CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING AND DEEP LEARNING TECHNIQUES

Main Project work submitted to Avinashilingam Institute for Home Science and Higher Education for Women

MASTER OF SCIENCE IN INFORMATION TECHNOLOGY

Submitted By

M. Kirubavathi (19PIT005)

Under the guidance of

Mrs. N. Krishnaveni M.Sc., M.Phil., Ph.D,

Assistant Professor, Department of Information Technology

AVINASHILINGAM INSTITUTE FOR HOME SCIENCE AND

HIGHER EDUCATION FOR WOMEN

SCHOOL OF PHYSICAL SCIENCES AND COMPUTATIONAL SCIENCES DEPARTMENT OF INFORMATION TECHNOLOGY

Coimbatore-641043

May 2021

DECLARATION

I hereby declare that the project entitled **"CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING AND DEEP LEARNING TECHNIQUES"** is a record of the original work done by Kirubavathi. M (19PIT005) under the guidance of Mrs. N. Krishnaveni M.Sc., M.Phil., SET., Assistant professor, Department of Information Technology, School of Physical Sciences and Computational Sciences, Avinashilingam Institute for Home Science and Higher Education for Women, in the partial fulfilment for the **degree of Master of Science in Information Technology** and this project has not formed the basis for any Degree/Diploma/Associates.

Place : Coimbatore

Date :

Signature of the Candidate

_____

Countersigned by

Mrs. N. Krishnaveni M.Sc., M.Phil., SET

Assistant Professor, Department of Information Technology,

School of Physical Sciences and Computational Sciences

CERTIFICATE

TO WHOMSOEVER IT MAY CONCERN

This is to certify the student **Ms.Kirubavathi M (19PIT005)** final year **MSC (IT)** In Avinashilingam Institute for Home Science and Higher Education for Women University, Coimbatore is permitted to do her project tin **"Credit Card Fraud Detection using Machine Learning and Deep Learning"** followed in our concern starts from March 2021 to April 2021.

Wish her for the best!

Signature of the HOD

Signature of the Guide

Signature of External Examiner

DURATECH SOLUTIONS
Products | Consulting | Training
Innovate Execute Empower

Date: 30-04-2021

TO WHOMSOEVER IT MAY CONCERN

This is to certify the student **Ms. Kirubavathi. M (19PIT005)** Pursuing her final year **MSC INFORMATION TECHNOLOGY in AVINASHILINGAM INSITITUTE FOR HOME SCIENCE & HIGHER EDUCATION FOR WOMEN, COIMBATORE** has completed her project entitled **"CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING AND DEEP LEARNING TECHNIQUES"** in our concern starts on February 2021 to April 2021

Wish her for the best!

DuraTech

DURATECH SOLUTIONS COIMBATORE

Project Manager

ARPEE Center, 320 N, NSR Rd, Saibaba Colony, Coimbatore-11.
http://duratechsolutions.in - info@duratechsolutions.in
+91 9994993885

ACKNOWLEDGEMENT

I would like to express my sincere thanks to God Almighty, for his constant love and grace that he has showered upon me, which kept me in good health, and sound mind without which my project would not have reached a successful end.

I would like to express my deep sense of reverential gratitude and sincere thanks to

Dr. S.P. Thyagarajan, Chancellor, Chancellor, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing all facilities during my course of study.

I owe my great deal of gratitude to **Dr.PremavathyVijayan M.Sc., M.Ed., Dip. Spl. Edn., M.Phil., Ph.D., Vice Chancellor,** Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for extending all resources that facilitated the smooth conduct of the project study.

I express my gratitude to **Dr. S. Kowsalya, Registrar,** Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, for providing all facilities and support necessary for the study.

I wish to extend my sincere thanks **Dr. K. Udaya Chandrika M.Sc., M.Phil., Ph.D., Dean, School of Physical Sciences and Computational Sciences**, for her support and valuable guidance.

I heartily thank my esteemed project guide **Mrs. N. Krishnaveni M.Sc., M.Phil., SET., Assistant professor, Department of Information Technology,** for imparting tremendous assistance and well-timed support for triumph of our project.

I would like to express my sincere gratitude to all the staff members of the Department of Information Technology, for their constant encouragement and for the opportunity to do our project in this esteemed university. Last yet importantly, i would like to thank my parents, family members, friends and all well-wishers for their kind inspiration, blessings and encouragement during the course of project.

ABSTRACT

Frauds in credit card transactions are common today as most of us are using the credit card payment methods more frequently. Each fraud is addressed using a series of machine learning models and the best method is selected via an evaluation. This evaluation provides a comprehensive guide to selecting an neural network algorithm with respect to the type of the frauds and we illustrate the evaluation with an appropriate performance measure. Another major key area that we address in our project is real-time credit card fraud detection finding the better accuracy. This paper aims in using the multiple algorithms of Machine learning such as support vector machine (SVM), k-nearest neighbor (Knn) and artificial neural network (ANN) in predicting the occurrence of the fraud. Further, we conduct a differentiation of the accomplished supervised machine learning and deep learning techniques to differentiate between fraud and non-fraud transactions.

The data which is being used in this study is analyzed in two main ways: as categorical data and as numerical data. The dataset originally comes with categorical data. The raw data can be prepared by data cleaning and other basic preprocessing techniques. First, classification data can be transformed into numerical data and then appropriate techniques are applied to do the evaluation. Secondly, classifiers data is used in the machine learning techniques to find the accuracy using various algorithm. Propose solution to the **finding accuracy using machine learning and deep learning techniques**. It is done using **Artificial Neural Networks** such as **better accuracy.**

INDEX AND CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

"Fraud detection is a set of activities that are taken to prevent money or property from being obtained through false pretenses." Fraud can be committed in different ways and in many industries. The majority of detection methods combine a variety of fraud detection datasets to form a connected overview of both valid and non-valid payment data to make a decision. This decision must consider IP address, geolocation, device identification, "BIN" data, global latitude/longitude, historic transaction patterns, and the actual transaction information. In practice, this means that merchants and issuers deploy analytically based responses that use internal and external data to apply a set of business rules or analytical algorithms to detect fraud.

Credit Card Fraud Detection with Machine Learning is a process of data investigation by a Data Science team and the development of a model that will provide the best results in revealing and preventing fraudulent transactions. This is achieved through bringing together all meaningful features of card users' transactions, such as Date, User Zone, Product Category, Amount, Provider, Client's Behavioral Patterns, etc. The information is then run through a subtly trained model that finds patterns and rules so that it can classify whether a transaction is fraudulent or is legitimate [16].

Machine Learning-based Fraud Detection:

- Detecting fraud automatically
- Real-time streaming
- Less time needed for verification methods
- Identifying hidden correlations in data

Conventional Fraud Detection:

- The rules of making a decision on determining schemes should be set manually.
- Takes an enormous amount of time
- Multiple verification methods are needed; thus, inconvenient for the user
- Finds only obvious fraud activities

AI Fraud Detection System Implementation Steps:

- If the probability is less than 10%, the transaction is allowed.
- If the probability is between 10% and 80%, an additional authentication factor (e.g. a one-time SMS code, a fingerprint, or a Secret Question) should be applied.
- If the probability is more than 80%, the transaction is frozen, so it should be processed manually.

Requirements for Fraud Detection with AI-based Methods

To run an AI-driven strategy for Credit Card Fraud Analytics, a number of critical requirements should be met. These will ensure that the model reaches its best detection score.

Amount of data.

Training high-quality Machine Learning models requires significant internal historical data. That means if you do not have enough previous fraudulent and normal transactions, it would be hard to run a Machine Learning model on it because the quality of its training process depends on the quality of the inputs. Because it is rarely the case that a training set contains an equal amount of data samples in two classes, dimensionality reduction or data augmentation techniques are used for that.

Quality of data.

Models may be subject to bias based on the nature and quality of historical data. This statement means that if the platform maintainers did not collect and sort the data neatly and properly or even mixed the information of fraudulent transactions with the information of normal ones, that is likely to cause a major bias in the model's results.

The integrity of factors.

If you have enough data that is well-structured and unbiased, and if your business logic is paired nicely with the Machine Learning model, the chances are very high that fraud detection will work well for your customers and your business.

Advanced Credit Card Fraud Identification Methods and Their Advantages

Advanced Credit Card Fraud Identification Methods are split into:

- Unsupervised. Such as PCA, LOF, One-class SVM, and Isolation Forest.
- Supervised. Such as Decision Trees (e.g. XGBoost and LightGBM), Random Forest, and KNN.

We've covered the basic vision of how Machine Learning for fraud detection works. Let's now dig deeper into the exact models that make it possible.

Unsupervised.

Unsupervised Machine Learning methods use unlabeled data to find patterns and dependencies in the credit card fraud detection dataset, making it possible to group data samples by similarities without manual labeling.

PCA (Principal Component Analysis) enables the execution of an exploratory data analysis to reveal the inner structure of the data and explain its variations. PCA is one of the most popular techniques for Anomaly Detection.

PCA searches for correlations among features — which in the case of credit card transactions, could be time, location, and amount of money spent — and determines which combination of values contributes to the variability in the outcomes. Such combined feature values allow the creation of a tighter feature space named *principal components*.

LOF (Local Outlier Factor) is the score factor that helps understand how high the chance is for a certain data sample to be an outlier (anomaly). This is another of the most popular Anomaly Detection methods.

To calculate LOF, the number of neighboring data points is considered to figure out its density and compare it to the density of other data points. If a certain data point has a substantially low density compared to its close neighbors, it is an outlier.

One-class SVM (Support Vector Machine) is a classification algorithm that helps to identify outliers in data. This algorithm allows one to deal with imbalanced data-related issues such as Fraud Detection.

The idea behind One-class SVM is to train only on a solid amount of legitimate transactions and then identify anomalies or novelties by comparing each new data point to them.

Isolation Forest (IF) is an Anomaly Detection method from the Decision Trees family. The main idea of IF, which differentiates it from other popular outlier detection algorithms, is that it precisely detects anomalies instead of profiling the positive data points. Isolation Forest is built of Decision Trees where the separation of data points happens first because of randomly selecting a split value amidst the minimum and maximum value of the chosen feature.

Subsequently, if we have a set of legitimate transactions, the Isolation Forest algorithm will define fraudulent credit card transactions because of their values — which are often very different from the values positive transactions have (i.e. they take place further away from the normal data points in the feature space).

Supervised

Supervised ML methods use labeled data samples, so the system will then predict these labels in future unseen before data. Among supervised ML fraud identification methods, we define Decision Trees, Random Forest, KNN, and Naive Bayes.

K-Nearest Neighbors is a Classification algorithm that counts similarities based on the distance in multi-dimensional space. The data point, therefore, will be assigned the

class that the nearest neighbors have.

XGBoost (Extreme Gradient Boosting) and **Light GBM (Gradient Boosting Machine)** are a single type of gradient-boosted Decision Trees algorithm, which was created for speed as well as maximizing the efficiency of computing time and memory resources. This algorithm is a blending technique where new models are added to fix the errors caused by existing models.

To classify a transaction as a fraudulent charge, the result (probability) of many Decision Trees is summarized — whereas every future tree improves its results based on of the errors made by its predecessors.

Random Forest is a classification algorithm that is comprised of many Decision Trees. Each tree has nodes with conditions, which define the final decision based on the highest value.

The Random Forest algorithm for fraud detection and prevention has two cardinal factors that make it good at predicting things. The first one is randomness, meaning that the rows and columns of data are chosen randomly from the dataset and fit into different Decision Trees. Say Tree Number 1 receives the first 1,000 rows, Tree Number 2 receives Rows 4,000 to 5,000, and the Tree Number 3 has Rows 8,000 to 9,000[16].

Deep learning vs Machine learning

The easiest takeaway for understanding the difference between machine learning and deep learning is to know that **deep learning *is* machine learning**.

More specifically, deep learning is considered an evolution of machine learning. It uses a programmable neural network that enables machines to make accurate decisions without help from humans.

But for starters, let's first define machine learning.

Machine learning

Machine learning is an application of AI that includes algorithms that parse data, learn from that data, and then apply what they've learned to make informed decisions.

An easy example of a machine learning algorithm is an on-demand music streaming service. For the service to make a decision about which new songs or artists to recommend to a listener, machine learning algorithms associate the listener's preferences with other listeners who have a similar musical taste. This technique, which is often simply touted as AI, is used in many services that offer automated recommendations.

Machine learning fuels all sorts of automated tasks that span across multiple industries, from data security firms that hunt down malware to finance professionals who want alerts for favorable trades. The AI algorithms are programmed to constantly be learning in a way that simulates as a virtual personal assistant—something that they do quite well.

Deep learning

Deep learning is a subfield of machine learning that structures algorithms in layers to create an "artificial neural network" that can learn and make intelligent decisions on its own.

The difference between deep learning and machine learning

In practical terms, deep learning is just a subset of machine learning. In fact, deep learning is machine learning and functions in a similar way (hence why the terms are sometimes loosely interchanged). However, its capabilities are different.

While basic machine learning models do become progressively better at whatever their function is, they still need some guidance. If an AI algorithm returns an inaccurate prediction, then an engineer has to step in and make adjustments. With a deep learning model, an algorithm can determine on its own if a prediction is accurate or not through its own neural network.

The differences between the two:

- Machine learning uses algorithms to parse data, learn from that data, and make informed decisions based on what it has learned
- Deep learning structures algorithms in layers to create an "artificial neural network" that can learn and make intelligent decisions on its own
- Deep learning is a subfield of machine learning. While both fall under the broad category of artificial intelligence, deep learning is what powers the most human-like artificial intelligence [17].

Benefits of Deep Learning

Following are the benefits or **advantages of Deep Learning**:
➨ Features are automatically deduced and optimally tuned for desired outcome. Features are not required to be extracted ahead of time. This avoids time consuming machine learning techniques.
➨Robustness to natural variations in the data is automatically learned.
➨ The same neural network based approach can be applied to many different

applications and data types.
➡ Massive parallel computations can be performed using GPUs and are scalable for large volumes of data. Moreover it delivers better performance results when amount of data are huge.
➡The deep learning architecture is flexible to be adapted to new problems in the future.

Drawbacks of Deep Learning

Following are the drawbacks or **disadvantages of Deep Learning**:
➡It requires very large amount of data in order to perform better than other techniques.
➡ It is extremely expensive to train due to complex data models. Moreover deep learning requires expensive GPUs and hundreds of machines. This increases cost to the users.
➡ There is no standard theory to guide you in selecting right deep learning tools as it requires knowledge of topology, training method and other parameters. As a result it is difficult to be adopted by less skilled people.
➡It is not easy to comprehend output based on mere learning and requires classifiers to do so. Convolutional neural network based algorithms perform such tasks.

Advantages of Machine learning

1. Easily identifies trends and patterns

Machine Learning can review large volumes of data and discover specific trends and patterns that would not be apparent to humans. For instance, for an e-commerce website like Amazon, it serves to understand the browsing behaviors and purchase histories of its users to help cater to the right products, deals, and reminders relevant to them.

2. No human intervention needed (automation)

With ML, you don't need to babysit your project every step of the way. Since it means giving machines the ability to learn, it lets them make predictions and also improve the algorithms on their own.

3. Continuous Improvement

As ML algorithms gain experience, they keep improving in accuracy and efficiency. This lets them make better decisions. Say you need to make a weather forecast model. As the amount of data you have keeps growing, your algorithms learn to make more accurate predictions faster.

4. Handling multi-dimensional and multi-variety data

Machine Learning algorithms are good at handling data that are multi-dimensional and multi-variety, and they can do this in dynamic or uncertain environments.

Disadvantages of Machine Learning

With all those advantages to its powerfulness and popularity, Machine Learning isn't perfect. The following factors serve to limit it:

1. Data Acquisition

Machine Learning requires massive data sets to train on, and these should be inclusive/unbiased, and of good quality. There can also be times where they must wait for new data to be generated.

2. Time and Resources

ML needs enough time to let the algorithms learn and develop enough to fulfill their purpose with a considerable amount of accuracy and relevancy. It also needs massive resources to function. This can mean additional requirements of computer power for you.

3. Interpretation of Results

Another major challenge is the ability to accurately interpret results generated by the algorithms. You must also carefully choose the algorithms[18].

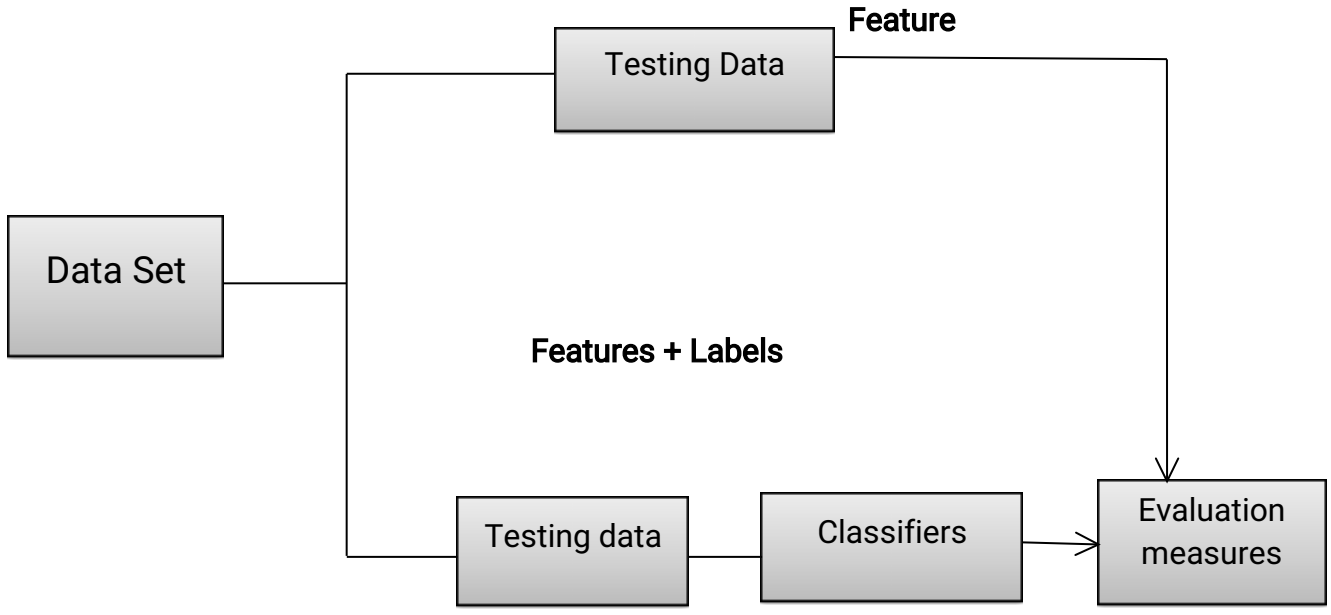


Figure 1.1 : Fundamental Steps in credit card fraud detection

1.2 OBJECTIVES OF THE PROJECT

- ✓ This project is based on applying different **Classification Techniques on Credit Card Fraud Detection** for finding the better accuracy.
- ✓ To apply the pre-processing to remove noise and cleaning data Processing of **Data Analysis.**
- ✓ To select the data and **Split the Features** in to training data and testing data.
- ✓ Building the Different types **Classification models** which is the Combination of Machine Learning and Deep Learning techniques.
- ✓ To Calculating the Random forest algorithm, Decision tree algorithm and Support Vector Machine using some **parameters and get the accuracy**.
- ✓ Finally **Comparing the Classifiers** in ANN techniques using Evaluation Measures.

1.3 ABOUT THE SOFTWARE

Spyder platform For data analysts, visualization and presentation of their hard worked ML projects are at least as important as the analysis part of their projects. Streamlit is an open-source Python library that makes it easy to build beautiful custom **web-apps for machine learning and data science.** It lets our app update live as you edit and save your file. All need is our favorite editor (I prefer Spyder which is included with Anaconda. Visit https://www.spyder-ide.org/ to look at this amazing editor) and a browser. will build a fraud detection model from scratch and look at the steps to deploy it using streamlit.

1.4 COMPANY DETAILS

Name of the Company : DURATECH SOLUTION
Name of the Manager : Mr. R.Srinivasa Prabhu
Company Address : Nehru Street, Ramnagar, Coimbatore-641009
Contact Number : 0422-4357105
Email Id : tn.cbe.ramnagar@duratechindia.com
Year of Establishment : 1988
Working hours : 9.30 AM - 7 PM
Website Address. : www.duratechindia.com
Domain :

1. Electronic Design Automation
2. Industrial Automation
3. IT Infrastructure Management
4. Software Development

1.5 SYSTEM REQUIREMENTS

Hardware Requirements

- ✓ RAM: 4GB and Higher
- ✓ Processor: HP
- ✓ Hard Disk: 500GB Minimum

Software Requirements

- ✓ OS: Windows10
- ✓ Python IDE : python 3.7.x and above
- ✓ Spyder Notebook
- ✓ Tensor flow tools, keras and pip to be installed for 3.7 and above
- ✓ Language : Python

CHAPTER 2

LITERATURE STUDY

2.1 BACKGROUND STUDY

The authors of Kuldeep randhawa, From: Journal of network and computer applications, 2018 There is much current research in the machine learning and statistics communities on algorithms for decision tree classifiers. A model can predict the value of the majority class for all predictions and achieve a high classification accuracy. The datasets with the largest range of error rates are range from 0.005 to 0.890. possible improvements such as more refined data and more accurate algorithm **[1]**.

The authors of Shashkant Gupta 2018, This paper points out an important source of inefficiency in Smola and Schölkopf's sequential minimal optimization (SMO) algorithm for support vector machine (SVM) regression that is caused by the use of a single threshold value. Implimented all these methods in C and ran them using the gcc on a P3 450 MHz Linux machine. The value t=0.01 was used for all experiements. SVM are usually used for binary classification, and can be extended to do multi-class regression **[2]**.

The authors of “Detection of online fake news using machine learning algorithm” Supervised learning problems can be further grouped into Classification and Regression problems. non-linear regression by constructing a linear regression function in a high dimensional feature space. The price lie between $5000 and $50,000 in units of $1000. Experimental results are then presented which indicates the performance of this algorithm relative to other algorithms **[3]**.

This paper presents The authors of Jolliffe and Jorge Cadima, published by 2016, Principal component analysis (PCA) is a technique commonly used for fault detection and classification Summary statistics from these PCA of all variables together. The authors have investigated a handful of methods but, to date, have found no satisfactory approach. This remain as an open topic and warrants further investigation **[4]**.

the algorithms by Marc Cleasen-2019, for Hyper-parameter tuning refers to the automatic optimization of the hyper-parameters of a ML model. *Hyper-parameters* are all the parameters of a model which are not updated during the learning and are used to configure either the model Hyper parameter optimization results on tasks of training neural network and deep belief networks. Convex task extends 0.018 above and below each point. MRBI task extends the 0.021 above and below each point. Researchers need not restrict themselves to systems of a few variables that can readily be tuned by hand **[5]**.

This paper presents a novel dynamic ensemble learning(DEL) published by Kazi Md 2017, Designing ensemble learners has been recognized as one of the significant trends in the field of data knowledge especially in data science competitions. target variables with having R-squared values of 0.92 and 0.88 respectively. For the future work, designing a methodology which incorporates finding best ensemble weights while tuning the hyper parameters of each base learner is recommended **[6]**.

This paper presents a Combining Unsupervised and Supervised Learning in Credit Card Fraud Detection. Published by Fabrizio Carcillo – 2017. paper concerns the integration of unsupervised techniques with supervised credit card fraud detection classifiers. The novelty of the contribution, beyond its applications in real and sizeable datasets of credit card transactions. The results are not convincing in terms of the global and local approaches. Amore promising outcome is obtained through the cluster approach (notablyin terms of AUC-PR) **[7]**.

This paper is Auto-Encoderand Restricted BoltzmannMachine Apapan Pumsirirat 2018 using method is deep learning based on auto-encoder (AE) is an unsupervised learning algorithm that applies back propagationby setting the inputs equal to the outputs. Result of Auto Encoder Model deep learning report of European Dataset based on H2O framework. AE and RBM can make more accurate AUC for receiver operator characteristics than that observable from the results from the European Dataset **[8]**.

The Analysis of Credit Card Fraud Detection Techniques: based on Certain Design Criteria Salvatore. J 2017., these techniques can be used alone or in collaboration using

ensemble or meta-learning techniques to build classifiers. Fuzzy Darwinian, has a very high accuracy with 100% true positive but with very low processing speed. The result of this survey enables us to build a hybrid approach for developing some effective algorithms which can perform well for the classification problem with variable misclassification costs and with higher accuracy **[9]**.

The Artificial Neural Network Tuned by Simulated Annealing Algorithm Azeem Ush Shan Khan 2018, Once a learning of training model is complete, the model is capable of classifying the unseen online transaction as fraudulent or non-fraudulent in real time. The result of 65% of total fraud case is correctly classified which is a very high percentage in comparison with geneticcombine Simulated Annealing and Genetic Algorithm to create a best model, it will gives better result than any other **[10]**.

A Novel Idea for Credit Card Fraud Detection using Decision Tree PrajalSave 2017, Address Mismatch and Degree of Outlierness are used to analyze the deviation of each incoming transaction from the normal profile of cardholder. we have found out validation of card are genuine and very low false alarm. further strengthened or weakened in the final step using Bayes‟ Theorem, followed by recombination of the calculated probability with initial belief of fraud using advanced combination heuristic **[11]**.

Using AdaBoost and Majority Voting KULDEEP RANDHAWA 2018, A hybrid model consisting of the Multilayer Percep-tron (MLP) neural network, SVM, LOR, and HarmonySearch (HS) optimization. as the rate of fraud detection varies from 32.5% for RT up to 83% forNB. The rate of non-fraud detection is similar to the accuracy rates, For future work which will reduce the number of losses incurred every day in the financial sector **[12]**.

Using Meta-Learning:Issues and Initial Results Philip K. Chan 2017, the methods of classifiers are all candidates to be base classifiers for meta-learning. the 4 base classifiers with the highest True Positive rates (each trained on a 50%/50% fraud/non-fraud distribution). conducted experiments with 50%/50% distribution to solve the

skewed distribution problem on other data sets and have also obtained good results **[13]**.

Application of Classification Models on Credit Card Fraud Detection Eunji kim 2019, neural network researchers have incorporated methods from statistics and numerical analysis into their networks. top decile captures about 59% of the responders using neural networks or logistic regression while only 38.94% of the responders using decision tree. basis for the intelligent authorized anti-fraud strategy, or refuse to authorize and launch investigations to suspicious transactions **[14]**.

Online E-CommerceTransactions Using Recurrent Neural Networks Shuhao Wang 2016, the sessions with the purchase database to filter out those sessions without an order ID. Then we obtain the manual labels whether a session is fraudulent or not. thus accumulate enough fraud samples to train an extremely detailed RNN model that captures not onlythe detailed click information but also the exact sequences. we can further improve the performance of CLUE by building aricher history of a user, including non-purchasing sessions **[15]**.

S.NO	TITLE OF THE PAPER	AUTHOR NAME & YEAR	ALGORITHM METHOD	RESULT	LIMITATION
1	A comparison of prediction accuracy, complexity and training	Kuldeep randhawa 2018	Based on Logistic Regression algorithm.	largest range of error rates from 0.5 to 0.890	possible improvements such as more refined data and more accurate algorithm.

2	Improvements the SMO algorithm to SVM regression	Shashkant Gupta 2018	Sequential Minimal algorithms for SVM.	The value t=0.01 was used for all experiement	extended to do multi class regression.
3	Meta-learing issues and initial results	J. Stolfo 2016	fraud catching rate and false alarm rate algorithm.	Results generated for 1600 runs.	highest true positive rates learned from 50% fraud distribution is the best method fund thus far.
4	Sensitive credit card fraud detection using bayes minimum risk	Alejandro correa bahnsen 2013	This method is compared with state of the art algorithms.	threshold of an algorithm is 50%	sensitive system gives rise to much better fraud detection results in the sense of higher savings.
5	A Realistic Modeling and a Novel Learning Strategy	Olivier Caelen 2018	real-world data stream containing more than 75 million transaction s.	Approaching α=0.1and α=0.9,	semi-supervised learning methods for exploiting in the learning process also few recent un

					labele transactions.
6	Optimized light gradient boosting machine	Altyeb altaher taha 2020	using an optimized light gradient boosting machine.	accuracy (97.40%),	parameter optimization strategy for enhancing the predictive performance of the proposed approach.
7	A Novel Approach Using Aggregation Strategy and Feedback Mechanism	Changjun Jiang 2018	detection process in order to solve the problem drift.	trained by AggRF+FB, AggRF,threshold= 50%.	propose a method to solve the adaptive capacity of the model.
8	Auto-Encoderand Restricted BoltzmannMachine	Apapan Pumsirirat 2018	algorithm that applies back propagation	Auto Encoder Model deep learning based on H2O framework.	AE and RBM can make more accurate AUC for receiver operator characteristics than that observable from the results from

					the European Dataset
9	Analysis of Credit Card Fraud Detection Techniques: based on Certain Design Criteria	Salvatore. J 2017	These ensemble or meta-learning techniques to build classifiers.	Fuzzy has a very high accuracy with 100% true positive.	some effective algorithms which can perform well for the classification problem with variable
10	Artificial Neural Network Tuned by Simulated Annealing Algorithm	Azeem Ush Shan Khan 2018	classifying the unseen online transaction as fraudu or non-fraud in real time	65% of total fraud case is correctly classified.	combine Simulated Annealing and Genetic Algorithm to create a best model, it will gives better result than any other.
11	A Novel Idea for Credit Card Fraud Detection using Decision Tree	PrajalSave 2017	used to analyze the deviation of each incoming transaction cardholder.	we have found out validation of card are genuine and very low false alarm.	calculated probability with initial belief of fraud using advanced combination heuristic
12	Using AdaBoost and Majority Voting	KULDEEP RANDHAW	A hybrid model	as the rateof fraud	For future work which will

		A 2018	consisting of the Multilayer Percep-tron n	detection varies from 32.5% for RT	reduce thenumber of losses incurred every day in the financial sector.
13	Using Meta-Learning:Issues and Initial Results	Philip K. Chan 2017	classifiers are all candidates to be base classifiers formeta-learning. Integrating very complex.	TruePositive rates (each trained on a 50%/50% fraud/non-fraud distribution).	conducted experiments with 50%/50% distribution tosolve the skewed distribution problem on other data setsand have also obtained good results.
14	Application of Classification Models on Credit Card Fraud Detection	Eunji kim 2019	neural network researcher s have incorporate d methods.	logistic regression while only 38.94% of the responders using decision tree.	basis for the intelligent authorized anti-fraud strategy, or refuse to authorize and launch investigations to suspicious transactions.

15	Online E-CommerceTransactions Using Recurrent Neural Networks	Shuhao Wang 2016	the sessions with the purchase database to filter out those sessions without an order ID.	RNN model that captures not onlythe detailed click information but also the exact sequences.	we can further improve the performance of CLUE by building aricher history of a user, including non-purchasing sessions.

Table 2.1 : Summary of Literature Review

Summary

The main goal of thesis was to compare certain machine learning algorithms and deep learning techniques for detection of fraudulent transactions. Hence, comparison was made and it was established Forest algorithm gives the best results 0.991% i.e. best classifies whether transactions are fraud or not. This established using different metrics, such as recall, accuracy and precision. For this kind of problem, it is important to have recall with high value. Feature selection and balancing of the dataset have shown to be extremely important in achieving significant results. Further Further research should focus on different machine learning algorithms such as genetic algorithms, and different types of stacked classifiers, alongside with extensive feature selection to get better results.

CHAPTER 3

METHODOLOGY

- **Classification** allows for the algorithm to learn from a small amount of labeled text documents while still classifying a large amount of unlabeled text documents in the **training** data.
- The goal of classification algorithm to predict a target value for a specific input data set.
- The **Artificial Neural Network** to find the fraud in the credit card transactions. Performance is measured and accuracy is calculated based on prediction [4].

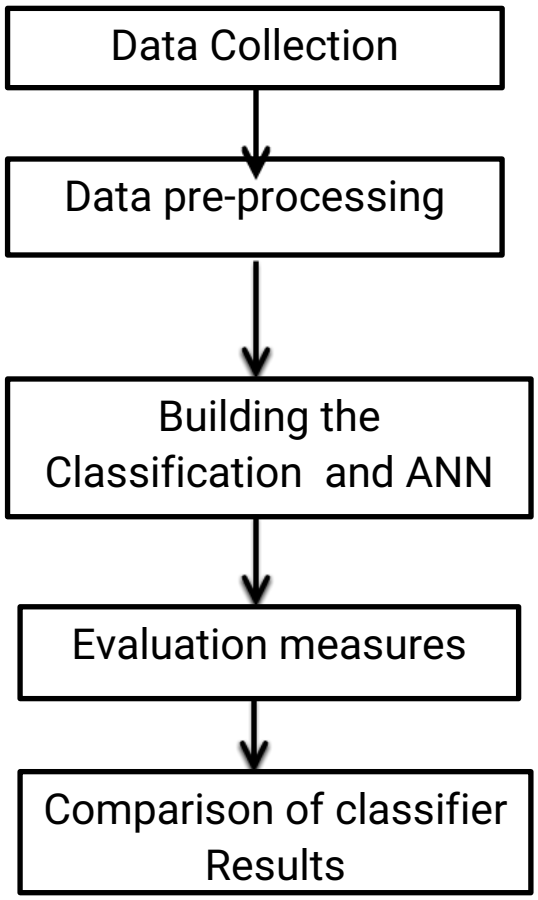


Figure 3.1 : Steps involved in credit card fraud detection

3.1 MODULES DESCRIPTION

3.1.1 Data collection

- **Data** can be collected using three main types of surveys: censuses, sample surveys, and administrative data. Each has advantages and disadvantages.
- **Data collection** is defined as the procedure of collecting, measuring and analyzing accurate insights for research using standard validated techniques.
- A researcher can evaluate their hypothesis on the basis of **collected data**.
- Each column represents a particular variable. Each row corresponds to a given member of the **dataset** in question.
- It lists values for each of the variables, such as height and weight of an object. Each value is known as a datum.
- Data collections from **Google -generated data**, such as Google Analytics or Google Sheets.
- A data source based on a **CSV file**. Metrics and dimensions typed directly into Data Studio.

3.1.2. Data Pre-processing

- To make the process easier, **data preprocessing** is divided into four **stages**: **data** cleaning, **data** integration, **data** reduction, and **data** transformation.
- **It** is a **data** mining technique that transforms raw **data** into an understandable format. Raw **data**(real world **data**) is always incomplete and that **data** cannot be sent through a model. That would cause certain errors. That is why we **need to preprocess data** before sending through a model.
-

Data Cleaning:

The data can have many irrelevant and missing parts. To handle this part, data cleaning is done. It involves handling of missing data, noisy data.

Regression:

Here data can be made smooth by fitting it to a regression function.The regression used may be linear (having one independent variable) or multiple (having multiple independent variables).

Data Reduction:

Since data mining is a technique that is used to handle huge amount of data. While working with huge volume of data, analysis became harder in such cases. In order to get rid of this, we uses data reduction technique. It aims to increase the storage efficiency and reduce data storage and analysis costs [5].

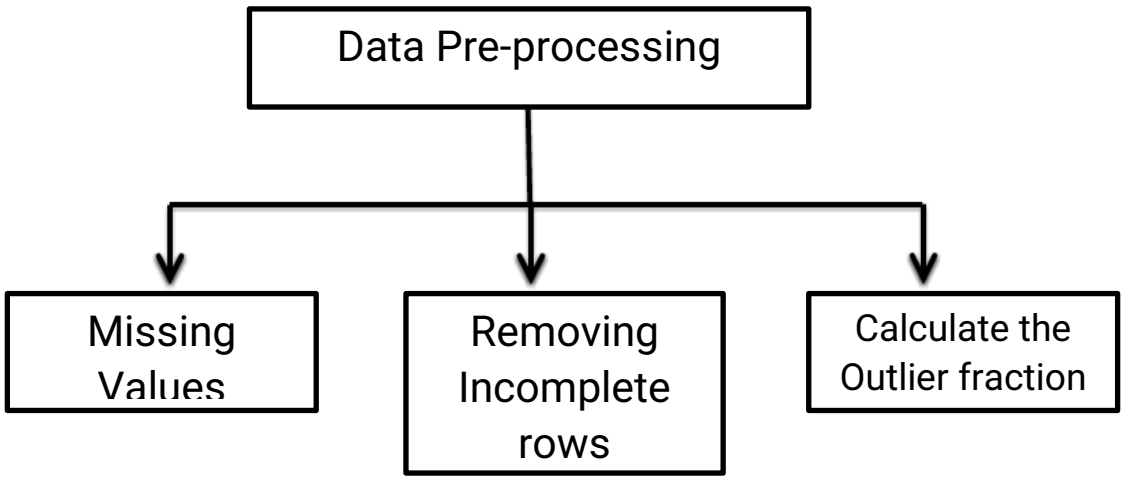


Figure 3.2 : Steps involved in Data pre-processing

3.1.3. Building the Classification Algorithm and ANN

- **Classification algorithms** used in **machine learning** utilize input training data for the purpose of predicting the likelihood or probability that the data that follows will fall into one of the predetermined categories.
- They are artificial **classification**, natural **classification** and phylogenetic **classification**.

- Classification is when the feature to be predicted contains categories of values. Each of these categories is considered as a class into which the predicted value falls and hence has its name, classification.
- **Classification neural networks** used for feature categorization are very similar to fault-diagnosis **networks**, except that they only allow one output response for any input pattern, instead of allowing multiple faults to occur for a given set of operating conditions.

  Classification algorithms include:

  Naive Bayes.

  Logistic regression.

  K-nearest neighbors.

  (Kernel) SVM.

  Decision tree.

  Ensemble learning

- Artificial **Neural Networks** (**ANN**) are multi-layer fully-connected neural nets.
- They consist of an input layer, multiple hidden layers, and an output layer.
- Training this **deep neural network** means **learning** the weights associated with all the edges.
- The **Artificial Neural Network** receives the input signal from the external world in the form of a pattern and image in the form of a vector. These inputs are then mathematically designated by the notations x(n) for every n number of inputs.
- **Neural networks** can be used for either regression or **classification**.
- Regression model a single value is outputted which may be mapped to a set of real numbers meaning that **only** one output **neuron** is required.
- **NLP** is concerned with how computers can process, analyze, and understand human languages [11].

3.1.4 Building the Classification Algorithms

1. Decision tree classifier

- Decision trees are statistical data mining technique that express independent attributes and a dependent attributes logically AND in a tree shaped structure.
- Decision tree usually separates the complex problem into many simple ones and resolves the sub problems through repeatedly using .
- will use the **DecisionTreeClassifier** class from the sklearn library to train and evaluate models. use **X_train and y_train** data for training purposes. X_train is a training dataset with features, and y_train is the target label.
- Credit card fraud detection is a **classification** problem. Target variable values of Classification problems have integer(0,1) or categorical values(fraud, non-fraud). The target variable of our dataset 'Class' has only two labels - 0 (non-fraudulent) and 1 (fraudulent).

ID3 (Iterative Dichotoniser3)

- ID3 is one of the most common decision tree algorithm.
- Algorithm iteratively divides attributes into two groups which are the most dominant attribute and others to construct a tree.
- Then, it calculates the entropy & information gains of each attribute. In this way, the most dominant attribute can be founded.
- After then, the most dominant one is put on the tree as decision node.
- Entropy & Gain scores would be calculated again among the other attributes.
- Procedure continues until reaching a decision for that branch.

Step 1: Import the Decision tree classifier from sklearn library.

Step 2: Then Split the training and test data

Step 3: Combine the train and test data and fit the values from class attributes.

Step 4: Using the some parameters for calculating and find the accuracy score, prediction value and f1 scores.

Step 5: Analyse the data set to summarise their main characteristics.

2. Random Forest Algorithm

- Random forest model is an ensemble of classification (or regression) trees.
- Random forests is a set of multiple decision trees.
- Deep decision trees may suffer from overfitting, but random forests prevents overfitting by creating trees on random subsets.
- Scikit-learn provides an extra variable with the model, which shows the relative importance or contribution of each feature in the prediction.
- Random forests creates decision trees on randomly selected data samples, gets prediction from each tree and selects the best solution by means of voting.

Random forest Introduce flexibility and converts High Variance → Low variance.

Step 1: Construct Bootsrapped dataset. Import the Random forest regression from sklearn library.

Step 2: Then Split the training and test data

Step 3: Combine the train and test data to fit the values from class attributes.

Step 4: Using the some parameters for calculating and find the accuracy score, prediction value and f1 scores.

Step 5: Analyse the data set to summarise their main characteristics.

3. Support Vector Machine (SVM)

- ✓ A support vector machine (**SVM**) is a supervised machine learning model that uses classification algorithms for two-group classification problems.
- ✓ **Accuracy** can be computed by comparing actual test set values and predicted values.
- ✓ To **evaluate** the **performance**, The first data sets is used to train the **SVM**, and the second learning data, which are not perfect (e.g. Noise) is taken for testing the **SVM** trained.

Step 1: Import libraries.

Step 2: Add datasets, insert the desired number of features and train the model.

Step 3: Predicting the output and printing the **accuracy** of the model.

Step 4: Finally finding the classifier for program.

Step 5: Analyse the data set to summarise their main characteristics.

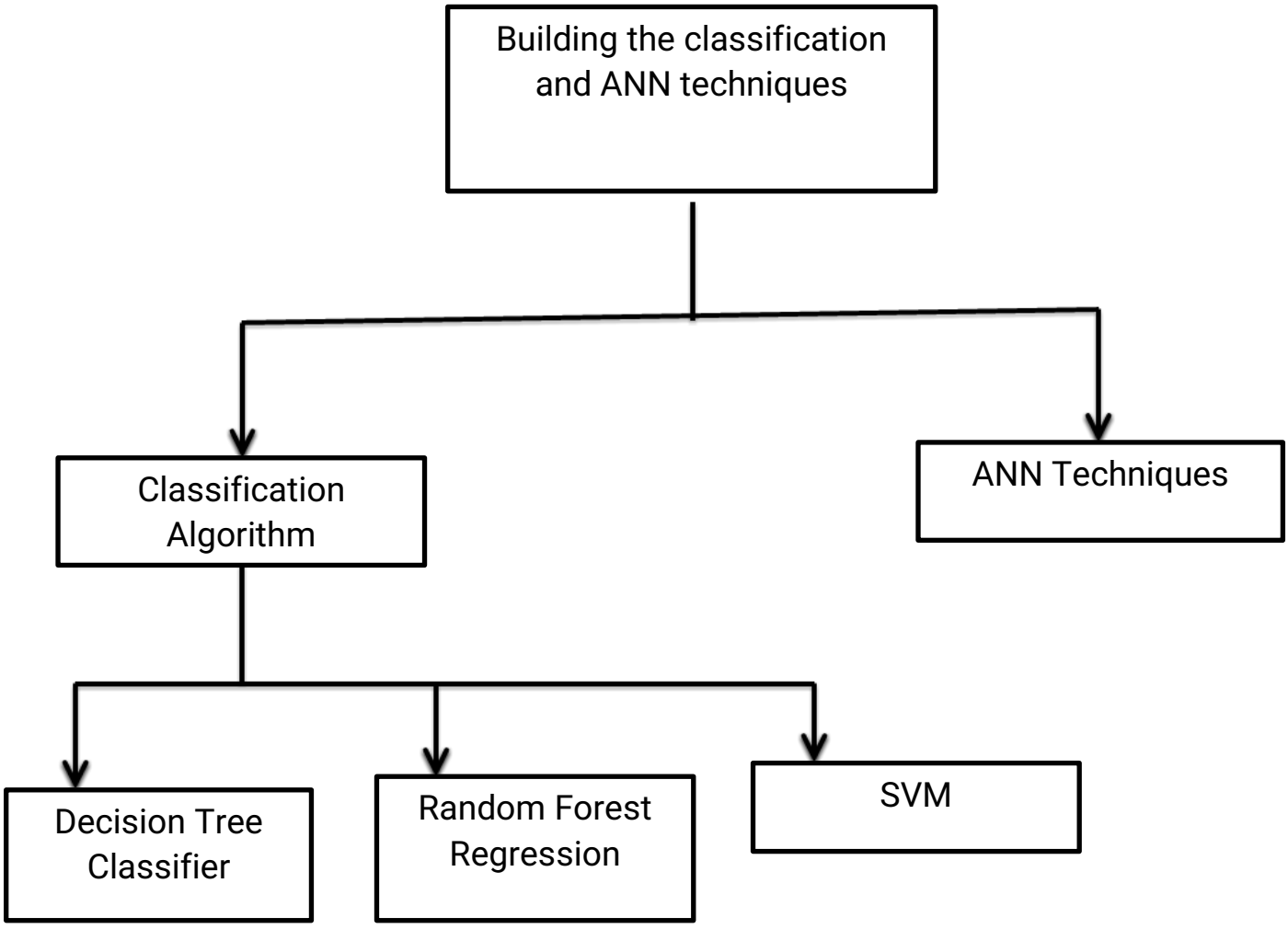


Figure 3.3 : Steps involved in classification and ANN techniques

3.1.5 Evaluation Measures

- The three main metrics used to evaluate a classification model are accuracy, precision, and recall.
- Accuracy is defined as the percentage of correct predictions for the test data.
- It can be calculated easily by dividing the number of correct predictions by the number of total predictions.
- Model evaluation aims to estimate the generalization accuracy of a model on future (unseen/out-of-sample) data.
- Methods for evaluating a model's performance are divided into 2 categories: namely, holdout and Cross-validation.
- Both methods use a test set (i.e data not seen by the model) to evaluate model performance.
- The accuracy of a machine learning classification algorithm is one way to measure how often the algorithm classifies a data point correctly. Accuracy is the number of correctly predicted data points out of all the data points [10].

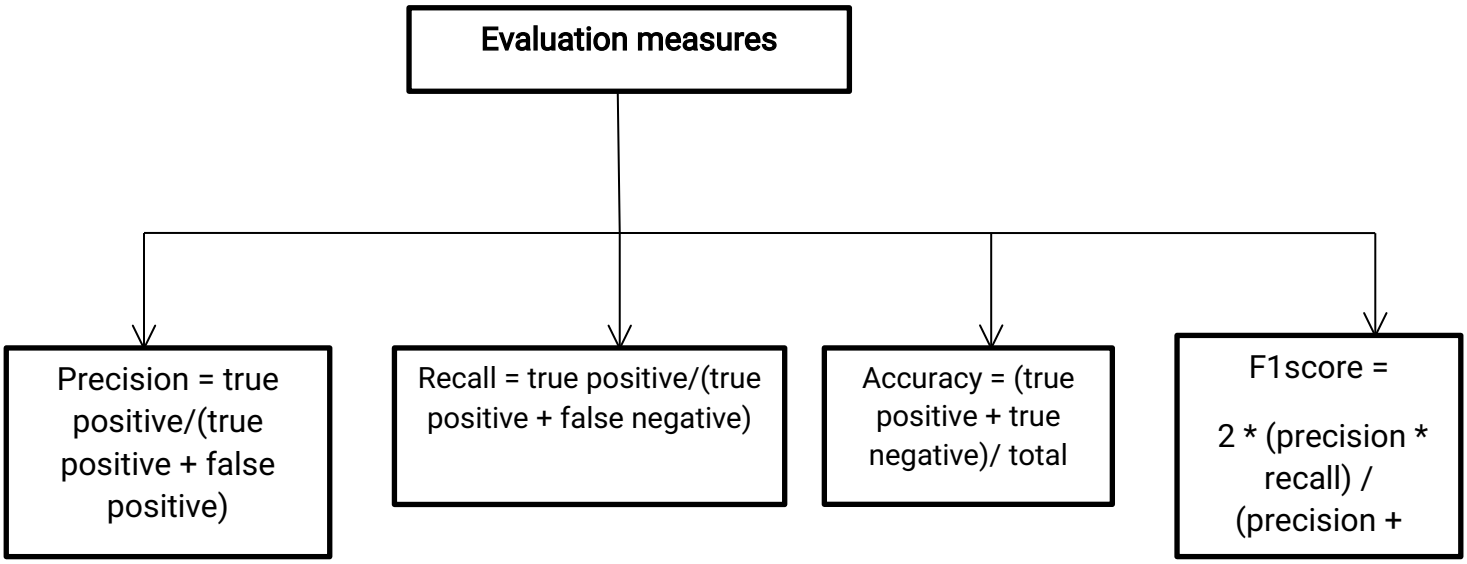


Figure 3.4 : Evaluation Measures

CHAPTER 4

IMPLEMENTATION

4.1. Data collection

The proposed system makes use of the dataset downloaded from this website: https://www.kaggle.com/c/1056lab-fraud-detection-in-credit-card. Dataset used is the transactions made by customer in a European bank in the year 2016-17.[1]

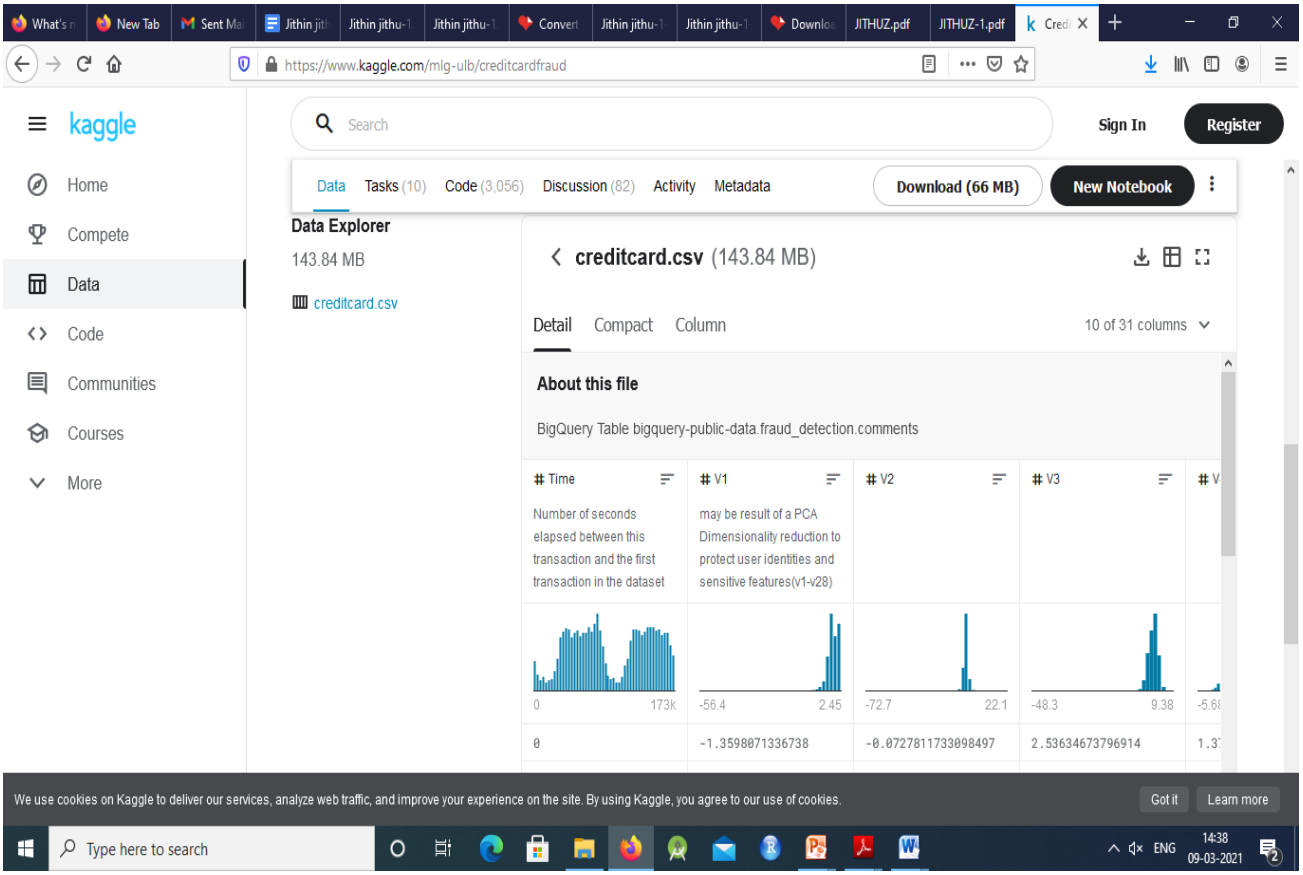


Figure 4.1 : creditcard.csv

Data set Description :

- It contains only numerical input variables which are the result of a PCA transformation.
- Features V1, V2, … V28 are the principal components obtained with PCA, the only features which have not been transformed with PCA are 'Time' and 'Amount'.
- Feature 'Class' is the response variable and it takes value 1 in case of fraud and 0 otherwise.

Attributes	Types
Time	Number
V1 to V28	Float
Amount	Float
Class	Number

Table 4.2 : Dataset Attributes and Types

1. **Time** : Number of seconds elapsed between this transaction and the first transaction in the dataset.
2. **v1 to v28** : Result of a PCA Dimensionality reduction to protect user identities and sensitive features(v1-v28)
3. **Amount** : Transaction Amount
4. **Class** : 1 for fraudulent transactions, 0 otherwise

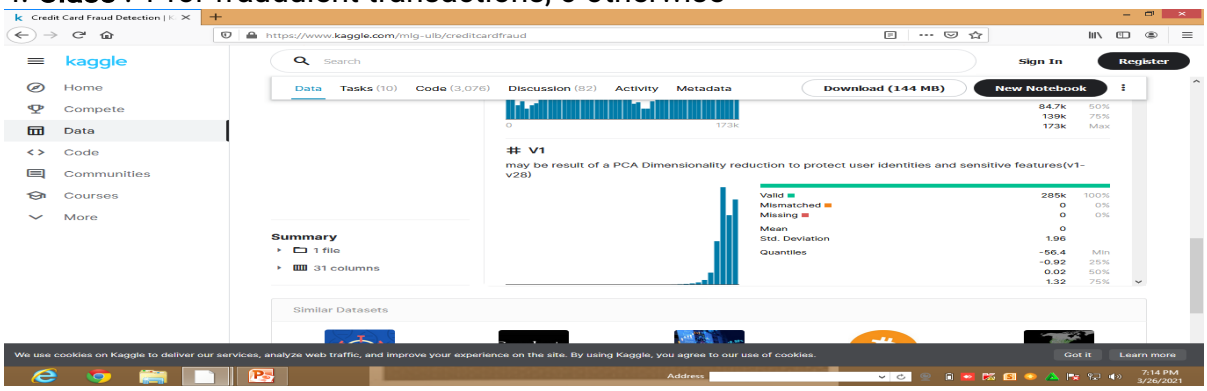


Figure 4.3 : v1 to v28 Description

4.2. Data pre-processing

For various reasons, the original database has a lot of dirty data, such as incorrect attribute values, duplicate records, null value, inconsistent values, various abbreviations, violations of referential integrity and so on.

In order to make better use of the data for data mining and decision support, it should be changed into high-quality data.

Therefore, a data pre-process procedure, which is Known as data cleaning,

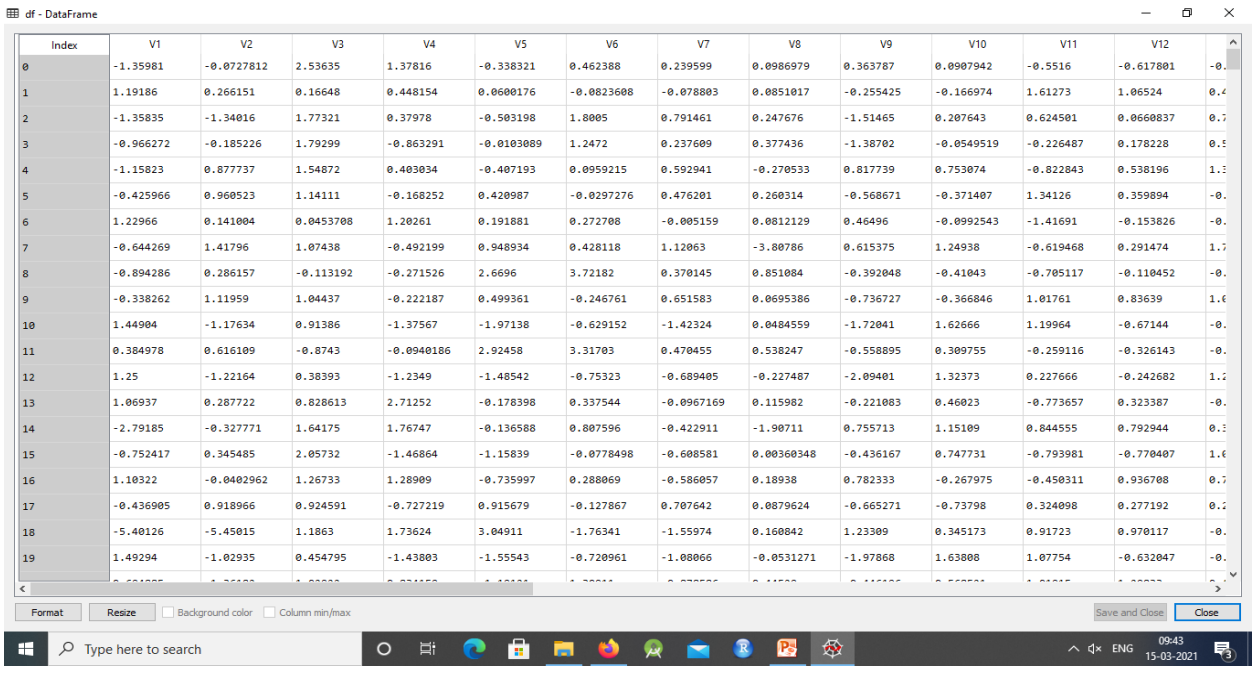
Will be used to clean up the dirty data before using the data.

Finding the missing values and duplicates data.

a. Importing the package

primary packages are going to be Pandas to work with data, NumPy to work with arrays, scikit-learn for data split, building and evaluating the classification models

df = pd.read_csv("C:\\Users\HP\\.spyder-py3\\project\\creditcard.csv")



Index	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	
0	-1.35981	-0.0727812	2.53635	1.37816	-0.338321	0.462388	0.239599	0.0986979	0.363787	0.0907942	-0.5516	-0.617801	-0.
1	1.19186	0.266151	0.16648	0.448154	0.0600176	-0.0823608	-0.078803	0.0851017	-0.255425	-0.166974	1.61273	1.06524	0.4
2	-1.35835	-1.34016	1.77321	0.37978	-0.503198	1.8005	0.791461	0.247676	-1.51465	0.207643	0.624501	0.0660837	0.7
3	-0.966272	-0.185226	1.79299	-0.863291	-0.0103089	1.2472	0.237609	0.377436	-1.38702	-0.0549519	-0.226487	0.178228	0.5
4	-1.15823	0.877737	1.54872	0.403034	-0.407193	0.0959215	0.592941	-0.270533	0.817739	0.753074	-0.822843	0.538196	1.3
5	-0.425966	0.960523	1.14111	-0.168252	0.420987	-0.0297276	0.476201	0.260314	-0.568671	-0.371407	1.34126	0.359894	-0.
6	1.22966	0.141004	0.0453708	1.20261	0.191881	0.272708	-0.005159	0.0812129	0.46496	-0.0992543	-1.41691	-0.153826	-0.
7	-0.644269	1.41796	1.07438	-0.492199	0.948934	0.428118	1.12063	-3.80786	0.615375	1.24938	-0.619468	0.291474	1.7
8	-0.894286	0.286157	-0.113192	-0.271526	2.6696	3.72182	0.370145	0.851084	-0.392048	-0.41043	-0.705117	-0.110452	-0.
9	-0.338262	1.11959	1.04437	-0.222187	0.499361	-0.246761	0.651583	0.0695386	-0.736727	-0.366846	1.01761	0.83639	1.6
10	1.44904	-1.17634	0.91386	-1.37567	-1.97138	-0.629152	-1.42324	0.0484559	-1.72041	1.62666	1.19964	-0.67144	-0.
11	0.384978	0.616109	-0.8743	-0.0940186	2.92458	3.31703	0.470455	0.538247	-0.558895	0.309755	-0.259116	-0.326143	-0.
12	1.25	-1.22164	0.38393	-1.2349	-1.48542	-0.75323	-0.689405	-0.227487	-2.09401	1.32373	0.227666	-0.242682	1.2
13	1.06937	0.287722	0.828613	2.71252	-0.178398	0.337544	-0.0967169	0.115982	-0.221083	0.46023	-0.773657	0.323387	-0.
14	-2.79185	-0.327771	1.64175	1.76747	-0.136588	0.807596	-0.422911	-1.90711	0.755713	1.15109	0.844555	0.792944	0.3
15	-0.752417	0.345485	2.05732	-1.46864	-1.15839	-0.0778498	-0.608581	0.00360348	-0.436167	0.747731	-0.793981	-0.770407	1.6
16	1.10322	-0.0402962	1.26733	1.28909	-0.735997	0.288069	-0.586057	0.18938	0.782333	-0.267975	-0.450311	0.936708	0.7
17	-0.436905	0.918966	0.924591	-0.727219	0.915679	-0.127867	0.707642	0.0879624	-0.665271	-0.73798	0.324098	0.277192	0.2
18	-5.40126	-5.45015	1.1863	1.73624	3.04911	-1.76341	-1.55974	0.160842	1.23309	0.345173	0.91723	0.970117	-0.
19	1.49294	-1.02935	0.454795	-1.43803	-1.55543	-0.720961	-1.08066	-0.0531271	-1.97868	1.63808	1.07754	-0.632047	-0.

Figure 4.4 : Importing view of data set

plot normal and fraud

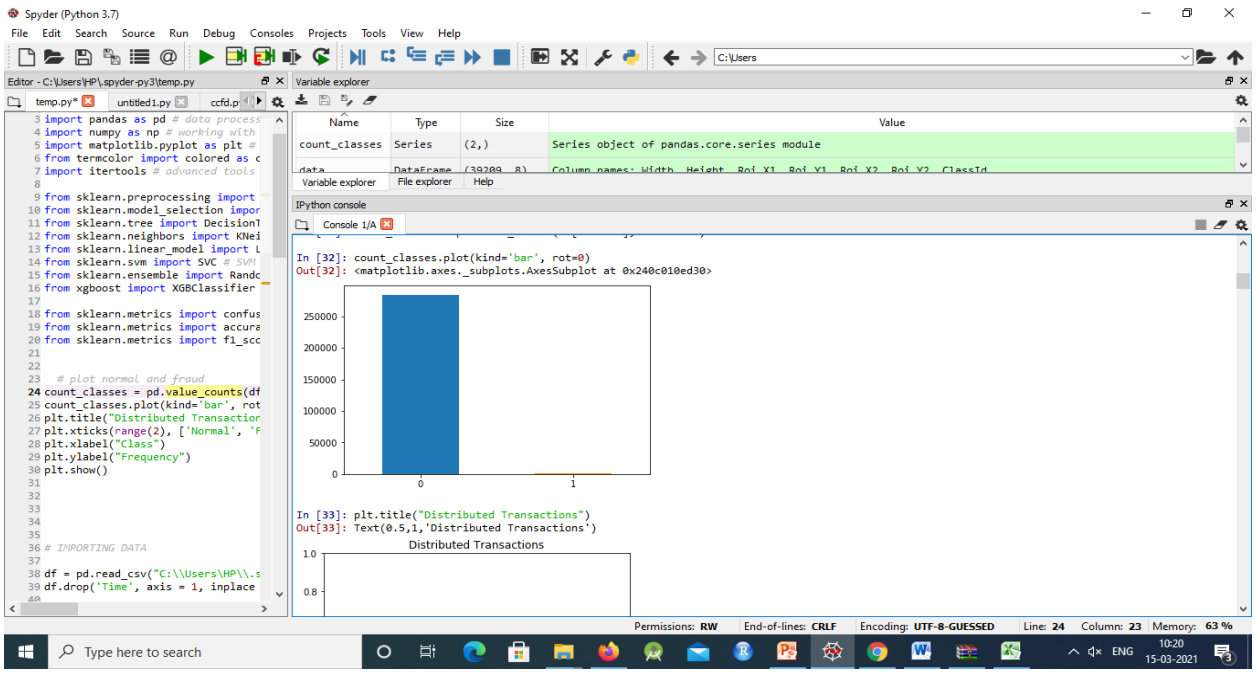


Figure 4.5 : plots view of fraud and non fraud

Pre-processing

print("missing values:", df.isnull().values.any())

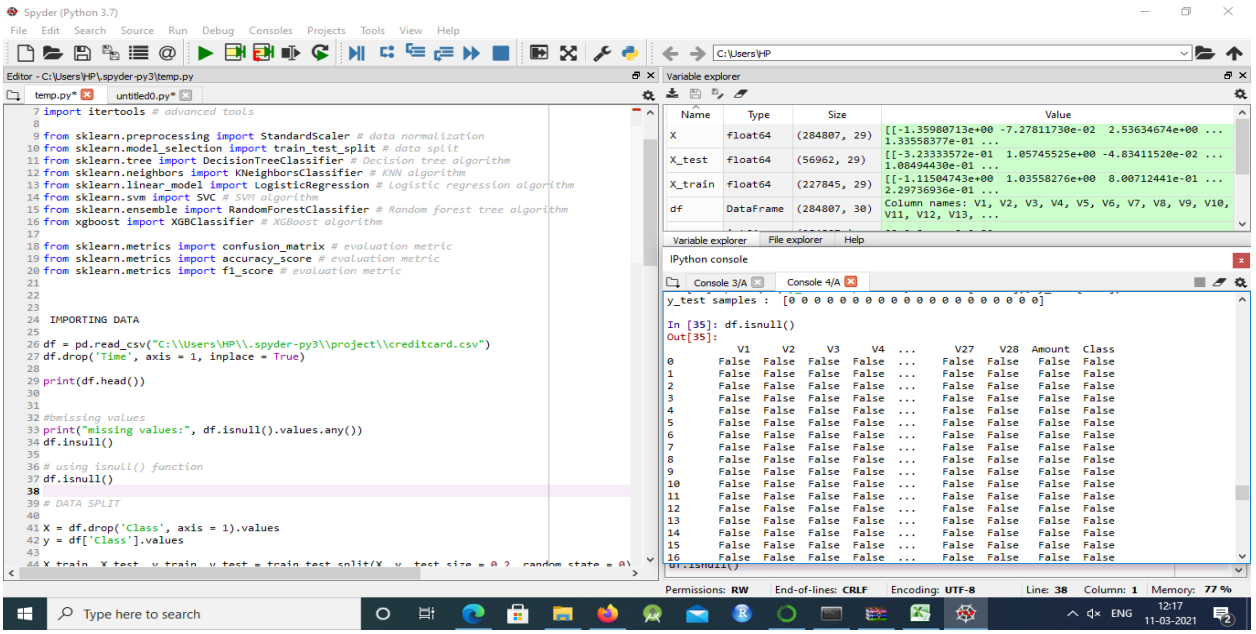


Figure 4.6 : Missing values views

df.dropna()

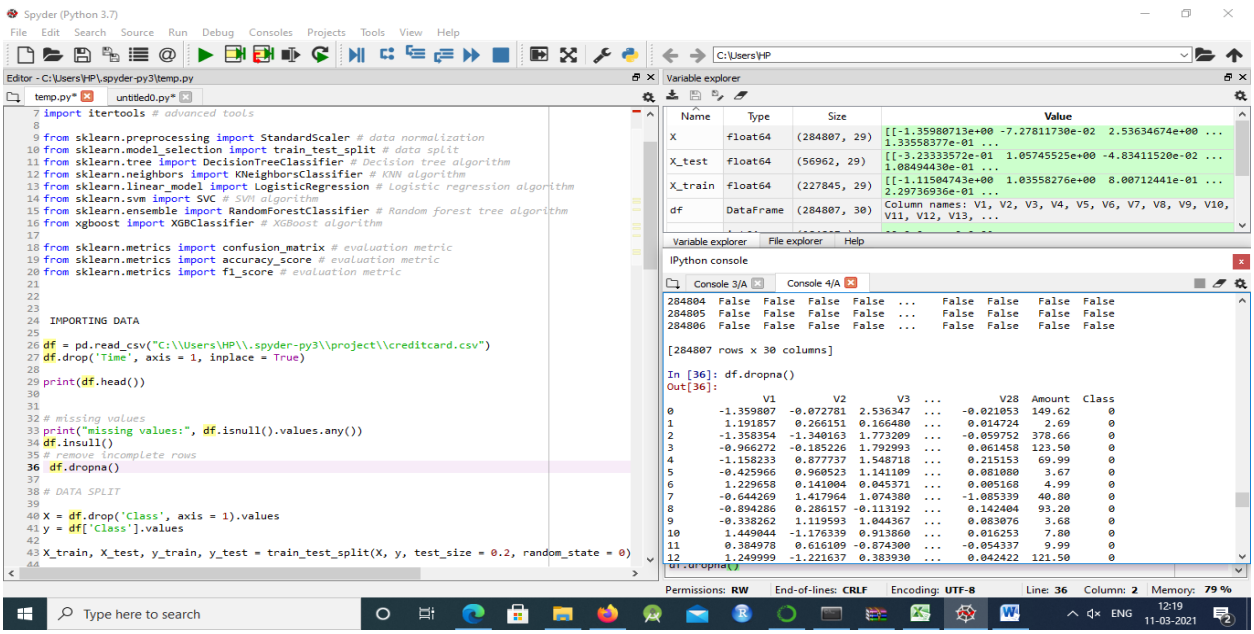


Figure 4.7 : Removing incomplete row

The outlier fraction is to be calculated.

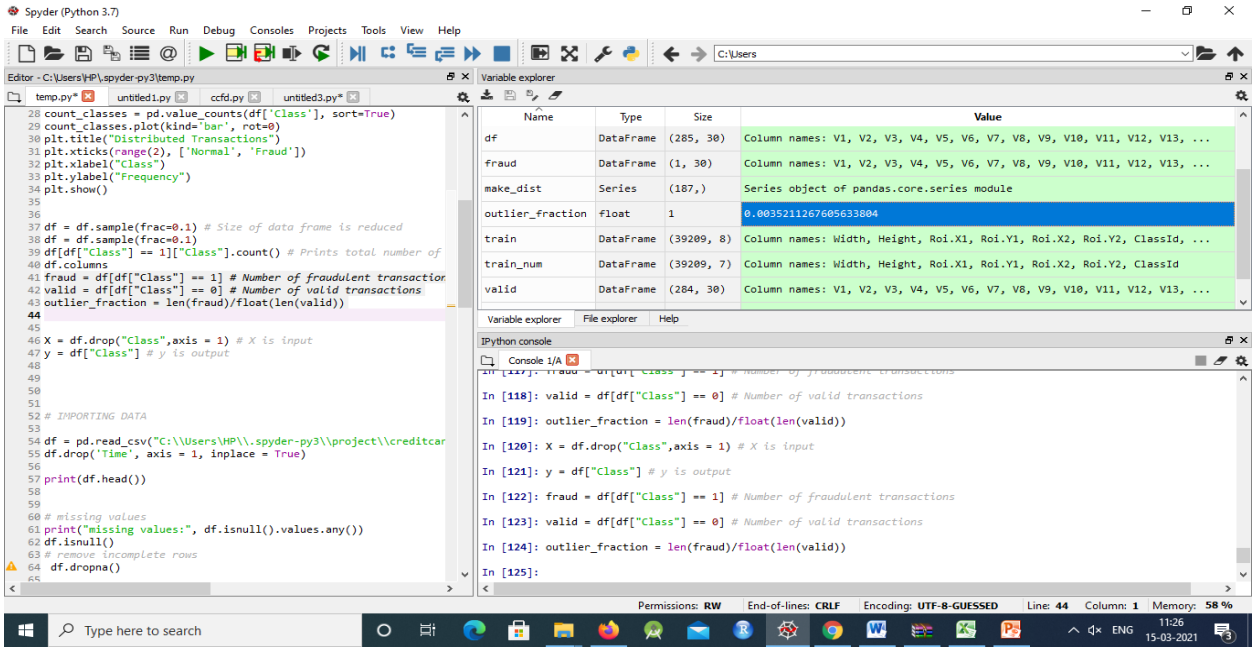


Figure 4.8 : View of Outlier fraction

Feature selection and Data spliting

- In this process, going to define the independent (X) and the dependent variables (Y).
- Using the defined variables, will split the data into a training set and testing set which is further used for modelling and evaluating.
- Split the data easily using the 'train_test_split' algorithm in python.

4.3. Build the classification and ANN technologies

- In order to build the different types of classification models namely Decision Tree, Logistic Regression.
- These are the most popular models used for solving classification problems. All these models can be built feasibly using the algorithms provided by the scikit-learn package.
- ANN It gives accuracy more than that of the machine learning algorithms. In this research work, data pre-processing, normalization and under-sampling carried out to overcome the problems faced by using an imbalanced dataset [8].

It works in four steps:

- Select random samples from a dataset.
- Construct a decision tree and random forest for each sample and get a prediction result, accuracy score and f1 score from each classification.
- Perform a vote for each predicted result, accuracy score and f1 scores.
- Select the prediction result with the most votes as the final prediction.

The model used is Decision tree classifier

The accuracy is 0.9991924440855307

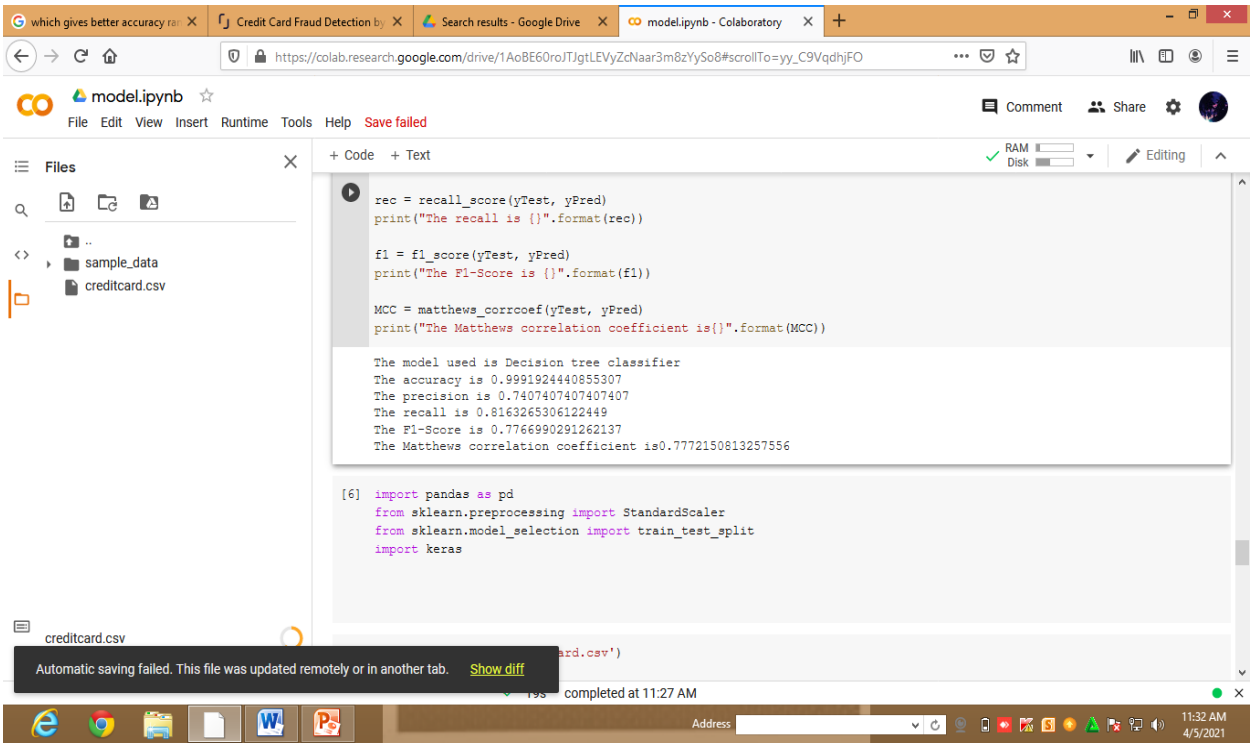


Figure 4.9 : Find accuracy using Decision tree classifier

The model used is Decision tree classifier

The accuracy is 0.9995786664794073

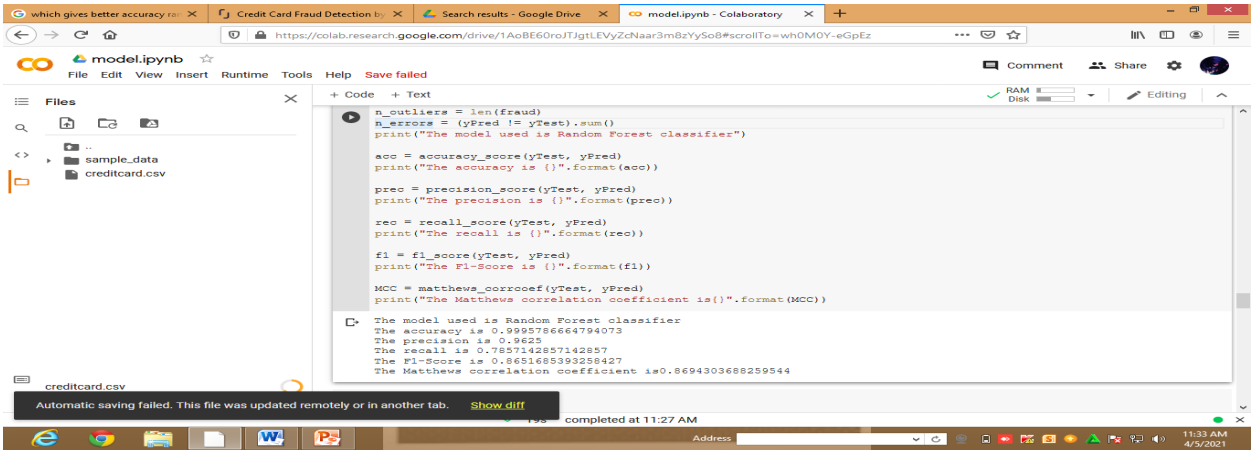


Figure 4.10 : Finding accuracy using Random forest classifier

The model used is SVM

The accuracy is 0.9998

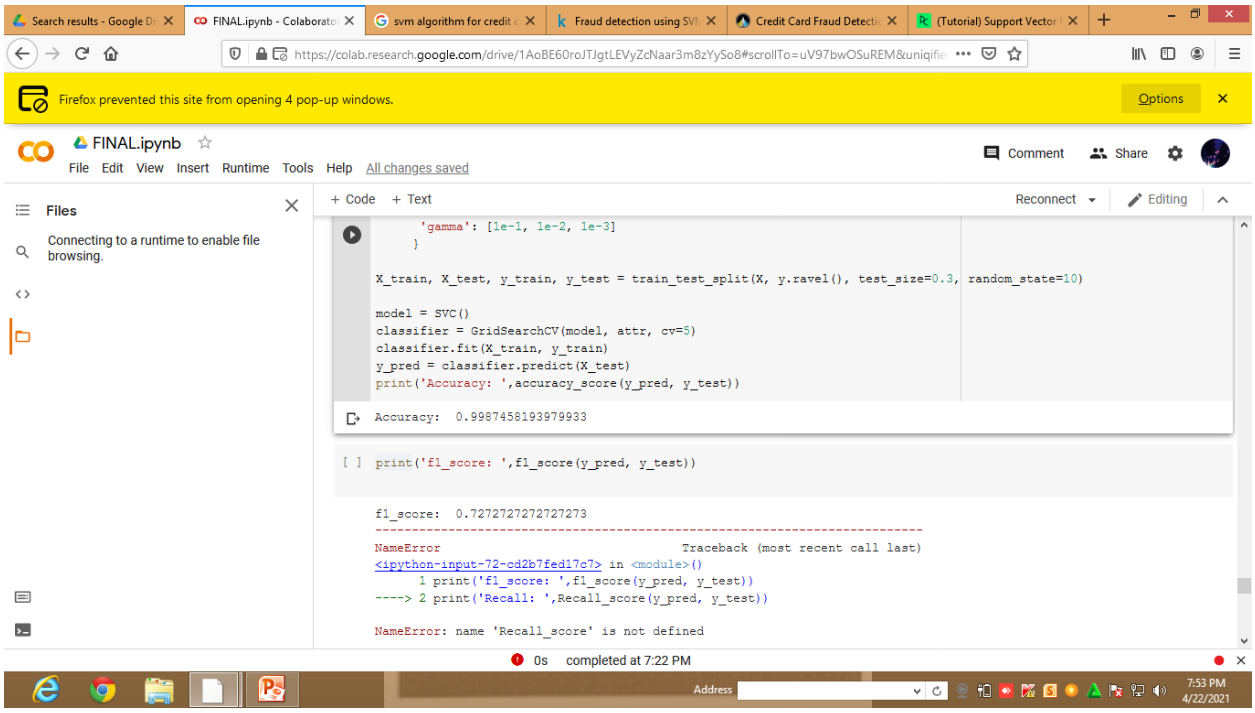


Figure 4.11 : Finding accuracy using SVM classifier

4.4. Evaluation measures

The end result is evaluated based on the confusion matrix and precision, recall and accuracy is calculated. It contains two classes: actual class and predicted class. The confusion metrics depends on these features:

True Positive: in which both the values positive that is 1.

True Negative: it is case where both values are negative that is 0.

False Positive: this is the case where true class is 0 and non-true class is 1.

False Negative: It is the case when actual class is 1 and non-true class is 0.

#Output

Epoch 1/2 5472/5472 [==============================] - 9s 2ms/step - loss: nan - accuracy: 0.9972 Epoch 2/2 5472/5472 [==============================] - 9s 2ms/step - loss: nan - accuracy**: 0.9973**

<tensorflow.python.keras.callbacks.History at 0x7f6197e95190>

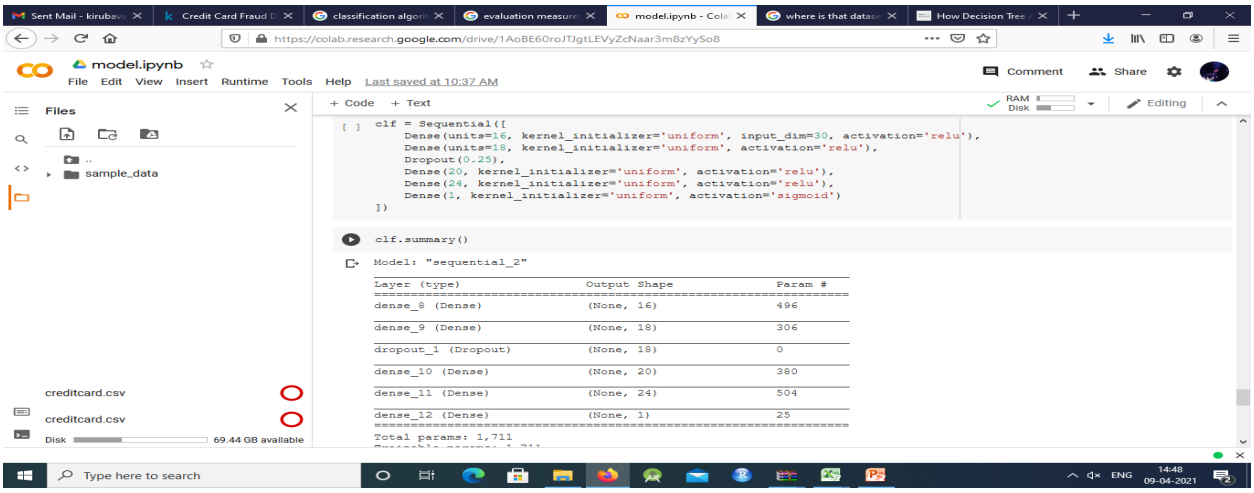


Figure 4.12 : Using ANN technique find the layers

CHAPTER 5

RESULTS AND DISCUSION

While comparing the **confusion matrix** of all the models, it can be seen that the **ANN model** has performed a very good job of classifying the fraud transactions from the non -fraud transactions followed by the Layers model. So we can conclude that the most appropriate model which can be used for our case is the ANN model and the model which can be neglected is the SVM model.

After a whole bunch of processes, we have successfully built different types of **classification models** starting from the **Decision tree model to the ANN model**. After that, we have evaluated each of the models using the evaluation metrics and chose which model is most suitable for the given case for giving better accuracy [15].

To build an ensemble, simply train multiple different ML models on the same data for the same task. At inference time, will apply all of the models to your input individually. If task is classification, can combine the results using a simple per class voting scheme or take the prediction with the highest confidence. For regression, just average out the results. Ensembles are an extensively field-tested and real-world-proven technique to **boost prediction accuracy.**

Method	Accuracy-score	F1-Score	Recall	Precision
Decision Tree Classifier	**0.991**	0.77669902912	0.8163265306	0.7407
Random Forest Classifier	**0.995**	0.86516853933	0.7857142857 1	0.9625
Support Vector Machine	**0.997**	0.72	1.0	0.571
ANN	**0.998**	0.917	0.889	0.946

Table 4.13 : Results of classifiers and ANN

This supervised learning dataset is suitable for history database for credit card fraud detection. Supervised learning such as **multilayer perceptron** in neural network that uses the prediction algorithm to identify whether new transactions are legal or illegal. When a credit card used, the neural network based on the fraud detection system checks for the pattern used by the fraudster and corroborates the pattern in question or checks for attributes that have been determined as illegal if the pattern matches with genuine transaction behavior, then the transaction is considered legitimate. **Neural networks** have many sub-techniques. So, if they pick-up this which is suitable for credit card fraud detection, the performance of the method will decline.

- **Accuracy** can be computed by comparing actual test set values and predicted values.
- **Random forests** consist of multiple single **trees** each based on a **random** sample of the training data. They are typically more accurate **than decision trees**.
- **Random Forest** is a great algorithm, for both classification and regression problems, to produce a predictive model.
- In this dataset using the decision tree classifier algorithm gives the accuracy is **0.9991924440855307 .** And random forest regression algorithm gives the accuracy is **0.9995786664794073.**
- Take the confusion matrix of the ANN Layer model as an example. The first row is for transactions whose actual fraud value in the test set is 0. can calculate, the fraud value of 56861 of them is 0.
- And out of these **56861 non-fraud transactions**, the classifier correctly predicted 56854 of them as 0 and 7 of them as 1. It means, for 56854 non-fraud transactions, the actual churn value was 0in the test set, and the classifier also correctly predicted those as 0. can say that our model has classified the non-fraud transactions pretty well.
- Its default hyperparameters already return great results and the system is great at avoiding overfitting.
- Moreover, it is a pretty good indicator of the importance it assigns to features.

6. CONCLUSION

- Using the classification algorithms and ANN techniques can find fraud transaction and **original transaction.**
- The detection can be done automatically so it will detect the fraud while transaction.
- Historical data is used so we can get the pattern of user's original transaction and fraud transaction.
- Classification algorithm SVM it gives the accuracy is 0**.997 %**, comparing the ANN it gives the accuracy is **0.998%** best accuracy.
- The accuracy of the detection is more accurate by using the **ANN techniques**
- Accuracy is **better than the existing system.**
- By using an artificial neural network which gives **accuracy approximately equal to 99% is best** suited for credit card fraud detection. It gives the accuracy more than that of supervised learning algorithms [15].

7. SCOPE FOR FUTURE ENHANCEMENT

- Further requirements and improvements can easily be done.
- There are large numbers of dataset present and many other classification algorithms are present.
- So the future work will be based on moves that the user provides to choose the time of movements.
- And also other classification algorithms that can be applied on the data set and comparison are made between the accuracy rate produced by the different algorithms.
- It allows software applications to become accurate in predicting outcomes. Moreover, machine learning focuses on the development of computer programs.
- The primary aim is to allow the computers learn automatically without human intervention [15].

8. BIBLIOGRAPHY

[1] C. L. Blake and C. J. Merz. (2017) UCI Repository of Machine Learning Databases. Univ. California, Dept. Inform. Comput. Sci.,Irvine, CA. [Online]. Available: http://www.ics.uci.edu/~mlearn/ML-Repository.html

[2] C. J. C. Burges, "A tutorial on support vector machines for pattern recog-nition," Data Mining Knowledge Discovery, vol. 3, no. 2, 2018.

[3] Delve: Data for Evaluating Learning in Valid Experiments [Online]. Available: http://www.cs.utoronto.ca/~delve

[4] (2018, Aug.) Improvements to Platt's SMO Algorithm for SVM Classifier Design, Accepted for Publication in Neural Computation. Control Division, Dept. of Mechanical and Production Engineering,National University of Singapore, Singapore. [Online]. Available :http://guppy.mpe.nus.edu.sg/~mpessk

[5] J. C. Platt, Advances in Kernel Methods: Support Vector Machines,38 B. Schölkopf, C. Burges, and A. Smola, Eds. Cambridge, MA:MIT Press, Dec. 2017. Fast training of support vector machines using sequential minimal optimization.

[6] National University of Singapore, Singapore. [Online]. KNN Available:http://guppy.mpe.nus.edu.sg/mpessk

[7] Feature Selection for Regression | Kaggle. Available online: https://www.kaggle.com/ohmets/feature-selection-for-regression/data (accessed on 10 June 2019).

[8] Maes, Sam, et al. "Credit card fraud detection using Bayesian and neural networks." Proceedings of the 1st international naiso congresson neuro fuzzy technologies. 2019.

[9] David J. Montana, "Neural Network WeightSelection Using Genetic Algorithms." BoltBeranek and Newman Inc. July2018.

[10] Philip K. Chan, Wei Fan, Andreas L.Prodromidis, and Salvatore J, "DistributedData Mining in Credit Card FraudDetection" IEEE December 2019.

[11] Sushmito Ghosh and Douglas L. Reilly,"Credit Card Fraud Detection with a Neural-Network." Nestor, Inc. IEEE (2018).

[12] Rajesh Parekh,, Jihoon Yang,and VasantHonavar, "Constructive Neural-NetworkLearning Algorithmsfor PatternClassification" IEEE 2010.

[13] Mubeena Syeda,Yan-Qing and Yi-Pan,"Parellel Granular Network For Credit CardFraud Detection". IEEE 2017.

[14] Erik Bothelius, "Fraud detection in theInternal Account System for PaymentService Providers." May 8, 2015.

[15] D.WHITLEY,“Genetic Algorithm AndNeural Network.”2017.'

[16] https://spd.group/machine-learning/credit-card-fraud-detection/

[17] https://www.zendesk.com/blog/machine-learning-and-deep-learning/

[18]https://www.rfwireless-world.com/Terminology/Advantages-and-Disadvantages-of-Deep-Learning.html