V26', 'V27', 'V28', 'Amount',
       'Class'],
      dtype='object')
```

Figure 4.1: Columns of dataset

	Time	V1	V2	V3	V4	V5	V6	V7	V8	V9
count	284807.000000	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05	2.848070e+05
mean	94813.859575	3.919560e-15	5.688174e-16	-8.769071e-15	2.782312e-15	-1.552563e-15	2.010663e-15	-1.694249e-15	-1.927028e-16	-3.137024e-15
std	47488.145955	1.958696e+00	1.651309e+00	1.516255e+00	1.415869e+00	1.380247e+00	1.332271e+00	1.237094e+00	1.194353e+00	1.098632e+00
min	0.000000	-5.640751e+01	-7.271573e+01	-4.832559e+01	-5.683171e+00	-1.137433e+02	-2.616051e+01	-4.355724e+01	-7.321672e+01	-1.343407e+01
25%	54201.500000	-9.203734e-01	-5.985499e-01	-8.903648e-01	-8.486401e-01	-6.915971e-01	-7.682956e-01	-5.540759e-01	-2.086297e-01	-6.430976e-01
50%	84692.000000	1.810880e-02	6.548556e-02	1.798463e-01	-1.984653e-02	-5.433583e-02	-2.741871e-01	4.010308e-02	2.235804e-02	-5.142873e-02
75%	139320.500000	1.315642e+00	8.037239e-01	1.027196e+00	7.433413e-01	6.119264e-01	3.985649e-01	5.704361e-01	3.273459e-01	5.971390e-01
max	172792.000000	2.454930e+00	2.205773e+01	9.382558e+00	1.687534e+01	3.480167e+01	7.330163e+01	1.205895e+02	2.000721e+01	1.559499e+01

Figure 4.2 Description of features

4.1 Implementing Data pre-processing

After getting an overview of each of the features, they were then visualised individually as shown in figure 4.4. Even though most of the exact features are unknown, the visualisations help in exploring their distributions.

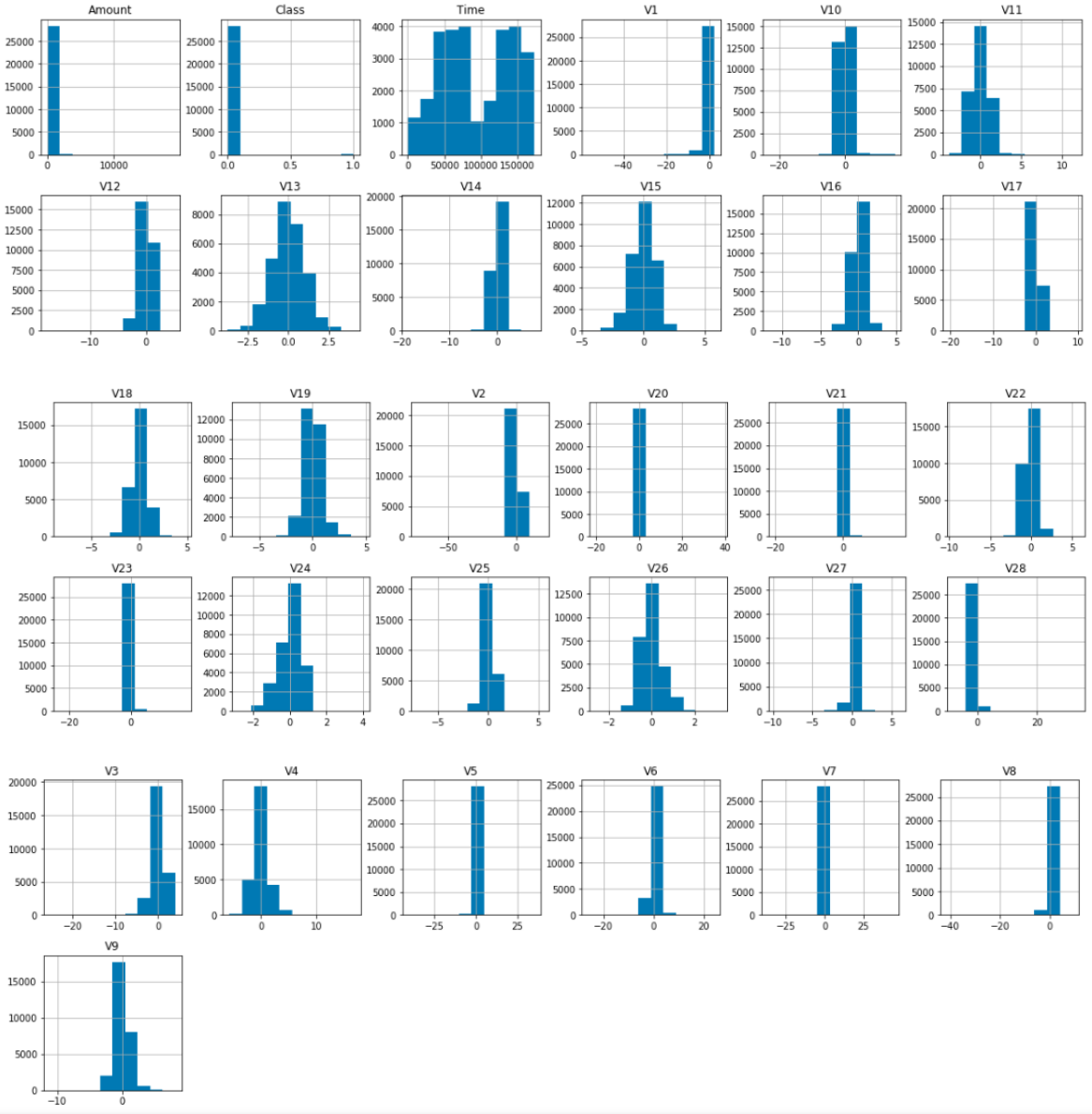

Figure 4.3: Feature visualisations

The data was checked for missing values and was found to be false hence made data exploration easier. As seen in the chart below, the data has almost zero fraud cases hence making the data skewed. Specifically, there were 284, 315 legitimate transaction with only

492 fraudulent transactions. This skewness was later transformed using SMOTE to bring the data to normal or make it balanced.

Figure 4.5: Fraud and valid transaction distribution

The details captured below shows that maximum amount stolen by fraudsters was about 2,125 euros and the maximum legitimate transaction in the dataset was about 25, 691 euros. The mean and standard deviation in the in the fraudulent transactions and the valid transactions is also shown below in figures 4.6 and 4.7.

```
Details of the fraudulent transaction

count     492.000000
mean      122.211321
std       256.683288
min         0.000000
25%         1.000000
50%         9.250000
75%       105.890000
max      2125.870000
Name: Amount, dtype: float64
```

Figure 4.6: Describes the fraudulent transaction

```
Details of the valid transaction

count    284315.000000
mean         88.291022
std         250.105092
min           0.000000
25%           5.650000
50%          22.000000
75%          77.050000
max       25691.160000
Name: Amount, dtype: float64
```

Figure 4.7: Describes the Valid Transactions

To perform feature engineering on the data set, the distribution of the two main features in the dataset was explored. Below is the distribution of amount in the fraud and valid transactions respectively. Also, an assumption made was that fraud happens at certain times of the day hence the distribution of time was visualised in both the valid and fraud transactions in figure 4.7. The time distribution clearly showed that time is not a factor in detecting credit card fraud. Since time of transactions does not seem to matter in the detection process, the time feature was dropped.

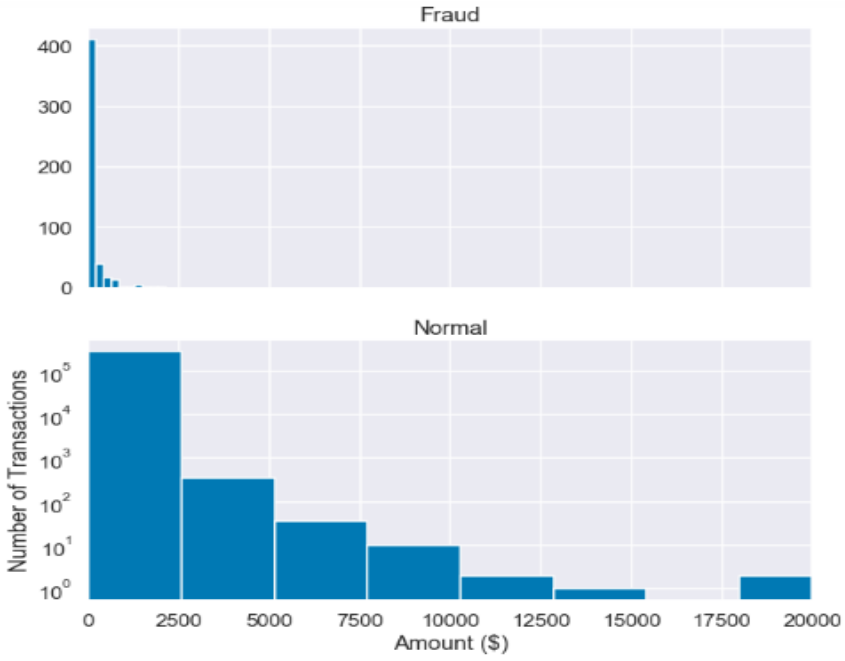

Figure 4.6: The distribution of amount in valid and fraud transactions

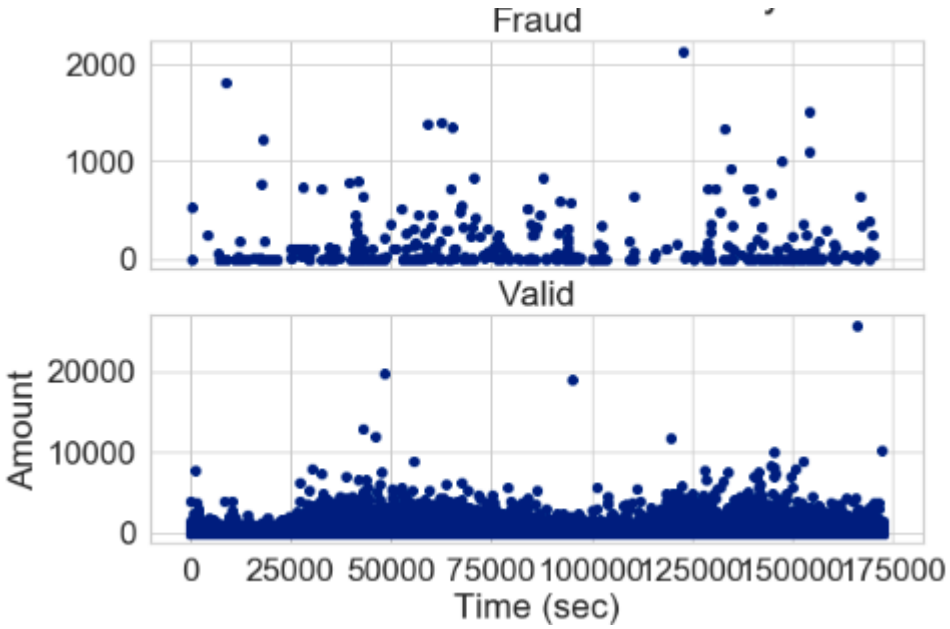

Figure 4.7: The distribution of time in fraud and valid transactions

4.2 Training the Model

After pre-processing, the model was trained using 4 interconnected layers with different number of neurons as used by [36]. The 4 layers was divided into two for each of the encoder and decoder model hence the encoder took the first two and the decoder took the last two. Each layer in the encoder and the decoder used the ReLU and tanh activation function and the first layer utilised L1 regularization. The model was run for 82 epochs with batch size of 32.

```
input_layer = Input(shape=(input_dim, ))
encoder = Dense(encoding_dim, activation="tanh",
                activity_regularizer=regularizers.l1(10e-5))(input_layer)
encoder = Dense(int(encoding_dim / 2), activation="relu")(encoder)
decoder = Dense(int(encoding_dim / 2), activation='tanh')(encoder)
decoder = Dense(input_dim, activation='relu')(decoder)
autoencoder = Model(inputs=input_layer, outputs=decoder)
```

Figure 4.8: Encoder and Decoder Model [36]

Chapter 5: Experiments and Results

From the implementations discussed in chapter 4, this chapter aims to expound on the results obtained from some experiments or testing done. This is to ensure the model is evaluated and to know whether or not it is performing according to the requirements of this study. To decide or to predict whether a transaction is fraudulent or not, a decision threshold must be defined. Thresholds are used in classification problems to categorize outcomes as either belonging to a certain group or not. In this case, a threshold determines if a transaction is above or below a normal range.

Hence it is very important to select the right threshold value to help the model make the right decision to detect the anomalies [37]. Selecting thresholds is very tricky and must be done correctly else the model will create unnecessary false alarms which is highly unwanted and makes the model unreliable in situations where it is set either too high or too low [37]. Also, thresholds are not always 0.5 which is the common value, they must be tuned to obtain the optimal value which is evaluated using the metrics. In this experiment, 3 thresholds are defined and are evaluated using the confusion matrix and some other metrics such a recall, area under ROC, precision and f1 score. Even though other studies mentioned the confusion matrix as the suitable performance measure, these other metrics were added to evaluate to obtain a wider perspective.

5.1 Threshold of 1.3

The matrix below shows that the model correctly predicted the legitimate class and correctly predicted some fraudulent class. Out of the number of test set, the number of correctly predicted legitimate class was 53008 transactions and the number of correctly predicted fraud was 89 transactions. Another good thing about this is that it has a good recall value of 0.91 and area under ROC of 0.92, even though area under ROC especially might

not exactly be a good metric. The problem with this is that it has a very high false positive number which is the number of valid transactions wrongly predicted as fraud. It Predicted a false positive value of 3856 which can cause a huge false alarm and hence must be decreased.

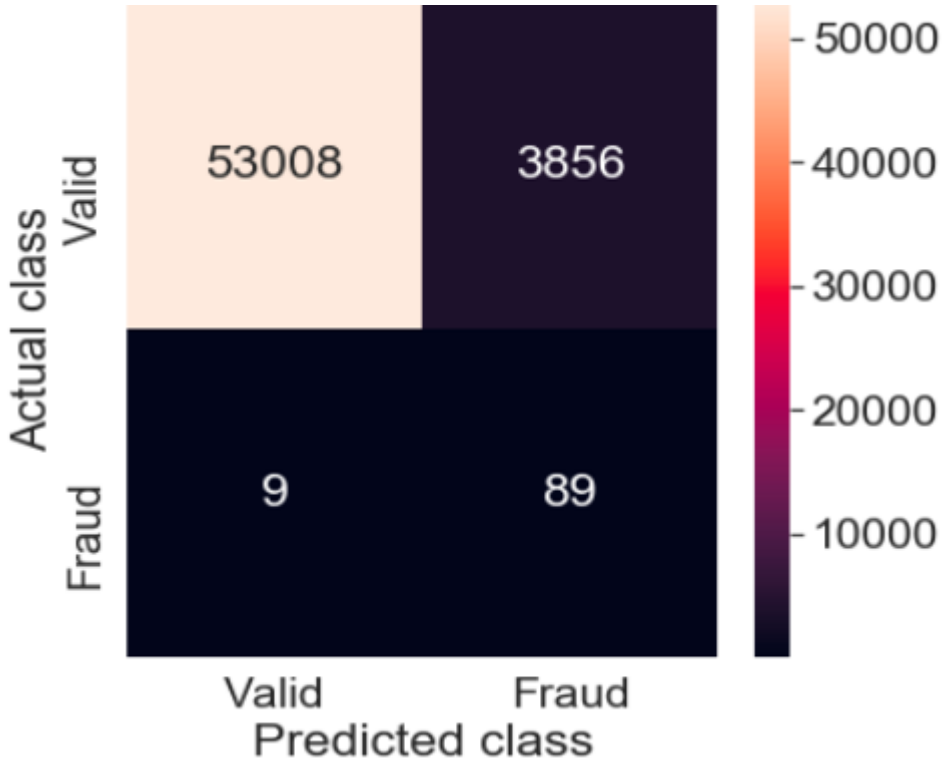

Figure 5.1: Confusion diagram using threshold of 1.3

```
print(classification_report(error_df.true_class,y_pred))
```

	precision	recall	f1-score	support
0	1.00	0.93	0.96	56864
1	0.02	0.91	0.04	98
micro avg	0.93	0.93	0.93	56962
macro avg	0.51	0.92	0.50	56962
weighted avg	1.00	0.93	0.96	56962

Figure 5.2: Summary evaluation report using threshold of 1.3

5.2 Threshold of 4.3

The model using this threshold as well correctly predicted some valid transactions and some fraud transactions. The number of correctly predicted valid transaction was 56019 and the number of correctly predicted fraud transactions was 69. The recall rate was not as good as the first as it gave a rate of 0.71 and the rest of the other metrics were incredibly low. However, the number of false positives which is the number of wrongly predicted valid transactions significantly reduced from 3856 to 845 transactions. This was leading towards the goal of reducing the number of wrongly predicted valid transactions.

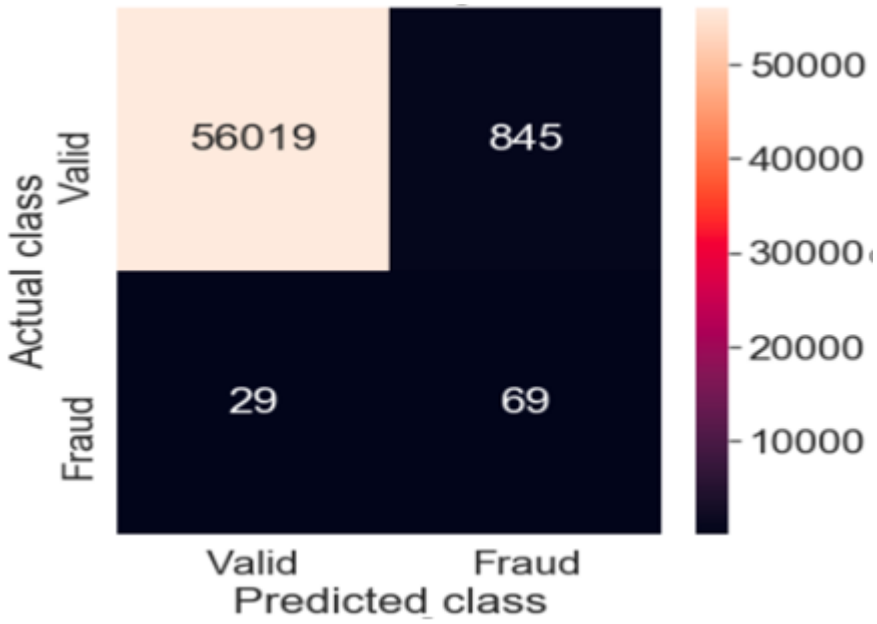

Figure 5.3: Confusion diagram using threshold of 4.3

```
print(classification_report(error_df.true_class,y_pred))
```

	precision	recall	f1-score	support
0	1.00	0.99	0.99	56864
1	0.08	0.70	0.14	98
micro avg	0.98	0.98	0.98	56962
macro avg	0.54	0.84	0.56	56962
weighted avg	1.00	0.98	0.99	56962

Figure 5.4: Summary evaluation report using threshold of 4.3

5.3 Threshold of 4.9

Since an increase in threshold from 1.4 to 4.3 significantly reduced the number of wrongly predicted valid transactions, the threshold was further increased to 4.9 and this also decreased the false positives from 845 to 697 transactions. The model also increased the correctly predicted number of valid transactions from 53008, 56019 to 56167 as shown below. Also, the cost of trying to reduce the number of wrongly predicted valid transaction was at the expense of correctly predicting the number of fraudulent transactions. It can be seen that the number of correctly predicted fraud transactions reduced from 89, 69 and now to 65 transactions in the test set. Even though reducing false positives is a goal, the main goal is for the model to correctly predict a fraudulent transaction hence threshold tuning was maintained at the value of 4.9. Again, the other evaluation metrics kept on scoring lower, but this is not an issue since the confusion matrix is a better evaluator in this case.

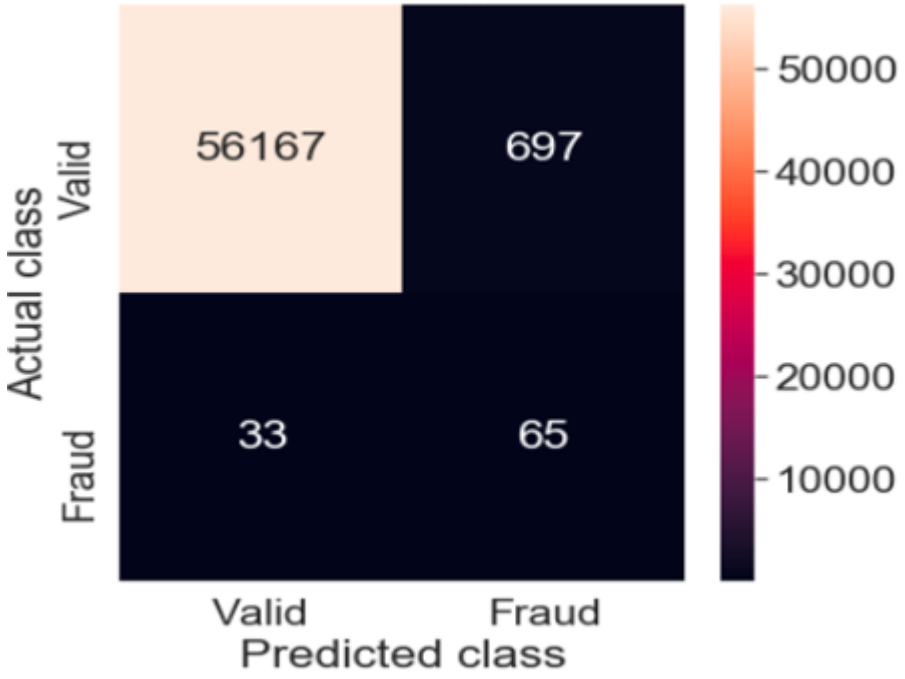

Figure 5.5: Confusion Diagram using threshold of 4.9

```
print(classification_report(error_df.true_class,y_pred))
```

	precision	recall	f1-score	support
0	1.00	0.99	0.99	56864
1	0.09	0.66	0.15	98
micro avg	0.99	0.99	0.99	56962
macro avg	0.54	0.83	0.57	56962
weighted avg	1.00	0.99	0.99	56962

Figure 5.1: Summary Evaluation report using threshold of 4.9

Chapter 6: Conclusion and Future Works

This study contributed to the research in detecting credit card fraud by providing an approach to the detection process. Current research has focused on exploring different data mining techniques that can predict fraud better. Others have also focused on credit card fraud prevention instead of detection. This study focused on finding an adaptive credit card detection model that also reduces the number of wrongly predicted legitimate transactions. The study was undertaken because a model that can easily adapted to the adapting nature of fraudsters was needed. Also, a model that reduces the inconvenience that comes with credit card detection systems was required as well. Most credit card detection systems wrongly flag a valid transaction which has caused so many confusions among credit card users.

This happens because data scientists and researchers usually focus on location and amount when performing feature engineering [38] in credit card detection. Hence if a user uses more than a 'normal' amount or the user changes location from a usual location, the user's transaction is automatically flagged as fraud. A principal research scientist said reducing false positives is the major challenge in credit card fraud detection [38]. Therefore, this research contributed to the technique that can be used in reducing false positives. The use of the autoencoder neural network also contributed to adaptability which is a nature of the neural network itself as mentioned in literature review.

Also, during the detection process, the distribution of time did not seem to affect fraud in any way hence it can be concluded that time of transaction does not help in detecting fraud. Time of transaction was dropped, and the amount of transaction was maintained for feature engineering. This answers the research questions; Is time a factor in conducting fraudulent activities? Is amount a factor in conducting fraudulent activities? It can also be concluded that the goal of this project has been partially met since the model needs to be

able to detect more fraud cases from the test set even though the number it predicts is a good amount.

6.1 Future Works

The system requirements of this project would be fully met if the system is functioning within an application where users make payments and receive payments. In order for users to be able to decline or permit a transaction, the detection model needs to be built within the application. Hence in the future, the detection model can be reconstructed as a plugin to be used in a website, web app or a mobile app. This would also allow the testing of the model when it comes to it ability to detect new variations of fraud.

References

[1] Ekrem D. and Hamdi O.M, "Detecting credit card fraud by genetic algorithm and scatter search," EXPERT SYSTEMS WITH APPLICATIONS, vol. 38, no. 10, pp. 13057-13063, 2011

[2] M. Rafalo, "Real-time fraud detection in credit card transactions," Data Science Warsaw, 2017.

[3] Z. Zojaji, R. E. Atani, and A. H. Monadjemi, "A survey if Credit Card Fraud Detection Techniques: Data and Technique Oriented Perspective," pp. 1-26, 2016.

[4] Linda D., Hussein A., and Pointon J., "Credit card fraud and detection techniques: a review," Banks and Bank Systems, vol. IV, no. 2, pp. 57-68, 2009

[5] Sharma A. and Panigrahi P.K, "A review of financial accounting fraud detection based on data mining techniques," International Journal of Computer Applications, vol. 39, no. 1, pp. 37-47, 2012

[6] Dheepa V. and Dhanapal R., "Analysis of Credit Card Fraud Detection Methods," International Journal of Recent Trends in Engineering, vol. 2, no. 3, pp. 126-128, 2009

[7] Rama K. K and Uma D. D., "Fraud Detection of Credit Card Payment System by Genetic Algorithm," International Journal of Scientific & Engineering Research, vol. 3, no. 7, pp. 1-6, 2012

[8] Gayathri R. and Malathi A., "Investigation of Data Mining Techniques in Fraud Detection: Credit Card," International Journal of Computer Applications, vol. 82, no. 9, pp. 12-15, 2013.

[9] K. B. Anderson, E. Durbin, and M. A. Salinger, "Identity Theft," *Journal of Economic Perspectives*, vol. 22, no. 2, pp. 171–192, Jun. 2008, doi: 10.1257/jep.22.2.171.

[10] M. Zareapoor and P. Shamsolmoali, "Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier," *Procedia Computer Science*, vol. 48, pp. 679–685, Jan. 2015, doi: 10.1016/j.procs.2015.04.201.

[11] L. Qibei. & J. Chunhua. (2011). Research on Credit Card Fraud Detection Model Based on Class Weighted Support Vector Machine. Journal of Convergence Information Technology, 6(1), 62-68.

[12] J. Piotr., A.M. Niall, J.D. Hand, C. Whitrow, J. David (2008). Off the peg and bespoke classifiers for fraud detection. Computational

[13] R.J. Bolton, D.J. Hand (2001). Unsupervised profiling methods for fraud detection. In Conference on credit scoring and credit control, Edinburgh

[14] D. S. Sisodia, N. K. Reddy, and S. Bhandari, "Performance Evaluation of Class Balancing Techniques for Credit Card Fraud Detection." IEEE Int. Conf. Power, Control. Signals Instrum. Eng., pp. 2747-2752, 2017.

[15] "The Nilson Report – Card Fraud Losses Reach $27.85 Billion." https://nilsonreport.com/mention/407/1link/.

[16] A. Thennakoon, C. Bhagyani, S. Premadasa, S. Mihiranga, and N. Kuruwitaarachchi, "Real-time Credit Card Fraud Detection Using Machine Learning," in *2019 9th International Conference on Cloud Computing, Data Science Engineering (Confluence)*, Jan. 2019, pp. 488–493, doi: 10.1109/CONFLUENCE.2019.8776942.

[17] S. Parusheva, "Card-not-present fraud – challenges and counteractions," *Narodnostopanski archiv*, no. 2, pp. 40–56, 2015. [Online]. Available: https://www.ceeol.com/search/article-detail?id=422707.

[18] R. J. Bolton and D. J. Hand, "Statistical Fraud Detection: A Review," *Statistical Science*, vol. 17, no. 3, pp. 235–249, 2002, Available: https://www.jstor.org/stable/3182781.

[19] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data mining for credit card fraud: A comparative study," *Decision Support Systems*, vol. 50, no. 3, pp. 602–613, Feb. 2011, doi: 10.1016/j.dss.2010.08.008.

[20] Z. Abu, K. I. Bolannwa, "Evaluation of Prospect and Challenges of Cashless Policy. The Case of Commercial Banks in Nigeria | ArmgPublishing." http://armgpublishing.sumdu.edu.ua/journals/fmir/volume-2-issue-4/article-9/.

[21] R. Bala and D. Garg, "Credit Card Fraud Detection using Logistic Regression and Bayesian Network," 2019.

[22] M. Zareapoor, K. R Seeja & A. Alam, "Analyzing Credit Card:Fraud Detection Techniques Based On Certain Design Criteria. International Journal of Computer Application", vol. 52, no. 3, pp. 35 – 613, Aug. 2012, DOI: 0.5120/8184-1538

[23] G. Potamitis, "Design and Implementation of a Fraud Detection Expert System using Ontology-Based Techniques. A dissertation submitted to the University of Manchester for the degree of Master of Science in the Faculty of Engineering and Physical Sciences".

[24] A. Sharma and P. K. Panigrahi, "A Review of Financial Accounting Fraud Detection based on Data Mining Techniques," *ArXiv*, 2012, doi: 10.5120/4787-7016.

[25] E. Turban, J.E. Aronson, T. P. Liang., & R. Sharda, 'Decision Support and Business Intelligence Systems", Eighth edition, Pearson Education, 2007.

[26] E. W. T. Ngai, Y. Hu, Y. H. Wong, Y. Chen, and X. Sun, "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature," *Decis. Support Syst.*, vol. 50, no. 3, pp. 559–569, Feb. 2011, doi: 10.1016/j.dss.2010.08.006.

[27] I.-C. Yeh and C. Lien, "The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients," *Expert Syst. Appl.*, 2009, doi: 10.1016/j.eswa.2007.12.020.

[28] Z. Zojaji, R. E. Attani, and A. H. Monadjemi, Survey of Credit Card Fraud Detection Techniques: Data and Technique Oriented Perspective," pp. 1-26, 2016.

[29] A. Thennakoon, C. Bhagyani, S. Premadasa, S. Mihiranga, and N. Kuruwitaarachchi, "Real-time Credit Card Fraud Detection Using Machine Learning," in *2019 9th International Conference on Cloud Computing, Data Science Engineering (Confluence)*, Jan. 2019, pp. 488–493, doi: 10.1109/CONFLUENCE.2019.8776942.

[30] J. West & M. Bhattacharya, "An investigation on experimental Issues in Financial Fraud Mining," Procedia Comput. Sci., vol. 80, pp. 1734-1744, 2016.

[31] R. J. Bolton and D. J. Hand, "Statistical Fraud Detection: A Review," *Statistical Science*, vol. 17, no. 3, pp. 235–249, 2002, Available: https://www.jstor.org/stable/3182781.

[32] V. N. Marivate, F. V. Nelwamodo, and T. Marwala, "Autoencoder, principal component analysis and support vector regression for data imputation," arXiv preprint arXiv:0709.2506,2007

[33] Z. Junyi, Z. Jinliang & J. Ping, "Credit Card Fraud Detection Using Autoencoder Neural Network," 2019.

[34] Ghosh and Reilly, "Credit card fraud detection with a neural-network," in *1994 Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences*, Jan. 1994, vol. 3, pp. 621–630, doi: 10.1109/HICSS.1994.323314.

[35] E. Aleskerov, B. Freisleben, and B. Rao, "CARDWATCH: a neural network based database mining system for credit card fraud detection," in *Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFEr)*, Mar. 1997, pp. 220–226, doi: 10.1109/CIFER.1997.618940.

[36] "Credit Card Fraud Detection using Autoencoders in Keras | TensorFlow for Hackers (Part VII)," *Curiousily*. https://curiousily.com/posts/credit-card-fraud-detection-using-autoencoders-in-keras/ (accessed May 20, 2020).

[37] O. Ismaila, “Investigating the Effects of Threshold in Credit Card Fraud Detection System,” *International Journal of Engineering and Technology*. Available: https://www.academia.edu/39140076/Investigating_the_Effects_of_Threshold_in_Credit_Card_Fraud_Detection_System.

[38] R. Matheson, “Reducing false positives in credit card fraud detection,” *MIT News Office*, Sep. 2018. Available: http://news.mit.edu/2018/machine-learning-financial-credit-card-fraud-0920.

[39] L. Delamaire, H. Abdou, and J. Pointon, “(PDF) Credit card fraud and detection techniques: A review,” *OAI*, Jan. 2009. Available: https://www.researchgate.net/publication/40227011_Credit_card_fraud_and_detection_techniques_A_review/references.