Appendix A

MSC APPLIED AI AND DATA SCIENCE

SEPTEMBER 2022

NSE UDOUKPONG

THE APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNIQUES TO DETERMINE CUSTOMER SATISFACTION AND CHURN RATES, A CASE STUDY OF THE TELECOMMUNICATION SECTOR.

SOLENT UNIVERSITY

FACULTY OF BUSINESS LAW AND DIGITAL TECHNOLOGIES.

SOLENT UNIVERSITY

FACULTY OF BUSINESS LAW AND DIGITAL TECHNOLOGIES

MSc Applied AI and Data Science

Academic Year 2021-2022

NSE UDOUKPONG

THE APPLICATION OF ARTIFICIAL INTELLIGENCE TECHNIQUES TO DETERMINE CUSTOMER SATISFACTION AND CHURN RATES, A CASE STUDY OF THE TELECOMMUNICATION SECTOR.

Supervisor: Dr. Drishty Sobnath

September 2022

This report is submitted in partial fulfillment of the requirements of Solent University for the degree of MSc in Artificial Intelligence and Data Science.

ACKNOWLEDGMENT

I would like to say a big thank you to God Almighty for the strength He has given to me to complete this thesis.

To my mother, thank you for the support, prayer, and words of encouragement.

To my sisters, and sister-in-law, Thank you for your advice, little jokes, and prayer.

To my brothers, Bro Mboutidem, Brother Aniekan, and My love (Bobo). I want to thank you for your financial support, for believing in me, and your endless advice. I am grateful.

To my family in Christ, Daddy Morgan, Mummy Juliet, Uncle Emmanuel, AuntyBukky, Memark, Seun. Thank you for your support. I am grateful.

To my friends, Kelvin, Onyinyechi, Peter, Pelumi, Gabriel, Jibola. Thank you.

Finally, to my supervisor, Dr. Drishty, I am grateful for your constant push, leadership, motivation, hard work and attention to my problems. Thank you for seeing the best in me. I am grateful.

ABSTRACT

Today, the availability of similar IT products and services to customers has made the telecom market highly competitive which has led to an issue of customer churn due to their inability of not being able to detect and predict the customers that are going to remain This research aim is to provide organizations with several Artificial Intelligence techniques on how to determine customer satisfaction and deter them from leaving. It has been proven that it is less expensive to retain existing customers than to engage with new customers because customers normally are faced with competitors throwing adverts, emails, and exclusive discounts that would encourage them to patronize their services/products. This research has also explored the strategies to retain customers by surveying the causes of customer dissatisfaction using reviews coming from customers which are considered one of the major approaches for customer satisfaction. The use of data mining/prediction tools such as Weka, and PowerBI were used for data visualization/ prediction, as well as using machine learning and deep learning models to determine various evaluation scores and how accurately they performed. For the machine learning model, the Random Forest had the highest performance score with 80% while the Roberta model had an accuracy score of 92%. To sum it up, an artefact was created to predict reviews so as to provide excellent guidance to organizations on how to improve customer satisfaction and deter churn rates based on the context of their words.

Table of Contents

Tables of Figures

ACRONYMNS AND ABBREVIATIONS

Area Under the Curve	AUC
Alite Bert	ALBERT
Artificial Intelligence	AI
Bi-directional Encoder Representations FromTransformer	BERT
Bi-directional long short-term Memory	BI-LSTM
Customer Relationship Management	CRM
Decision Tree	DT
Deep learning Model	DLM
Exploratory Data Analysis	EDA
Extreme Gradient Boosted Machine	XGBOOST
Gated Recurrent Unit	GRU
Gradient Boosted Machine	GSM
Information Technology	IT
K-Nearest Neighbour	KNN
Linear Regression	LR
Logistics Regression	LR
Machine Learning	ML
Natural Language Processing	NLP
Natural Language Processing Toolkit	NLTK
Principal Component Analysis	PCA
Random Forest	RF
Recurrent Neural Network	RNN
Robustly Optimised Bert Pretraining Approach	ROBERTA
Support Vector Machine	SVM
Valence Aware Dictionary for sEntiment Reasoning	VADER
Waikato Environmental for Knowledge Analysis	WEKA

CHAPTER ONE

INTRODUCTION

1.1 Background Information

The experience a customer encounters with a company's service or product indicates if a customer is going to churn or remain. The telecommunication sector is a competitive industry where companies need to strive ahead to be above their fellow competitors and the perfect way to do that is through customer satisfaction. This implies, creating unique value for customers that would prevent them from leaving. Customer satisfaction is a measure that helps in determining how elated a customer is with the services, products, and potency of the organisation. Kortler(2002), defined customer satisfaction because of the client's contentment or disapproval of an organisation's products or services which can either lead to a customer's choice to either remain or churn.

According to the Management study guide, UK, Customer Relationship Management(CRM) started in the early 1970s when organisations saw the need to be more empathetic towards customers and not just be concerned about the profit they are gaining or products that they are presenting to customers. This rise led to an era of customer satisfaction which allowed businesses to be more intentional in looking for ways they can strategically satisfy customers. The advent of CRM has brought up diverse changes coupled with technology, great software implementation, and business intelligence tactics that can help companies in gaining information that would increase profit as well as improve customer satisfaction.

Customers are regularly faced with competitors throwing adverts, emails, and exclusive discounts that would encourage them to patronise their services/products. Research has proven that to attract new customers cost more than customers who have been with companies for long. (Holmound and Kock,1996). According to Reichheld and Sasser (1990), there is a 25-85% increase in a company's profit if a customer's loyalty to an organisation is 5%. This means that the profit made by an organization grows consistently if customers remain.

In the book "A Churn Prediction Model Using Random Forest" Ullan et al(2019), ascertained that customers are open to many choices of similar services or products, it is difficult to retain new customers. This means that the use of advertisement and the cost of retaining new customers cannot be overemphasised while the reverse is the case with existing customers. They went further to explain that if an organisation can identify the major cause of existing customers being dissatisfied, there would be a little amount spent as compared to retaining new customers.

The impact that comes with customers' dissatisfaction is that they leave. Verbeke (2011), opined that when a customer leaves, the prestige of the company is affected. However, when customers' reasons for being dissatisfied are targeted, trust, loyalty, and profit increases because they will be willing to advertise the company to their colleagues, family, and friends. Reichheld and Sasser (1990), in their book "Zero definitions: Quality comes to services", estimated that 60% of profit in a company can be linked to loyal customers due to their recommendations.

Engel et al (1996), also gave an in-depth explanation, they explained that with the way the market is brimming with competition, discerning, and adjusting to customers' choices is no longer a choice but important to sustain organisational growth as well as to survive in a market that is led with similar and better services or product that can easily appeal to customers. They emphasised the need for companies to put in action by prioritising customers first to avoid a high increase in churn rate. If organisations can ask themselves questions like, how can we know when customers are going to leave? Why are they leaving to organizations that offer similar services? it would be easy for an organization to execute customers' needs and desert itself from competitors.

While there are several interventions from researchers to use AI technologies like the use of applying machine learning models to predict and analyse how customers can be satisfied and to reduce churn rate, there is still a gap of specifically creating a technique that can help companies identify core reasons why customers leave which needs to be filled. Hence, this study seeks to delve into and narrow the knowledge gap on how to satisfy customers and identify those that are leaving the companies due to their dissatisfaction. From the implementation stage of working with powerful software tools, alongside applying machine and deep learning models, to building a demo user-interface, down to adopting sentiment analysis, and wrapping up with the recommendation for future references. This research will outline the steps that was undertaken in employing artificial intelligence techniques to determine customers reasons for leaving and how companies can recognize when a customer is dissatisfied with their product/services.

1.2 Statement of the problem

Today, the availability of similar Information technology (IT) products and services to customers has made the telecom market highly competitive which has led to an issue of customer churn (Customers switching from one IT product or Internet service provider to another). Nguyen (2011), in her book, opined that the fact that customers can convert to any IT product or service provider without contentions has made it easy for customers to leave and fostered customers to be in pursuit of a better and affordable service. Millions of companies have experienced the “customer-losing problem” due to their inability of not being able to detect and predict the customers that are going to remain. Companies that encounter these setbacks in detecting these customers always try to attain new customers which in most cases are highly pricey and time-consuming.

For companies, losing customers to their competitors is a big failure which is a problem. Therefore, effectively applying artificial intelligence techniques in determining customer churn rate would aid companies to diagnose customers that have the maximal chance of leaving as well as deter them from churning. Hence, this study seeks to examine the application of artificial intelligence techniques in determining customer churn rate.

1.3. Justification of the Study

Churn management is one of the key features in customer relationship management because of the attention it gets from companies due to its unique components on how customers can be retained. Balling and Poel (2012), in their book “Customer event history for Churn Prediction: How long is long enough”, explained that the reason churn management has gotten the attention of companies is because of the value it brings to companies.

This study explores the strategies to retain customers. More precisely, the study surveys the causes of customer dissatisfaction which is considered one of the major approaches capable of increasing customer churn rate.

This study has both economical and factual significance that will contribute to existing literature and body of knowledge, help companies understand several Artificial Intelligence measures asides from the traditional method of applying machine learning models to determine customer churn rate, and retain customers that will eventually yield profit to the organization. It will help to understand how organizations can achieve different yet similar customers with the utilization of key strategic tools. Lastly, the findings from this study are expected to serve as a guide to organizations for financial prowess and dexterity because of the success that could be achieved after with the utilization of this method.

1.4. Aims and Objectives

The aim of this study is to build a demo user interface that can serve as a guide to companies on how to satisfy and determine customer churn.

The research will do the following:

1. This project will build a system which would allow companies input data for prediction and provide guidance to organizations on how to improve customer satisfaction based on the output.
2. To Apply several Artificial Intelligence techniques to determine customer churn rate.
3. Comparative analysis with the application of different machine learning and deep learning models to determine the best performance accuracy
4. Demonstrate essential elements of customers, their actions, and how businesses are to respond when faced with similar scenarios.

1.5 Delimitation

1. This study was limited to Network providers in the UK
2. 120 customers through survey, using different mobile networks were used in the study.
3. Not all customers completed the customer questionnaire survey.
4. Data collection for the research transpired in May 2022 from a website called Data world.
5. The survey collected from customers is for the sole purpose of tackling the research questions and is restricted to public interpretation.

1.6 Definition of Key Terms

Customer Churn: This can be defined as the process by which customers leave an organization by switching to another organization that offers related products/services. (Mustafa, 2021).

Customer Satisfaction: It is a response of pleasure, and fulfilment got from an organization over a product or service (Martins, 2012).

Artificial Intelligence: This can be seen as the “ability to make machine intelligence” (Nils Wilson, 2010).

Sentiment Analysis: This is the application of an algorithm to extract text data into categories for the purpose of gaining insight. (Pang and lee, 2008).

Machine learning: It is a computing algorithm, whose constant aim is to mimic and learn human intelligence (Murphy, 2015).

CHAPTER TWO

LITERATURE REVIEW

The main aim of this chapter is to discuss the empirical review of the research topic, the theoretical framework that is underpinned, and the conceptual framework. This chapter will be written based on this knowledge. There are several works and journals that have been done on customer satisfaction and churn prediction. It gives an overview of various sources of works and research journals that have contributed to the body of knowledge on customer satisfaction and churn prediction.

2.1 Concept of Customer Churn

Customer relationship management originated from relationship marketing. Researchers like Kortler (1997) and Berry (1983) analysed the best way to create a satisfying relationship with customers, stakeholders, and employees and opined that marketing birthed the idea of customer relationship management. Customer satisfaction is the focus of any organization. This has led to research and studies from various scholars to analyse the concept of customer satisfaction. In the book, "Your Customer's Perception of Quality", Baboo Kureeman and Robert Fantina (2011), explained customer satisfaction as a method by which companies listen to customers' "perception of quality". They explained that a customer can only be regarded as being satisfied if the company's products/ services are impacting positively on the customer's impression of quality. They opined that organizations make huge mistakes when they try to convince customers to accept what they perceive as quality.

Cook sees customer satisfaction to be when the organization has been able to capture the economical and emotions of the customers. To her, economical aspects involve effort, affordability, and time while the emotional aspects involve their perception of the brand, the product/services, and the employees. In other words, customers connect on an emotional level with brands which in most cases annul the sense of thoughts and actions when making decisions. Oliver (1999) further explained that customer satisfaction is the result of loyalty from the two parties. That is, a customer can be satisfied if they are pleased with the company's services/ product consistency which transcends to the customer's loyalty, on the other hand, companies can only satisfy customers if they remain loyal to the services and products they render.

In their book, Goldenstein and Sheldon (2009), believe that customer satisfaction comes from companies receiving unbiased, constructive feedback from surveys that involve attributes that are important to the companies. They opined that companies may feel they are satisfying customers by surveying and improving on products or services that might not be important to the customers, leading to dissatisfaction and customer churn. Companies can only do better if they are open to constructive criticisms from customers by discovering and measuring features that are important to their customers and not dwelling on biased surveys.

Grigoroudis et al (2004), believe that customer satisfaction comes from the mental analysis of the customer that is linked to their behavioural analysis. They explained the customer's mind as a "black box" controlled by an expectation of what they want. According to him, this expectation can either be turned into a satisfactory or dissatisfactory outcome which is linked back to the perception of what they want.

In addition to the customer satisfaction analysis, Churchill et al, (1982), Oliver (1980), Vavra (1997); and Oliver (1996) all have a common definition and approach to customer satisfaction, which is that customers are satisfied based on experience from product and services of a given company which meets their expectations.

2.2. Customer Churn

In the telecommunication sector, churn is when a company loses a customer to another service operator (Houiekh and Elhaj 2020). Churn can be linked to predicting when customers leave their current company. Hence, companies seeking sustainability and growth must make an active plan in creating an inevitable working strategy that would retain or acquire new customers. Pooja et al (2020) defined customer churn as a process that occurs when a customer cease doing business with a company. Zhang (2020) defined churn as "a loss of a customer in favour of a competitor". According to her, with the use of key strategies and the application of a machine learning model, predicting customers' churn can help in identifying the reasons for their disservice.

The major objective of telecom companies is to stay relevant in a competitive market and maximise profits. However, this can only be achieved if the customer and the company have a better relationship. As explained earlier, customer dissatisfaction is the major reason customer churn. In the telecommunication sector, the market is saturated, making it challenging to retain customers mainly because of the leverage for customers to switch to different networks or products. For companies to effectively identify customers who are going to leave, Ammar et al (2017) explained that identifying these customers at the right time would impact the company's profit and growth. Additionally, Nguyen (2011), opined that several companies have lost their customers due to their inability to detect customers who are going to remain. In essence, companies must understand the behaviour of customers to retain existing customers and prevent them from leaving.

Furthermore, Hemlata et al (2020) explained that though every industry experience problem with customer churn, the telecom market has the highest percentage of 90% compared to other industries, this is because the sector is highly competitive which has made companies proactive in determining customer churn by analysing the behaviour of the customer and putting in the effort that would retain them. Retaining them is vital because it has been proven by researchers like Holmlund and Kock (1996) that it cost more for a company to gain new customers than to keep existing ones.

Mattison (2001) went further to analyse the 2 distinct types of churns, which are external and internal churn. According to him, external churn can be classified under the voluntary and non-voluntary variables. The non-voluntary falls under circumstances that are beyond the customer's control e.g. Death, criminal cases, etc. while voluntary churn could be a factor like traveling out of the country. On the other hand, the internal churn involves changing service providers. Hadden et al (2006) described the voluntary churners as "deliberate churners" because this is based on their behavioural attitude that most organizations are trying to resolve.

McCann et al (1990); and Brandusoui et al (2013) explained in their respective books that the key challenge for businesses is to find a perfect way of identifying similar traits that future churners possess and to deploy a retention process that can efficiently retain them. To achieve this effectively, the application of machine learning models that can accurately predict customers who are going to churn has been adopted by companies. With churn prediction, companies do not feel the burden of traditionally finding ways to detect customers that are likely to leave. The application of information technology has eased off the stress that companies face and rather, bring in more success, profit, and growth in a highly competitive market.

2.2.2 Factors Affecting Customers Churn

Understanding the reason customers churn helps customer relationship management in businesses to improve their productivity effectively, improve the marketing strategies of the company as well as help in knowing the relevant ways a customer can churn based on similar patterns that have been portrayed.

Tianyuan et al (2022) discussed how expenses and price of the data, monthly fees, and network services efficiency are major reasons why a customer can churn in the telecommunication sector. They went further to explain that these are significant factors that can be used to measure and analyse customer behaviour. They opined that if companies can structure these key factors, they are guaranteed a loyal customer. They went further to explain that when customer dissatisfaction level is on the increase, there is a high tendency for the customer to churn and when companies consistently stick to customer satisfaction, there will be a low percentage of customers likely to churn.

2.2.2.1 Price Factor

Customers always love products that have lower prices. Zhao et al (2021) opined that although a company’s product/ services can be high. However, it should meet the expectations of customers. That is, companies should create a long-lasting relationship by delivering values despite the price to maximize growth and profit from the customers. It is essential to note that the price of a company’s product /services can negatively determine a customer’s willingness to leave.

2.2.2.2 Product Factor

The product can also be a determinant of a customer to churn This can be through the following reasons:

I. The needs/demand of the customers are not fully considered during the production of these products and not consider the effect the product will have on the market (Zhao et al 2021).

II. The value of the product has reduced from its initial state. (Zhao et al (2021) explained that most customers have a “first-time impression” behavioural attitude towards a product from a company which determines if they will remain loyal. They explained that if an existing customer feels that the value of a product has been reduced, the customer might leave. The customer's behavioural traits can only be impressed by the value of the company’s product /services if the company remains consistent.

2.2.2.3 Business Factor

In the telecommunication sector, what most customers desire is to be able to purchase all the needed products and services they want from one service provider. For example, if customers find it difficult to get a bundle option that would save cost, and create value, they might look elsewhere which can increase churn. Gerpott et al (2015) stated that customers would want the company to have all that they need. They stated that lack of this plan can increase customer churn rate and can affect bringing back lost customers. Wangehein et al (2017) also discussed that when companies do not have diverse service contracts, there will be little customer demand that would be met which might affect customer churn.

2.2.2.4Service Factors

Wangeheim et al(2017) stated that the more customer is satisfied with the service of a company, the lesser they churn. If an excellent relationship is maintained between a customer and an enterprise, there will be a low level of churn rate. In essence, if there is a failure of quality service, dissatisfied customers would prefer to patronize the company's rivals that would use better services as a strategy to attract customers. It is vital to note that investing in relationships with a customer is key because the lower the customer's dissatisfaction, the lower the churn rate would be.

2.3.Customer Segmentation

Dolnicar et al(2018) defined customer segmentation as categorizing customers with similar characteristics. According to them, these groups of people help companies able to target and niche out specific marketing strategies that would get their attention. They went further to explain that customer segmentation is mostly used in the telecommunication marketing sector. Grouping these customers can be either through similar behavioural patterns. E.g. demographics etc.

Wedel and Kamkura(2000) believe that companies should involve more in market segmentation because it involves various factors such as socio-economic, cultural, the customer's personality, and geographic variables which would make it easier for businesses to determine similar attitudes from customers and know how to target their needs. They further explained that implementing customer segmentation through using of marketing segmentation software can help companies to understand their customers as well as know the best way to access them based on what they like, their number of purchases, and their interests.

Kortler and Keller (2006) discussed that an effective segmentation should be evaluated based on the company's objectives, resources, marketing strategies, growth potential, responses, and sizes, towards retaining the customers, which should be in cooperation with what the company can handle. They went further to explain that implementing segmentation should be well thought of and handled as critical because with the company's using their research on marketing campaigns and advertisement, they need to be sure that they are focused on identifying the right set of people without making mistakes.

2.3.1 Reasons behind customers segmentation

Bruce et al (2007) in their book stated that market segmentation is a great strategy that helps guide firms on how to strategically market products/ services to customers. The market is saturated with a unique group of customers that have different characters, behaviour, lifestyles, and needs, that might need to be approached based on their preferences. Segmenting customers would give ideas to organizations on the best way they can satisfy customers coming from different works of life.

Additionally, segmentation helps organizations in improving their productivity level as well as helps suppliers in producing products based on the results that have been derived from the analysis they have carried out on customers. Bruce et al (2007) presented a case study of how one of the biggest grocery stores in Turkey segmented its customers. Here, customers were segmented based on three different approaches. This included the **Value-based customer**; these customers were regarded as loyalists. They were defined as the loyal segment because of the number of purchases and years they have bought from the grocery stores, The **Behavioural-based Approach;** was based on the purchase lifestyle of customers that had similar patterns repeatedly over the past year, and The **Activity-Level based Approach**; was based on the customer's level of purchase. From their illustration, there were 3 categories of customers; the active, passive, and normal. This helped in understanding the level where a group of individuals fell under and how to apply the "migration analysis" in increasing their activities. These different segmentation approaches gave the groceries stores an insight into how they should improve their productivity.

Also, it is pertinent to note that with organizations employing the segmentation of customers, they are going to achieve extraordinary results based on the analysis and surveys that the organization has carried out. Also, it helps businesses to strategically know the customer's group that they should be intentional about and how they can employ marketing measures on the different group.

2.4. Impact of Customer satisfaction on Organizations

As previously stated, customer retention is derived from the sustained satisfaction that the customer receives due to the organization's services and products. Moshan et al (2011) emphasised in their book that one of the impacts of customer satisfaction is customer loyalty. They stated that satisfaction makes a customer remain loyal and indicates that the customer is pleased with what the organization is doing to retain its customer.

In their various research work, Ryals (2015); Kolis and Jirinova(2013), and Payne and Frow (2006) had a similar view on the impact of customer satisfaction in organizations. They explained that customer satisfaction helps in bringing better performance to organizations which helps the company grow. According to their research, this performance is a result of their creativity, value creation, and strategic development to sustain customers which eventually transforms into financial growth.

In addition, Rizwan et al (2022) explained in their work that customer satisfaction helps in building a great reputation for the company in a competitive market such as the telecommunication sector. Just like Takeuchi (1983) explained, building a great reputation is an important feature that involves rendering great service quality. From his explanation, service quality cannot be quantified because when a customer has a subjective insight about a company's product/services as being, in their favour, it does not only create a good reputation but also fosters customer loyalty, exposes the company to new customers and earn the trust of the customers which is a significant impact.

2.5. Customer to Company Relationship

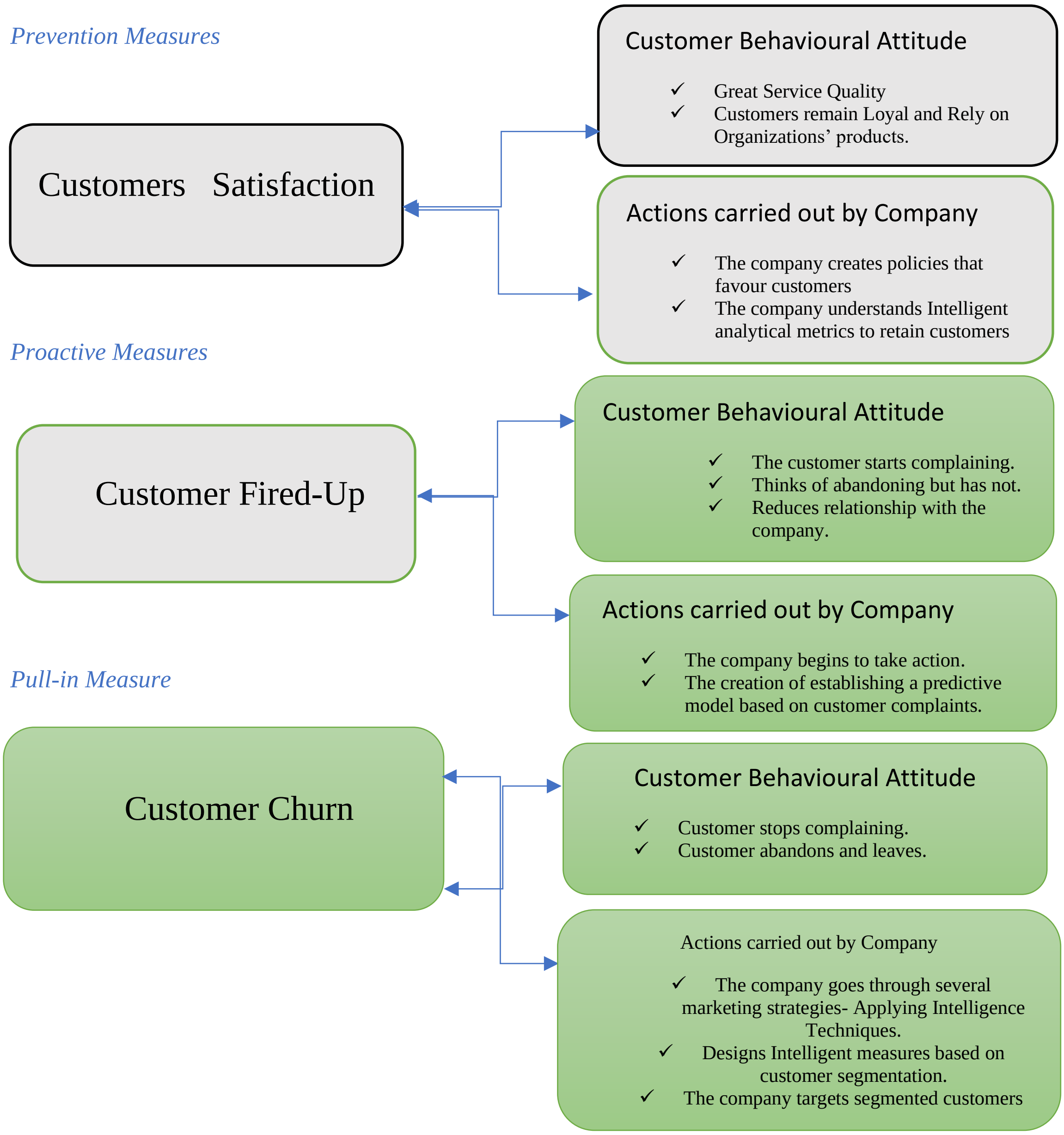


Figure 1 The above Conceptual Framework created by me illustrates the stages in a customer's life cycle. The figure gives three illustrative analyses of a customer's profile I called the 3Ps. The Preventive Measure. The Proactive Measure, and the Powell-in Measures. The figure also shows the relationship the customer has with the company. The diagram displays how a customer can switch from one stage to the other if there are no adequate measures in place to resolve their problems.

2.6. The Application of Artificial Intelligence in Business for customer satisfaction.

In a competitive market such as the telecommunication sector, Artificial Intelligence techniques such as the application of machine learning (ML), deep learning (DL), Natural language Processing (NLP), and Data Visualisation tools, are great instruments for solving business problems. Garcia et al (2017), in their book, demonstrated that to derive insight, retain customers and understand the reasons why customers are not satisfied with a company's products/ service, the application of artificial intelligence technique plays a vital role in assisting the company to understand the mechanisms of how a customer can remain loyal. They went further to explain that using artificial intelligence techniques helps to analyse customers' reasons as to why they want to abandon a company's product/service. This encompasses the exploratory process, the analysis process from different angles, the utilization of different tools, and the prediction process.

Thomas (2019), in his book, explained that during Juniper Research, companies employed different methods such as natural language, and recommendation methods to satisfy customers' needs. As a result, an estimated 3.25 billion digital assistants have been developed and identified to satisfy customers' needs. Baier et al., (2018); Koeheler, (2016); Leimeister and Bitter, (2019), gave a similar discussion in their research, that companies integrate technologies into their daily operation to achieve productivity and observed that when they utilize these intelligent techniques, they create trust from the customers.

In the book, "How 50 successful companies used Artificial Intelligence to solve problems", Bernard Marr and Matt Ward (2019) demonstrated that businesses can utilize AI to create better and more intelligent products and services, improve their business processes, and change the techniques on how they interact with their customers. With regards to customers, they stated that with the use of AI, businesses are given the chance to understand their customers, predict the best products or services that they might want, and provide a better way of achieving their results. On the aspect of the product and services, the application of AI techniques helps in creating intelligent products/services for customers.

2.7. Constraint Affiliated with the Application of Artificial Intelligence

In a study conducted by Venkatesh et al (2011), they stated that as much as it is the desire for companies to satisfy their customers, within the context of Artificial Intelligence techniques, companies need to use the appropriate measures in protecting and securing their customer's information as well as still have the ability and expertise to sustain the nature of their products/ services by giving out the best to their customers. They went further to explain that customers have the right to their information, and it is important that when applying artificial intelligence techniques ethical principles should be followed, and customers should still have the power over decisions regarding their information. Applying artificial Intelligence techniques always requires data to discover hidden patterns and insights. However, Bhatt, (2014) stated that it comes with the risks of companies being accused of misusing or abusing the data. He described that unauthorized use of individual information can be regarded as abuse that can affect the organization.

Marr (2019) urges companies to approach Artificial intelligence strategically by having a concept of how to get data. According to him, data is an asset that is used to solve problems. However, it must be done in conformity with the legal and privacy of customers to avoid encountering issues that would affect the policies of the company.

2.8. Leveraging Churn Analysis

There are several ways that researchers have used to leverage churn analysis. These includes the use of machine learning, data visualization, and deep learning by incorporating the customer's data to draw insight and predict the risk to which a customer can churn. Despite the great outcome of these approaches, there have been some limitations in this research work.

Nhi et al (2021) explained that only using structured data such as customers' demographics and their history can be expanded. They suggested the use of unstructured data and believe that companies using customer interaction with customer services can help in revealing more insight. According to them, even though the call lengths and frequencies to which they call have been used by other researchers to determine if a customer was going to leave or remain, the content of the call from the customer has not been utilized. To them, applying Natural language processing (NLP), and incorporating it with machine learning can apprehend a comprehensive reason a customer wants to leave a company and can also be used for prediction. However, to have access to customers' calls, there needs to be permission taken from the customers which is a protection of their rights. This limitation led to inferior quality of text transcript which made the advanced model that was implemented ineffective.

Zhou et al (2018) also gave an in-depth analysis by explaining that using customer activities and online reviews using big data text analytics has proven successful. They explained that this can be used to capture the emotions of customers based on their reviews. Hybrid models that captured the customers' text were used. However, with the use of these hybrid models, it was unclear how the researchers used the text information derived from the customers to accurately determine if a customer was going to leave and how they can be retained.

2.9. Related work on Churn Prediction in Mobile Telecommunication

In the telecommunication sector, customer churn is a concern faced by companies, which has made companies seek other means such as the use of artificial intelligence to predict the reason a customer might leave. Several approaches have been used to predict churn in the telecommunication sector. Most of the approaches utilized are machine learning, deep learning, and data mining to evaluate and analyse the reasons why a customer wants to leave.

Gavril et al (2016) made use of an advanced method of data mining to predict why a customer might leave a company, with the use of over 3000 call datasets gotten from customers' call transactions. The researchers applied the Principal Component Analysis dimension (PCA) and made use of three machine learning algorithms namely; the Neural Networks, Support Vector Machine, and Bayes Network. The Area Under Curve (AUC) was used to measure how the model performed. For the Bayes model, they had 99.10%, Neural Network was 99.5%, and SVM was 99.70%.

Idris et al (2012), with the use of two datasets, made use of AdaBoost to predict customer churn in telecommunications. For the cell2cell dataset, their accuracy performance was 63% while the orange telecom dataset had 89%.

He et al (2009), used the Neural Network Algorithm to predict why customers leave, making use of one of Chinese largest telecom datasets, their overall accuracy performance was 91.1%.

Abdelrahim et al (2019), in their work, "Customer Churn Prediction in telecom using machine learning in big data platform", used four algorithms to predict customers that were going to churn. The models that were used included Decision Tree, Random Forest, Extreme Gradient Boosted Machine (xgboost), and Gradient Boosted Machine (GSM) The models were tested using the Syriatel dataset and the social network analytics dataset. The result they got was XGBOOST: 93.3%, GSM: 90.89%, Random Forest:87%, and Decision Tree: 83%. While having GSM outperforms other algorithms.

Hiziroglu et al's (2014) research were on determining customer churn with the use of segmentation and data mining. With the use of Logistic regression and Decision Trees, their Area under the ROC Curve (AUC) performances were 82.6% and 87.3% respectively. After segmenting the data, the AUC score was

2.4% improvement for logistics regression and slightly the same for decision trees. From their research work, segmenting the model to determine future churners, and its accuracy performance can help companies target marketing campaigns to retain customers that were going to churn.

The work by Mohammad et al (2021), was on churn prediction, to compare different machine learning models. They did a comparative analysis of churn prediction making use of the K-Nearest Neighbour (KNN and Decision Tree algorithm. From the research, DT (Decision Tree) had the highest accuracy score with 92% while K-Nearest Neighbour was 86%.

Li (2020), researched on customer churn using mobile communication technology, and was structured using data mining technology, he believed that businesses could derive effective decisions making on customers through data mining, which will determine if a customer is going to leave or remain. He used the decision tree model and the logistic regression. From his research, the decision had an overall accuracy score of 83% while the logistic regression was 79%.

Adnan et al's (2017) research on rule-based customer churn prediction was implemented using the rough-set theory. According to him, several researchers have made use of different mechanisms. However, their result was not effective. In his research, they used four different mechanisms, namely: The Exhaustive, The Genetic, The Covering, and the LEM2 Algorithms. The Genetic Algorithm had the highest efficiency in the research with an overall accuracy performance of 98%, Exhaustive: 92%, The Covering: 87%, and LEM2 with 93%.

Cheng et al's (2019) study were on credit cards to show the importance of retaining customers that would sustain the market in the banking sector. The research produced a system that could help in detecting customers that would churn and provide an indicator to companies by stating problems that could make them lose their customers. From the system that was built, the overall test confidence score was 96%.

The research by Pooja et al (2020) was to develop models that can help to predict if a customer was going to churn. Using fintech companies as their case studies, they made use of the neural network, SVN, RF, and Linear Regression. The neural network had the highest accuracy score of 91%, SVM: 84%, RF: 79%, and LR (Linear Regression): 67%. A summary of relevant work is presented in a Table format below.

Table1. Summary of related work

Author	Industry	Variables	Research Methodology
Garvil et al (2016)	Mobile Communications	Calls transactions, Time, Age, Text	Neural Network, PCA, Support Vector Machine
Idris et al (2012)	Telecoms	Age, demographics, Bills, and Churn	Ada Boost
He et al (2009)	Telecommunications	Bills, Tenure, Payment information	Neural Networks
Abdulrahim et al (2019)	Telecommunications	Demographics, frequency, bills, profit	Decision Tree Random forest XGBOOST Gradient Boosted Machine (GSM)
Hiziroglu et al (2014)	Customer Relationship Management	Monthly bills, partners, Churn, call intake	Decision Tree, Logistics Regression
Mohanmmed et al (2021)	Mobile Telecommunication Technology	Calls- intake Monthly bills Tenure, Frequency	K-Nearest Neighbour Decision Tree
Quang Li (2021)	Mobile Telecommunication Technology	Call intake Monthly bills Tenure Frequency	Logistic Regression Decision Tree
Adnan et al (2017)	Telecommunication	Age, demographics, profit, Tenure, and Payment records	Rough-set theory Genetic Algorithm LEM2 Algorithm Exhaustive Algorithm Covering Algorithm
Li Cheng et al (2019)	Credit Cards	Payment bills, Monthly bills, Tenure, frequency, and Age	A model (96%)
Pooja et al(2020)	Bank Sectors	Age, Payment, Frequency	Neural Network

Figure 2: Table displaying Summary of Related work.

2.10. SUMMARY

Customer satisfaction is a necessity in a field such as the telecommunication sector. Customer churn is derived because of when a client is not satisfied with the products/services of a company. This chapter gave an extensive literature review on customer satisfaction. It x-rays the impacts of customer satisfaction on organizations and how practicing artificial intelligence has helped solve problems.

CHAPTER THREE

METHODOLOGY

3.1 INTRODUCTION

This section will describe the technique and the process used to achieve this project's objectives. Kothari (2004) defined research as a methodology scientific and systematic examination of any topic in the quest for the truth. This is usually carried out with the use of research, comparison, observations, and experimentation. This project used the mixed method approach to understand the topic in a greater in-depth.

Research handbook (2010) defined research as a specific way an individual investigates, writes, and answers questions that pertain to the study, and does a piece of work. Shorten and Smith (2017), pointed out that when researchers use quantitative and qualitative data, it is known to be a mixed method approach. Creswell (2019), in his book, explained that the mixed method approach helps researchers use several ideas and unravel questions that exist during research studies. Guerra-Santin et al (2016), went further to explain that the mixed method is always used when researchers want to pragmatically abduct the social and scientific practice of the quantitative and qualitative methods. According to them, it can simply be referred to as combining two distinct ways of answering research questions. These approaches can be implemented during various processes during the research. This can include, data visualization and analysis, or data collection.

According to Keyson et al (2016), the qualitative data and quantitative approach can be used in three distinct ways:

1) Through building one data type on the other, which is known as connectivity;

2) Comparing or relating results coming from the same date types which can be referred to as merging;
3) Explaining the result derived from the data can be referred to as embedding.

The interpretation of mixed methods is mostly utilized when researchers want to successfully curb the limitation coming from the two distinct approaches. Utilizing the mixed method helps give in-depth, broader ideas that always yields impressive results.

The purpose of this research is to apply artificial intelligence techniques in satisfying customers and determining e churn rate. This will be by building an algorithm that will fixate on term retention techniques, utilizing technical tools to segment and visualize customer's behavioural attitudes and demonstrate the customer-company relationship step actions that need to be taken when faced with similar scenarios. This chapter analyses and discusses the data collection, data features and descriptions, and its intended outcomes and dissemination.

3.2 Data Collection

Two datasets were used to carry out this research work. The secondary source was gotten from a website known as Dataworld(https://data.world/secun/fmcgdataanalysis/workspace/file?filename=WA_Fn-UseC_-Telco-Customer-Churn.csv). The dataset used is one of the biggest network providers in the United Kingdom known as Telco. This was used in carrying out data visualization as well as Machine learning Prediction. The data consisted of 21 columns and 7,043 rows Figure 42 and one label was identified as the independent feature known as Churn, which was to determine customers that were going to leave or remain. Each row in the data was filled with customer information.

The primary data had 80% of its responders from WhatsApp groups. The data was restricted to how satisfied responders were with their service users. It consisted of 11 columns and 120 rows Figure 43 and one feature label was used to understand users that were not satisfied with their network providers.

The reason for this research was to use several Artificial Intelligence techniques to identify customer satisfaction in the telecommunication sector and to help businesses understand how to identify ways to retain their customers. To achieve this, businesses and customers' reviews were a unit of analysis. The materials that were used in this research were to help in providing reliable results to bring out the best conclusion at the end of this research. It was also to find and review studies that have researched and documented Customers Satisfaction and Churn rate between 2005-2022 as I believe will help in gaining more insight and helps in improving the objectives of this study. However, vital theories and research works have been used outside this time bracket. Additionally, the geographical location of this research work was based in Southampton on the account that respondents from the survey were residents in this location, and on the notion that the outcome of this work would have a positive impact and relevance to environmental researchers who might feel the need to do further research.

3.3. Data Description

The primary data was developed on 21 June 2022, 12:56 pm, and was shared from 21 to 29 June 2022. 120 respondents completed the survey, and three were dropped because the respondents did not answer all the important questions that were needed for the research. The interpolate and replace function was not used because the missing values were reviews that needed the opinion of responders that filled the survey. However, 117 respondents were valid to work on the research because they responded to the important questions. Respondents were asked 10 questions based on their network providers to understand factors that have led to them retaining or leaving their current network. The secondary data was downloaded on 26 June 2022, 3:30 pm, and was already in a CSV format. The data consisted of a complete dataset because there were no missing values.

3.3.1 Experimental Data

As previously explained above, in section 3.2(Data Collection), 2 datasets were used in this research. The secondary data was obtained from the data world while the primary data was through surveys. Both data estimate if a customer was willing to remain or leave a network provider service and the information that was received was based on the relationship, they have had with the network providers.

The secondary data feature includes the following: Gender, Senior Citizen, Partner, Dependent, Tenure, PhonseService, Multipleline Service, OnlineSecurity, Device Protection, Tech Support, Streaming Movies, Contract, Paperless billing, Payment Method, Monthly Charges, Total Charges, and Churn.

While the primary data feature includes the following: Churn, Network Service, Country of Origin, Gender, Age range, how long have you been using your current Network, have you tried any other Network asides from the one you are currently on, what are the three bad things about your current Network, what are the three good things about your current Network.

3.3.1.1 SECONDARY DATASET TABLE

Variable	Description	Data Type	Type of Variable
Churn	Based on Responder’s Response	Target	Nominal
Gender	Either Male or Female	Independent	Nominal
Partner	If responders have a partner or not	Independent	Nominal

Dependent	Is the responder a dependant or has dependants	Independent	Nominal
Tenure	How long has dependant been with Network Provider	Independent	Ordinal
Phone Service	If responders have other phone services	Independent	Nominal
Multiple Line Service	Does the respondent have another network service	Independent	Nominal
Internet Service	What type of service did responder opt for	Independent	Nominal
Online Security	Based on Yes/No	Independent	Nominal
Online Backup	Either yes/No	Independent	Nominal
Device Protection	Either yes/No	Independent	Nominal
Tech Support	Either yes/No	Independent	Nominal
Streaming Movies	Either yes/No	Independent	Nominal
Contract	Agreement	Independent	Ordinal
Paperless Billing	Method of Charges	Independent	Nominal
Payment Method	Method of Charges	Independent	Categorical
Monthly Charges	Based on Charges	Independent	Nominal
Total Charges	Based on Charges	Independent	Nominal

3.3.1.2PRIMARY DATASETS

Variable	Description	Data Types	Type of Variable
Churn	Based on Yes/ No	Target	Nominal
Network Service	Network providers	Independent	Categorical
Country of Origin	Responder Country	Independent	Categorical
Gender	Based on Male or Female	Independent	Nominal
Age range	Based on responder's Age	Independent	Interval
How long have you been using your current Network	Based on years	Independent	Continuous
Have you tried any other Network asides from the one you are currently on	Based on Yes/No	Independent	Nominal
What are the 3 bad things about your current Network	Based on comments	Independent	Categorical
What are the 3 good things about your current Network	Based on comments	Independent	Categorical
Would you recommend your current Network to family, friends, and colleague?	Based on yes/No	Independent	Nominal

The primary dataset was divided into two, using the excel tool. This was solely for the purpose of extracting the reviews gotten from the responders during the survey.

3network	Uk	Female	30-45	24	No	Cheap bun	They can b	Yes	Yes
Lebara	Nigeria	Female	25-30	12	No	Availability	Costly,limi	Yes	No
Lycamobil	Nigeria	Female	23- 25	7	No	Fast netwo	Issues with	Yes	No
Vodafone	Nigeria	Female	30-45	24	No	Good cove	None	Yes	No
EE	Italy	Female	25-30	24	No	Reliable, c	Network f	No	Yes
EE	Nigeria	Female	30-45	0-6	No	Pay as you	Expensive,	No	Yes
2	Italy	Male	45 and abc	24	Yes	Reliability,	Na	Yes	Yes
Vodafone	UK	Female	45 and abc	24	Yes	1, The curr	None, real	Yes	No
Vodafone	Nigerian	Male	30-45	24	No	Fairly wide	Poor cove	No	Yes
Virgin	Nigeria	Female	30-45	24	Yes	Flexible, gc	A bit exper	Yes	Yes
Smarty	Nigeria	Female	30-45	24	Yes	1)Good se	I would lov	Yes	No
BT	Nigeria	Male	30-45	24	Yes	Cost savin	Recent pri	Yes	No
BT	UK	Female	30-45	24	Yes	Cost, good	Poor inter	Yes	No
3network,	Nigeria	Male	30-45	7	No	Good netw	Internatio	Yes	Yes
3network	Nigeria	Female	30-45	24	No	It's fast, nc	Expensive	Yes	No
EE	Nigeria	Female	25-30	12	No	Itâ€™s aff	Nothing I c	Yes	No
Giffgaff	Nigeria	Male	30-45	0-6	Yes	Good	Occasion	Yes	No
Giffgaff	Nigeria	Male	25-30	7	No	Ease of us	None I car	Yes	No
3network	Ghana	Female	45 and abc	24	Yes	Good rece	Nothing th	Yes	No
Vodafone	Nigeria	Female	45 and abc	24	Yes	Good	A bit more	Yes	No
EE, 02	United Kin	Female	25-30	7	Yes	Sufficient	Horrible da	No	Yes
EE	Ghana	Female	45 and abc	24	No	Reliable. G	Long wait	Yes	No
EE, 3netwo	Nigeria	Male	30-45	12	Yes	Connectivi	Sim with p	Yes	No

Figure 3: Real Dataset of Primary Dataset

Network S	Review
EE	Connectivity, network strength, internet
Giffgaff	Cheap bundles
EE	Easy to contact support. Easy to use app. Fair prices
EE	Gifting data to family
EE	Cost , sometimes singal goes off
Giffgaff	Network signal not everywhere
EE	Lack of signal sometimes.
EE	High priced contracts, not enough network coverage in remote areas, comes across as
Giffgaff	Affordable price, easy connections, Good app interface
Giffgaff	Reliable
Lebara	Bonuses, WiFi availability, cost
Telco	Cheap subscription, network speed and network coverage
Telco	I got better deal, good customer satisfaction
Lebara	Steady, fast and reliable
EE	Speed, coverage
Lebara	Coverage, strong network,
Lebara	No credit check necessary, decent cost, ideal for international calls
Telco	Reliable, Cheap and wide coverage
Giffgaff	None
Giffgaff	No 5g
Lebara	International bundle cost
Telco	Network stability, distance to recharge, broadband availability in my area.
Telco	Poor service in some area
Lebara	Non
EE	Nothing
Lebara	High data tariff,
Lebara	Long Wait time to reach customer service

Figure 4 Using Excel to Extract the Reviews

3.3.2. Data Understanding

Data pre-processing is important when it comes to the implementation of the models. The data is downloaded and gotten into a specific format at this stage. Preparing the data for analysis involves adopting a programmatic pattern of cleaning any form of vagueness and reading it into a data frame. it helps in giving better clarity when performing the pre-processing and cleaning using python.

A technique called "Exploratory Data Analysis" (EDA) is especially important in carrying out this project. To perform an EDA, the data was read into a data frame to get an insight about the data. etc. Descriptive statistics of the data were performed which displays the count, mean, and standard deviation of each column of the data.

```
data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 120 entries, 0 to 119
Data columns (total 11 columns):
 #   Column                                                Non-Null Count  Dtype
---  ------                                                --------------  -----
 0   Timestamp                                             120 non-null    object
 1   NetworkService                                        120 non-null    object
 2   Country of Origin                                     120 non-null    object
 3   Gender                                                120 non-null    object
 4   Age range                                             120 non-null    object
 5   How long have you been using your current network provider  120 non-null    object
 6   Have you tried any Network asides the one you are currently on  120 non-null    object
```

```
data.describe()
```

	SeniorCitizen	tenure	MonthlyCharges
count	7043.000000	7043.000000	7043.000000
mean	0.162289	32.370865	64.761692
std	0.368742	24.559231	30.090047
min	0.000000	0.000000	18.250000
25%	0.000000	9.000000	35.500000
50%	0.000000	29.000000	70.350000
75%	0.000000	55.000000	89.850000
max	1.000000	72.000000	118.750000

Figure 5: understanding my data

```
data.isnull().sum()

Timestamp                                                                  0
NetworkService                                                             0
Country of Origin                                                          0
Gender                                                                     0
Age range                                                                  0
How long have you been using your current network provider                 0
Have you tried any Network asides the one you are currently on             0
What are the three good things about your current Network                  0
What are the three bad things about your current Network                   0
Would you recommend your current network to family, friends, colleagues.   0
Churn                                                                      0
dtype: int64

updated_data= pd.read_csv("Reviews_Sentiment.csv")
```

Figure 6 No missing values primary data

```
#checking for missing values
data.isnull().sum()

customerID          0
gender              0
SeniorCitizen       0
Partner             0
Dependents          0
tenure              0
PhoneService        0
MultipleLines       0
InternetService     0
OnlineSecurity      0
OnlineBackup        0
DeviceProtection    0
```

Figure 7 : No missing values for secondary data.

The diagram above displays an output after writing a few lines of codes to understand my data.

3.4 Data Analysis

The literature review in this research was done by the collection of data, analysis of the data, pre-processing, and evaluation of the data. The approach of this research is based on using inductive reasoning. Inductive reasoning is the means of drawing conclusions through specific cases. (Hamad, 2007). To gain the results that were needed for the success of this research, a few steps were taken which included: (1) report types; (2) study design, and (3) Behavioural factors (Technologies, training, networking, mentorships, information, and cultural attitude). All these, were vital for my research as they gave me room as the researcher to study, evaluate, and understand the research.

3.4.1. Data Visualisation

Various data visualisation techniques were implemented to gain an improved perspective of the data. This includes the use of a heatmap to understand the correlations between the target and dependent variables, and the use of a histogram, and pie chart. This is to gain more insight into the dataset. For instance, Figure 46 with a few lines of code, shows the most frequent words in our dataset. Also, it can be seen in Figure 8 Customers that have opted for the month-to-month plan have a high percentage of leaving the organisation while customers that did the 2-year monthly plans have the least. This indicates that customers with the longer plan are not affected by the high charges compared to customers going for the monthly plan.

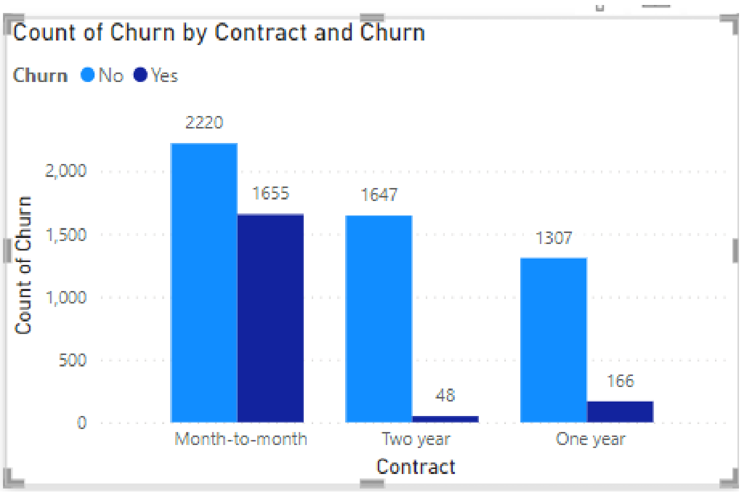


Figure 8: Displaying count of customers according to their Contract

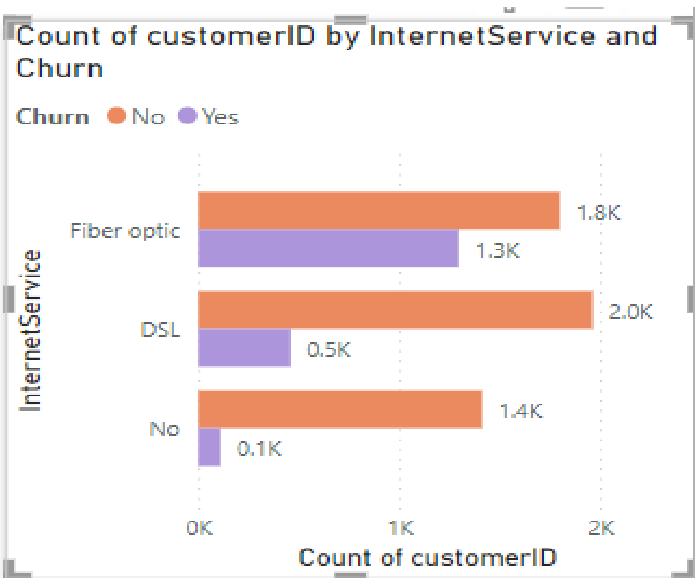


Figure 9: Displaying count of customers according to their Internet service

Also, Figure 9 gives an insight into how network providers are to monitor the services they offer to customers because they might not be satisfactory to the customer. The Fiber Optic plan had the highest rate of customers willing to churn, and this might be because of the poor network, high charges, or low/no discount.

Additionally, in Figure 10, the visuals tell us that there were more respondents that agreed to remain. These responders were between the ages of 18 – 60. Responders that were not willing to remain were little in numbers, majorly younger people between 20-30. This tells us that most times, young people can be easily influenced by peer groups and families that are not using the same network service that they are using. The only exception to not being influenced is when they have experience with the plan and are certain that the network providers are good. Also, older people have the tendency to remain.

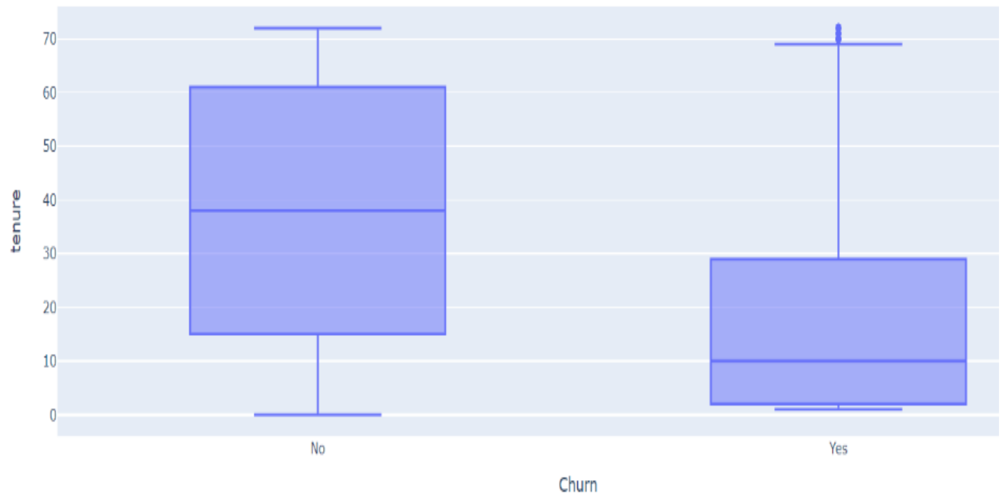


Figure 10: Displaying count of customers according to their Internet service

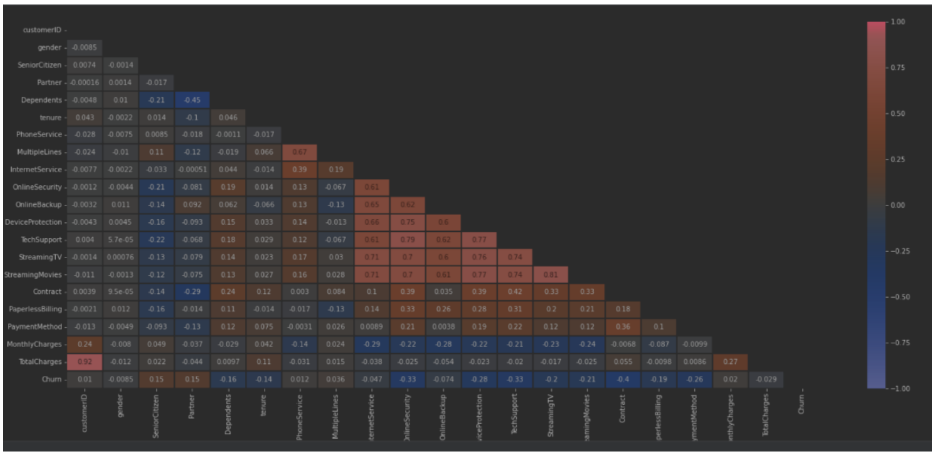


Figure 11: Correlation Matrix

I went further to understand my features well. From the plot, we can say that Tenure, Total charges, monthly charges, internet services, and Streaming TV, Tech support all have strong correlations and are factors that can make customers remain or leave. We can also see that between Churn, the target variable, and Internet Online backup there is a strong relationship but in an opposite direction.

To understand the total count, the Weka figure shows how the male gender are more than the female gender.

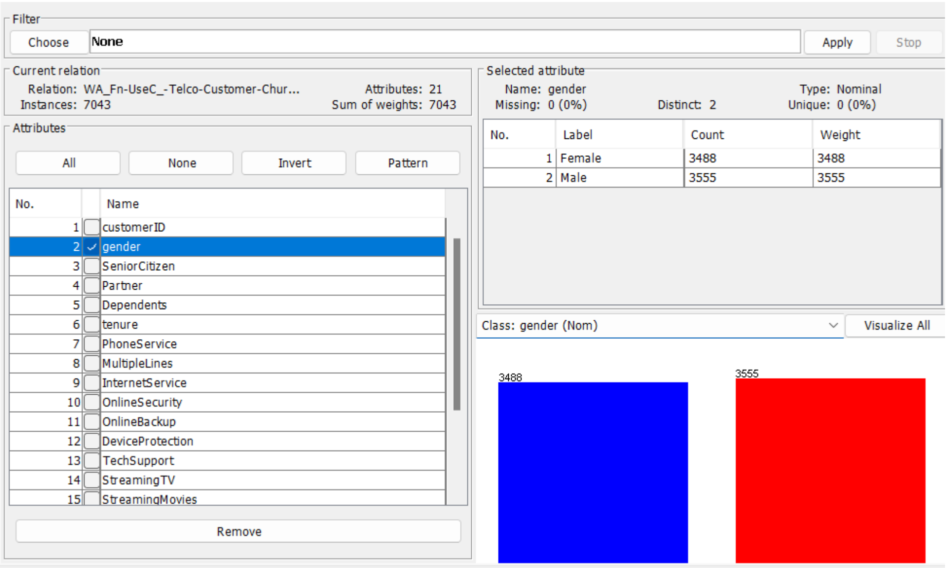


Figure 12: Gender Count

The dataset also displayed the count of customers in the data that would remain and those that would leave. It shows that there is a high number of customers that would remain rather than leave.

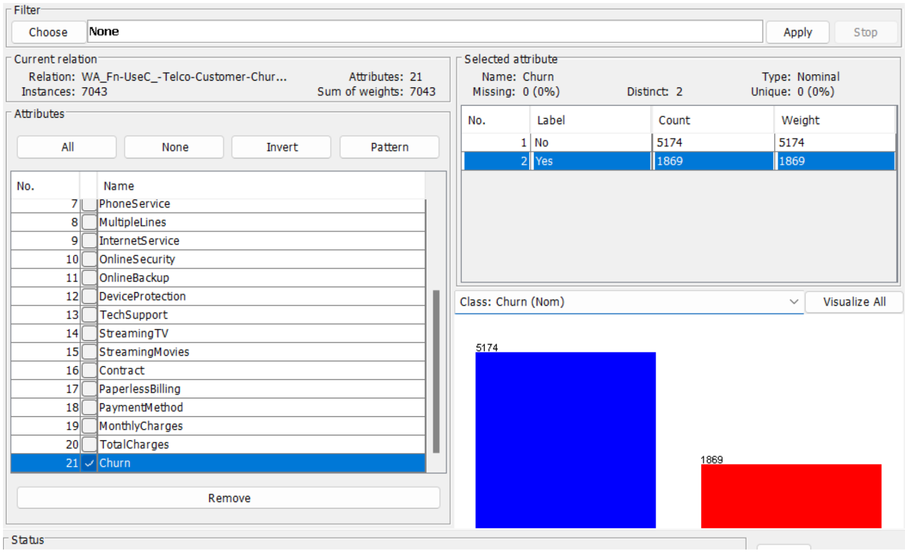


Figure 13: Churn Count

3.5. Instrument Used

The Excel tool was used to format the primary dataset into a CSV file (comma-separated values). The primary data was developed by creating a survey through google docs and was shared using a popular network platform known as WhatsApp. According to Statista, as of January 2022, the social network had over 2 billion users who can interact with friends, create a group chat, and share similar interests. 80% of the responders were from WhatsApp groups.

Additionally, the resources used in this research, are study findings gotten from previous works. Secondary Literature was gotten from published works from Research Gate, direct, Google, Google Scholar, ProQuest, E-books, ScienceDirect, etc. This is a cross-sectional study because resources were collected from numerous studies, reports, theses, and journals from different countries. Different tools such as Weka, Python, Power-Bi, Excel, and Streamlit were used to work on the datasets to gain insights. It is important to note that just as the Weka tool was for customer segmentation, and PowerBI for visualising the data, Python was used to train the deep learning models in preparation for the sentiment analysis. This will expatiate below.

3.6. Implementation and Algorithm of Artificial Intelligence Techniques Used

During this research, there will be a prototype development. This will include the use of Data visualisation tools, machine learning models, deep learning models, and the application of natural processing language to gain insights. The programming language referred to as Python will be solely used in achieving these results. To work on my data, the classification method was used. Classification analysis is a supervised machine learning technique that can be used to predict categorical labels. Its output is always discrete. These are data that have qualitative characteristics. One of the objectives of this research is to use suitable models that would work for my research study.

3.6.1 Sentiment Analysis

Sentiment analysis is categorized under the natural language processing field which can also be regarded as an application of artificial intelligence techniques. It majors on the polarity behind a given statement by assisting one to fully understand the reasons behind a subject matter. Philander (2016) in his book “Twitter Sentiment Analysis: Capturing Sentiment from Integrated Resort Tweet” sees sentiment analysis as the extraction of customers’ thoughts and views from data. Decisions are carried out based on the comprehensive analysis of opinions that have been pulled out. This study, shows how sentiment analysis can be used to create standard real-time metrics for customer satisfaction using survey data.

Figure 14: Sentiment Analysis displaying tweets from Twitter from [Tutorial Power BI: Sentiment Analysis con Cognitive Services - Vandalytic - tu blog de Power BI y más en español]

3.6.2. Random Forest

This was created by Leo Breeman and Adele Cutler in 2001. It is regarded as an ensemble combination of multiple decision trees that is implemented to yield forth a result. The Random Forest Classifier has been used by several researchers to predict as well as gain insight through their predictions and this has brought up accuracy in their results. Its major benefit is that it handles large datasets and lacks sensitivity. The Random Forest is a model that falls under ensemble learning which makes the model work fast. The random forest has unique features: the “voting strategy” and the “bagging.” The voting strategy helps to correct undesired decision trees to avoid overfitting when training the data while the bagging helps to select samples and replaces them with a training dataset that fits into the tree samples.

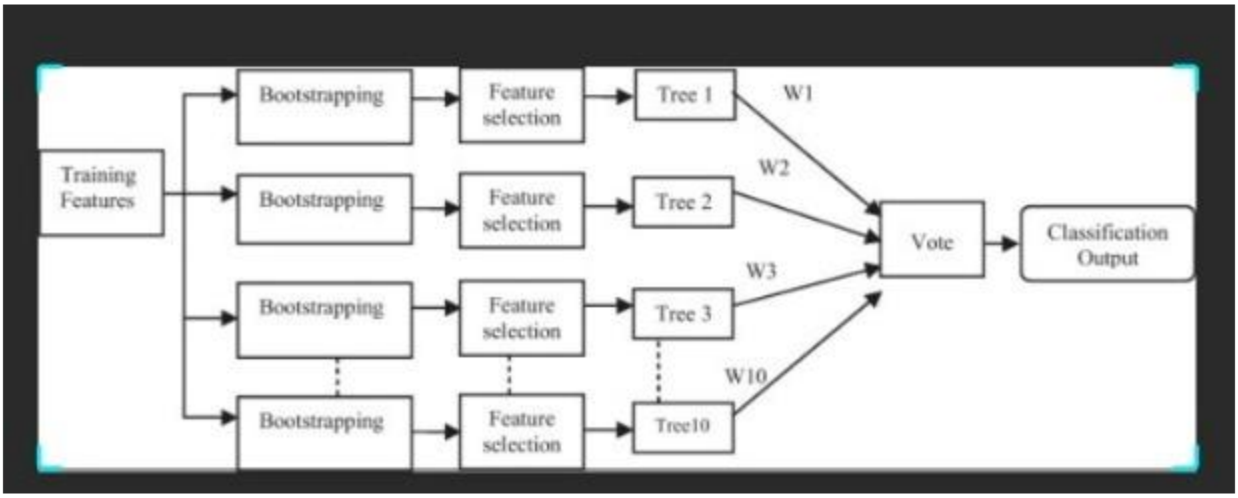


Figure 15: Random Forest Diagram from [Random Forest Classifier - an overview (pdf) | ScienceDirect Topics]

3.6.3 Support Vector Machine (SVM)

The support Vector Machine was created by Vladimir Vapnik in 1999. The support vector is used for both regression and classification models. It is used to determine a line that can separate data into classes. Its primary features are Hyperplane, Boundary line, kernel trick, and the margin. It has its uniqueness from the fitting of the best line within a threshold. The uncommon feature of the SVM is its ability to address overfitting, it also works with lots of features and numbers of observations which the datasets used in this project possess. Support Vector Machine is memory efficient, explicit, and creates accurate boundaries. (Nhu et al 2022).

Decision Boundary Line

The intention of the support vector Machine is to set a line that would divide the data points into classes. The decision boundary line is meant to be broad and feasible as possible for the distance in-between every point and the boundary line to be maximized.

Hyperplane

The hyperplane is used to separate the data into classes. The farther the distance is from the hyperplane, the better the support vector because it gives room for accurate prediction while we have the Margin which is the distance between the hyperplane and vector.

Kernel Trick

The kernel trick is used in the support vector machine. This is used to build non-linear and linear transformations in the dataset. Here, the kernel trick applies the non-mathematical operations and transforms the dataset's current characteristics into new ones.

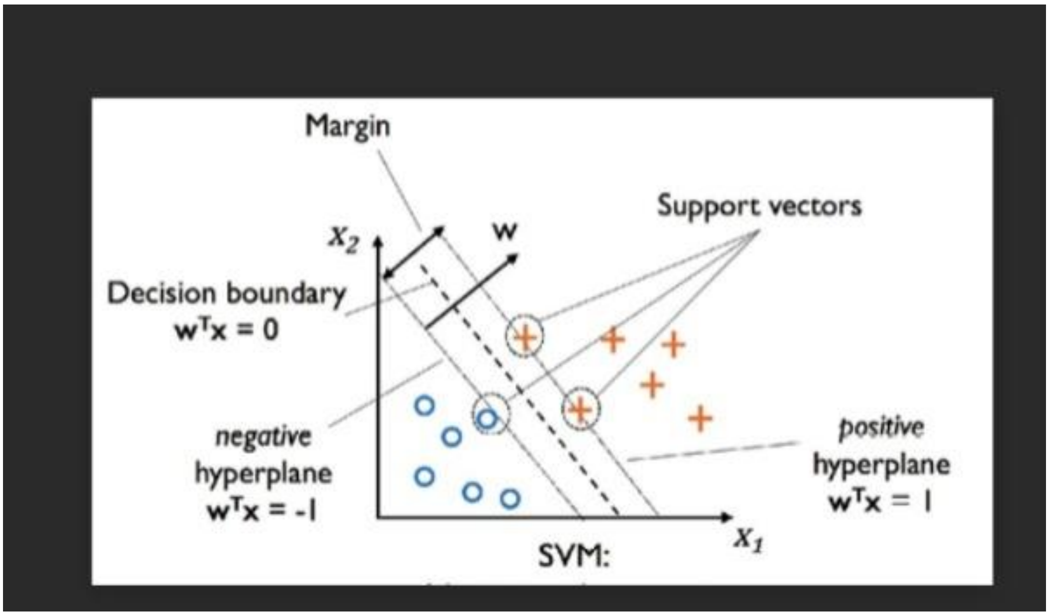


Figure 16: Support Vector Machine Diagram from [SVM Algorithm as Maximum Margin Classifier - Data Analytics (vitalflux.com)]

3.6.4 Decision Tree

The decision tree is a machine learning model that is used for regression and classification projects. The term comes from splitting the dataset into small sets till the data is divided into a single value. The decision model is in form of a tree format that extends like a flow-chart. The decision tree comprises of the root node, leaf nodes, and branches. The node starts with a single node that splits into several choices that are referred to as branches, which come along with their consequences. The decision tree uses what we know as “recursion split” and “pruning.” For the recursion, it is used to split the data which starts from the root, where the number of features in the data determines the number that would be split. The splitting stops when all data have been duly classified. What the pruning does, is increase the performance of the decision tree model by removing branches that contain little or no feature importance to the target variable. The process either starts from the top or bottom.

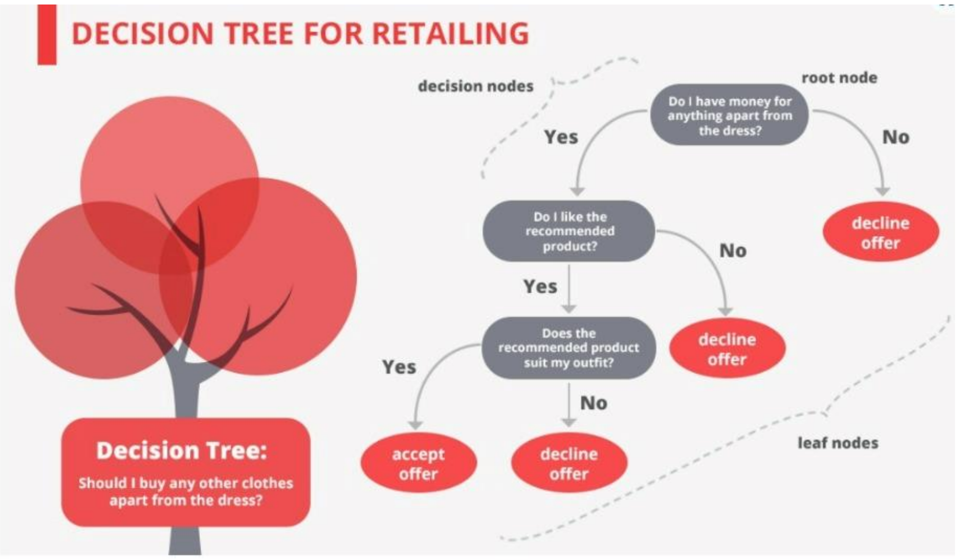


Figure 17 : Decision Tree Diagram from [How AI, machine learning, and big data reshape SMB in retail (innovecs.com)]()

3.6.5.1. Dimension Reduction

Lubaba et al (2022) defined dimension reduction as a statistical way of reducing the number of variables in a data. The variable is removed is for the purpose of reducing redundancy in the dataset. In dimension reduction, only the vital variables are not removed from the dataset. There are several methods for applying dimension reduction and they include the following; i) Random Forest: ii) Principal Component Analysis iii) Locally Linear Embedding iv) Isometric Mapping. This research would be making use of the Principal Component Analysis (PCA).

3.6.5.2. Principal Component Analysis

This can be referred to as an unsupervised variable reduction algorithm used in removing redundancy in a dataset by analysing how important each feature is to the variable. It gets on by selecting a set of factors from the existing dataset.

```
components = None
pca = PCA(n_components=components)
pca.fit(scaled_x_train)
print('Variances(Percentage):')
print(pca.explained_variance_ratio_ * 100)
```

Figure 18: snippet code displaying Principal component analysis

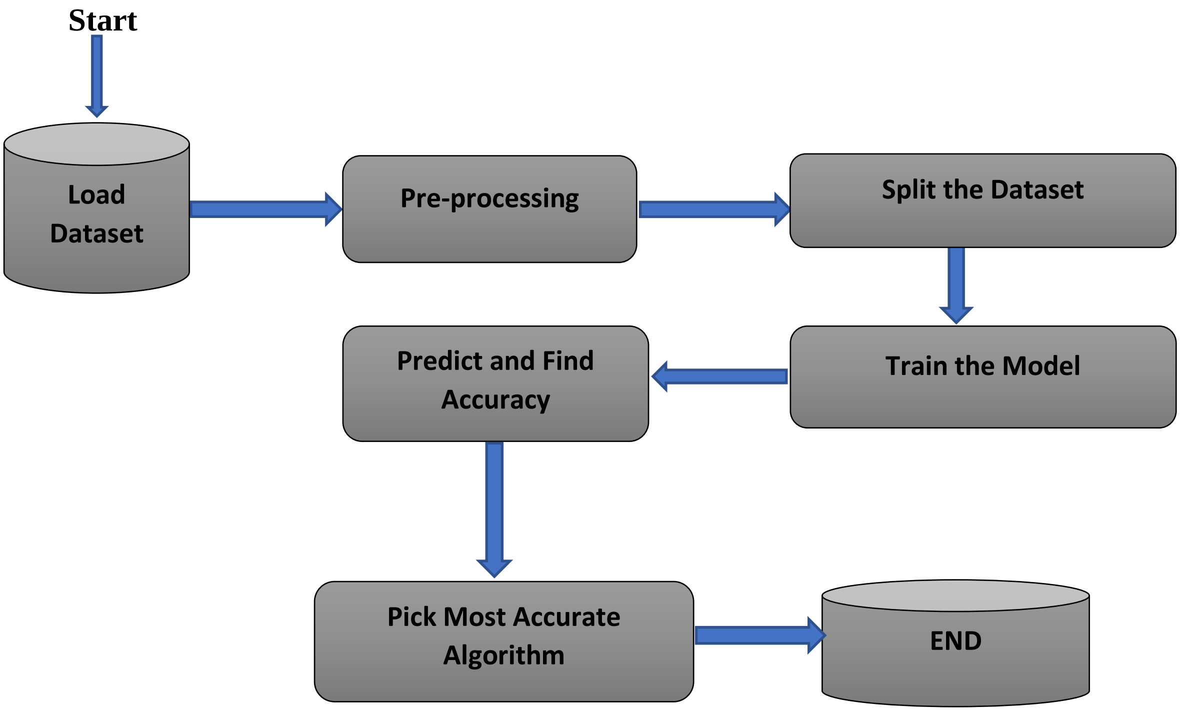
3.6.5.3 Feature Scaling

This is a scaling method that removes data duplication in a dataset before prediction. It is a method of unifying the data as well as helping in reducing the training time of the data. It is an important stage when pre-processing because it helps in fixing the performance of the model. This research made use of the Normalisation technique. To normalize, the use of MinmaxScaler is imported from the sklearn library.

```
scaler = preprocessing.MinMaxScaler(feature_range=(0,1))
scaled_x_train = np.array(X_train).reshape(len(X_train),45)
scaled_x_train = scaler.fit_transform(scaled_x_train)
scaled_x_test = np.array(X_test).reshape(len(X_test),45)
scaled_x_test = scaler.transform(scaled_x_test)
```

Figure 19: snippet code exhibiting the minmaxscaler()

3.6.5.4. ARCHITECTURAL FRAMEWORK



3.6.6 Deep Learning Model

The transformer is primarily used in Natural Language Processing, it is regarded as a deep learning model, built by Google brains in 2017, to solve problems related to Natural Language Processing. It is regarded as a model that substituted the Recurrent Neural networks due to the limitations that the RNN had by not being able to process input data all at once. The transformer model came into place to cover this loophole. The transformer model helps in processing data at once and with the attention Mechanism, the transformer model performs well without recurrent measures, which is not possible with the RNN. The goal of the attention Mechanism is to attract relevant information to reduce redundancy, and this is to yield forth a greater output. Just as the RNN comprises the LSTM (Long Short-Term Memory), BI-LSTM, and GRU, the transformer has the BERT, ROBERTA, and AlBERT. They are regarded as pre-trained models. During this research, the ROBERTA model will only be discussed as it fulfils the objective of this research. (Mayukh et al , 2021).

3.6.6.1 Pre-Trained Language Model

With Natural language processing, comes interpretation, translations, and sentiment analysis. Deep learning has always been a powerful tool in National Language processing. However, the setback linked to DLM, is that it requires extensively large datasets which makes the tasks of training the model challenging and time-consuming. These issues led to the emancipation of pre-training. It is a method where millions of tweets have been applied to models to learn and understand the features attached to the model.

3.6.6.2 Roberta Model

The Roberta Model stands for “Robustly Optimised Bert Pre-Trained Approach.” It was built in 2018 to perform the limitations of the BERT model through the modifications of hyperparameters. The Roberta was trained with an extensively larger dataset after google brains realised that the BERT model was undertrained. This yielded froth momentous results after the training.

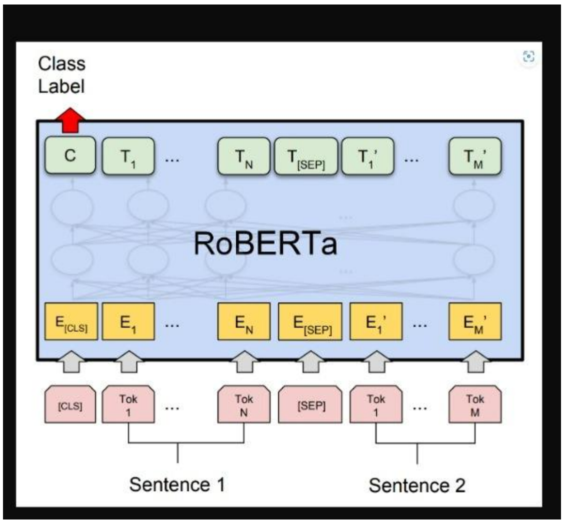


Figure 20; Architectural structure from [roberta-multichoice.png-0000000927-36fb4743.png (1134×1002) (paperswithcode.com)]

3.6.6.3. Vader Model

"Valence Aware Dictionary sEntiment Reasoner Model" is a pre-trained model that makes use of words in a sentence by combining the words to specify how positive, negative, or neutral the words are. The Vader fully relies on the lexical approach to determine the emotion behind the words. The Valence score is a major feature of the Vader model. This is a computational summing up of each word to determine the polarity of the words. The valence score is also regarded as the "Compound Score. (Mayukh et al, 2021).

```
import nltk
nltk.download('vader_lexicon')
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
sia= SentimentIntensityAnalyzer()
```

Figure 21: snippet code of the vader model

3.6.6.4 Aspect-Based Method

This is a type of model that helps to identify specific words, to show how related the sentence is. Rohan et al (2021), described Aspect based Analysis to be challenging and complicated as it encompasses identifying, classifying, and aggregating words. He also explained that what makes ABSA complex, is the ambiguity of the sentiment in the text. He further explained that the key feature of ABSA is the ability to capture key roles in a sentence for easy understanding of the context of the word which makes the information hold more meaning for better and more productive performance. The ABSA was implemented in this research work to see its performance.

```
aspects = []
for sentence in reviews:
  doc = nlp(sentence)
  descriptive_term = ''
  target = ''
  for token in doc:
    if token.dep_ == 'nsubj' and token.pos_ == 'NOUN':
      target = token.text
    if token.pos_ == 'ADJ':
      prepend = ''
```

Figure 22 snippet code of aspect based

3.6.6.5. Deep learning Model Pre-processing

Prior to training the model, the data was cleaned through the following process.

Tokenization: This is the breaking down of words into smaller words to reduce the size of the vocabulary.

Token- Porter: This helps in breaking down the sentence into simpler words.

Frequency: It helps to understand how many words have occurred in the text.

Part of Speech: This is done by assigning parts of speech such as nouns, prepositions, adjectives etc. to each word for the system to understand and for a better analysis.

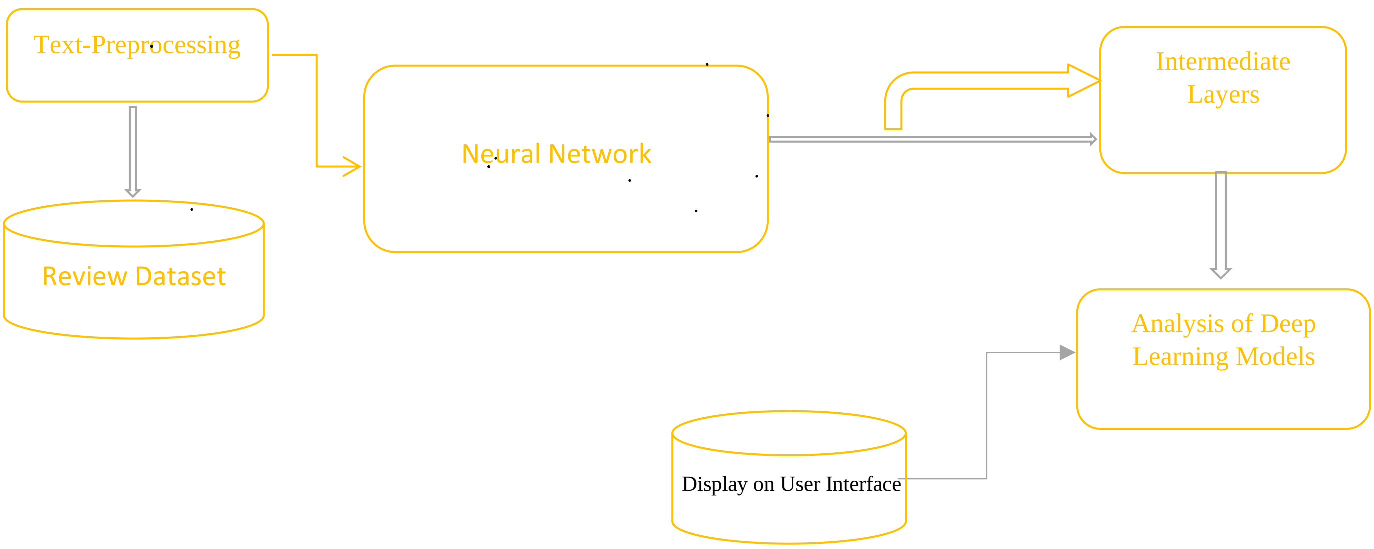
```
#each of the words have been given their part of speech
tokens= nltk.word_tokenize(words)
tagged= nltk.pos_tag(tokens)
tagged[:10]

[('I', 'PRP'),
 ('got', 'VBD'),
 ('better', 'RBR'),
 ('deal', 'NN'),
 (',', ','),
 ('good', 'JJ'),
 ('customer', 'NN'),
 ('satisfaction', 'NN')]
```

Figure 23: snippet code of the part of speech

Chunk: This is to group each word into bigger entities or tokens of words into chunks.

Stop Words: This is referred to words that cannot give meaning to a statement. Words like “at,” “the,” “a,” and “all.” They do not have meaning and therefore are removed.

3.6.6.6 ARCHITECTURAL FRAMEWORK OF DEEP LEARNING MODEL



3.6.7. Clustering Analysis

Clustering Analysis can be regarded as one of the powerful tools when we are talking about unsupervised learning. Ryan (2015) explained clustering as a “Machine learning Algorithm” where results are derived from features in a dataset. Clustering is an important way to identify groups and similar patterns through the combination of related characteristics from your data to gain more insight and use the results you have gotten to solve problems related to your target. There are several literature reviews and works on predictive analysis on customer satisfaction and how to reduce churn rate. However, little or no work has been done on how to specifically use clustering analysis through the application of tools to segment customers, understand, satisfy, and determine those who are going to leave based on their complaints. Just like I mentioned earlier, an organisation needs to understand that satisfying customer cannot be generalized because we have different customers with diverse needs. This is based on the survey that was carried out. Hence, in as much as it is important to predict customers that are going to leave, it is also vital to understand why they are not satisfied with the product or services. This is where clustering becomes vital so that organizations can use customer's needs to figure out when a customer is going to leave or remain. This would also help companies on how and who to channel their adverts on. I would be using the Weka software that was created by IBM. It is an interactive statistical software and data visualization tool to gain greater insights from a dataset.

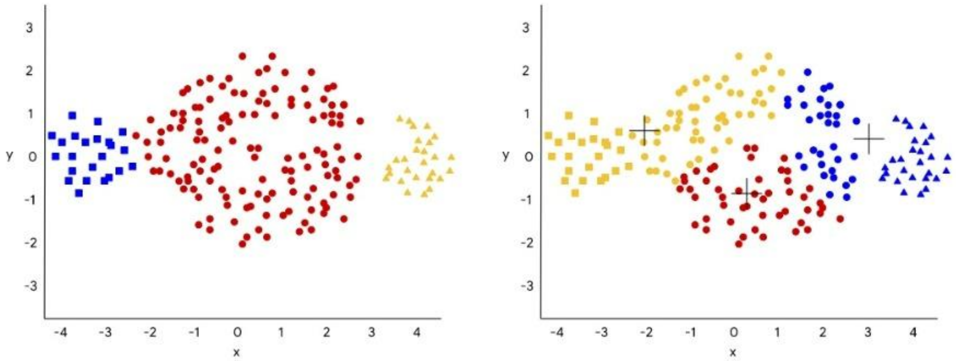


Figure 24: Clustering analysis diagram from [k-Means Advantages and Disadvantages | Clustering in Machine Learning | Google Developers]

3.6.8 Software Used

Several Scientists have produced various methods for creating a perfect approach for analysing and solving problems. Some of these tools include the following; Python language, SPSS, Tableau, Weka, PowerBi, Excel, and Streanlit, tkinter. However, for the purpose of this research, I will be focusing on Weka, Python, PowerBI, and Streamlit.

Weka: Weka is known as "Waikato Environment for Knowledge Analysis" (WEKA), It is a software that is used for machine learning problems. Its features include Clustering, data pre-processing, visualisation, prediction, and ranking. The Weka was used for segmenting customers.

PowerBI: This is a tool founded by the Microsoft team to create interactive visuals that can be used to tell a story and share insight that can lead to success. The PowerBi was used in this research to modify my data and to create insights.

Python: This is a programming language created by Guido van Rossum, using codes to solve problems. It helps in giving meaning to the data.

Streamlit: It can be referred to as an “open-source library” that helps Machine learning Engineers, data scientists, and developers to create an interface for easy and better interaction. The streamlit helped to display what I created; it also gives room for users to input words to see the output of what they have inputted.

```
input_text = st.text_input("Get Predictions", value="", max_chars=None, key=None, type="defaul
if input_text:
    reason: list = []
    key_list = reasons.keys()
    for key in key_list:
        for k in reasons[key]:
            if input_text.__contains__(k):
                if reason.__contains__(key):
                    continue
                else:
                    reason.append(key)
    for n in networks:
        if input_text.__contains__(n):
            network = n
```

Figure 25 snippet code for streamlit

Applying the Roberta model on the streamlit was because of its high-performance score. The code above displays a for loop written to iterate through a list that contains names of network providers and a dictionary that contains specific words that were used by respondents during the survey.

Figure 26: snippet code displaying user's output

The intention is to make it easy for companies to know how they can satisfy their company and to know when a customer is dissatisfied, and likely to leave. The demo was tested using students, and respondents that were used in the survey.

3.6.9. Libraries and Platform

Libraries are also known as modules that has been created to make programming easier for programmers. The following modules will be used in this study.

Pandas: This can be referred to as a module used in python for solving problems that include finances, analytics, and statistics.

Numpy: This is a scientific method that is used in python that helps in providing multidimensional objects. It can help to transform and manipulates objects.

Sklearn: This is a library that is used in python for machine learning problems. With sklearn comes machine learning models such as Random Forest, decision tree, regression, classification etc.

Seaborn: This is a library used to create statistical, informative, and attractive visuals for easy understanding.

Matplotlib: This is a library that is used to create interactive visuals.

Spacy: This is an "open-source Library" that has been built for natural language processing to gain greater insight. Spacy is mostly used for advanced sentiment analysis projects.

Natural Language ToolKit(NLTK); This is a platform that is used to build programs that work with understanding human languages. Its features come with tokenization, porterstermming, stopwords, parsing, chunking, etc.

```
porter = PorterStemmer()
def tokenizer(words):#breaking down the text with the tokenizer
    return words.split()

def tokenizer_porter(words):
    return[porter.stem(words)for words in words.split()]
```

Figure 27: snippet code of tokenization and porterstermming through the use of Nltk library

TextBlob: This is a library in Python that helps to interpret and process text in data.

WordCloud: This is a platform that helps in displaying words in varied sizes. The larger words are the most frequent words in the data.

```python
#seeing how well the sentiment are distributed
words= ' '.join( [reviews for reviews in updated_data['Reviews']])
word_cloud = WordCloud(width = 500, height=300, random_state =21, max_font_size= 119).generate(words)
plt.imshow(word_cloud, interpolation = 'bilinear')
plt.axis('off')
plt.show()
```

Figure 28; snippet code to display word cloud in python

3.7. Ethics

The ethics consideration form was filled out and approved by the University to ensure that documentation of research was carried out accurately and the author used were credited for their work (Figure 44).

3.8. Intended Outcomes and Dissemination

Regarding the findings and result that was concluded in the study, the researcher intends to use the information to inform readers about the different techniques that can be enforced to determine customer satisfaction and the way businesses can enhance their relationship with customers.

CHAPTER FOUR

RESULT AND DISCUSSION

4.1. Evaluation Metrics

In this research, to get the model prepared, the categorical columns were encoded to numerical values. The one-hot encoding was used because it gives room for flexibility to pick the column to be encoded. Three models were used for prediction which is; The Random Forest Model, Decision Tree Model, and Support Vector Machine. The data were normalised by using the Minmaxscaler. This was able to set the feature range between 0 and 1. To ensure prediction, the model needs to be evaluated to understand its accuracy and performance. For the classification model, the precision, recall, and f1-score are vital. This is determined by the True Positive, True Negative, False Positive, and False Negative,'

True Positive: N0 of customers that are in the churn group and are predicted correctly.

True Negative: N0 of customers that want to churn and predicted correctly.

False Positive: N0 of customers that are not willing to churn but predicted incorrectly.

False Negative: N0 of customers that have churn but predicted incorrectly as customers that have not churn.

The Recall tells the ratio of real customers that have churned correctly and identified by the algorithm.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}$$

Precision: Tells the predicted churners that are correct.

$$\text{Precision} = \frac{TP}{TP + FN}$$

F1 Score: This is getting the sum of the average derived from the precision and recall.

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

4.2. Results

Random Forest

For the Random Forest, after prediction, the result showed that the performance score was good. The F1 score had 87%, the Precision score had 81%, the Recall score was 94%, and the Accuracy score was 80% as shown in Figure 29 below.

```
class_report = classification_report(y_test, rfc_y_prediction)
print('Report of Classification;\n', class_report)

 Report of Classification;
               precision    recall  f1-score   support

           0       0.81      0.94      0.87      1555
           1       0.70      0.40      0.51       555

    accuracy                           0.80      2110
   macro avg       0.76      0.67      0.69      2110
weighted avg       0.78      0.80      0.78      2110
```

Figure 29 : Random Forest Before Application of PCA.

After prediction, the Principal Component Analysis was applied to see how well the model would perform. From the diagram, the accuracy score dropped from 88% to 77%, while the recall score became higher with 96% as shown in Figure 30. This was after tuning the model.

```
confusion_matrix(y_test, rfc_y_prediction)
class_report = classification_report(y_test, rfc_y_prediction)
print('Report of Classification;\n', class_report)

 Report of Classification;
               precision    recall  f1-score   support

           0       0.78      0.96      0.86      1555
           1       0.67      0.25      0.37       555

    accuracy                           0.77      2110
   macro avg       0.73      0.60      0.61      2110
weighted avg       0.75      0.77      0.73      2110
```

Figure 30: Random Forest with the use of Principal Component Analysis

Support Vector Machine

The support vector machine as shown in Figure 31 had 87% as the F1score, and an 86% as the precision score. Accuracy measure was 80%.

```
Report of Classification;
              precision    recall  f1-score   support

           0       0.84      0.90      0.87      1555
           1       0.64      0.51      0.57       555

    accuracy                           0.80      2110
   macro avg       0.74      0.70      0.72      2110
weighted avg       0.79      0.80      0.79      2110
```

Figure 31 Support Vector Machine

After applying the PCA, the score reduced from 80% to 78% as revealed in Figure 32.

```
[0 1 0 ... 0 1 0]

77.77251184834122
```

```
class_report = classification_report(y_test, svc_y_prediction)
print('Report of Classification;\n', class_report)
```

```
Report of Classification;
              precision    recall  f1-score   support

           0       0.81      0.91      0.86      1555
           1       0.62      0.40      0.49       555

    accuracy                           0.78      2110
   macro avg       0.72      0.66      0.67      2110
weighted avg       0.76      0.78      0.76      2110
```

Figure 32: After the Principal Component Analysis

My expectation was to have the score increase after tuning, but it still had the same score after tuning.

Decision Tree

The Decision Tree model was used for prediction. The accuracy score that was gotten after the model was 79% for accuracy score, 91% for f1score,

```
class_report = classification_report(y_test, dt_y_prediction)
print('Report of Classification;\n', class_report)
```

```
Report of Classification;
              precision    recall  f1-score   support

           0       0.83      0.91      0.87      1555
           1       0.66      0.46      0.54       555

    accuracy                           0.79      2110
   macro avg       0.74      0.69      0.71      2110
weighted avg       0.78      0.79      0.78      2110
```

Figure 33: Decision Tree

and 83% for precision. With the PCA application, the accuracy score reduced to 76%.

```
class_report = classification_report(y_test, dt_y_prediction)
print('Report of Classification;\n', class_report)
```

```
Report of Classification;
               precision    recall  f1-score   support

           0       0.81      0.89      0.85      1555
           1       0.57      0.41      0.48       555

    accuracy                           0.76      2110
   macro avg       0.69      0.65      0.66      2110
weighted avg       0.75      0.76      0.75      2110
```

Figure 34: Decision Tree after the application of the PCA.

Weka Tool

With Weka, the F1 score was 84%, 73% for the precision score. This demonstrates that there is a slight disparity in the performance score using the Weka technique.

Performance Score

Model	F1 Score	Precision	Recall	Accuracy	Accuracy After PCA
Random Forest	87%	81%	94%	80%	77
Support Vector	87%	86%	90%	80%	78%
Decision Tree	91%	83%	91%	79%	76%
Weka Tool	84%	73%	1.0	73%	

Figure 35: Performance score for supervised learning model

4.3. Deep learning Models Results

The sentence text “I have a better deal, good customer satisfaction”, had a score of 82% from Vader, the Roberta Model was 92%, and 62% for the Aspect based.

Performance Score for deep learning Model

Model	Review	Neutral	Positive	Negative	Score
Vader	I have a better deal, good customer satisfaction	36%	68%	0.0%	82%
Roberta	I have a better deal, good customer satisfaction	0.9	90	0.00%	92%
Aspect Based	I have a better deal, good customer satisfaction	70%	50%	0.0	65%
Vader	Facing Network issues, mobile data connectivity issues	1.0	0.0	0.0	0.0%
Aspect Based	Facing Network issues, mobile data connectivity issues	0.0	0.0	0.0	0.0
Roberta	Facing network issues, mobile connectivity issues	0.0	0.1	80%	99%

Figure 36: Performance score of deep learning Model

4.5. Discussions

Scholars like Samira et al (2016), in their work "Comparison of supervised machine learning techniques for customer churn prediction based on analysis of customer behaviour", made use of the Random Forest, the Support Vector, and the Decision Tree for prediction. After their prediction, the Random Forest and Support Vector had the highest score. They went with the Random Forest. In this research, the RF also came out with the highest scores after my prediction, and Decision had the lowest accuracy scores. From this, the Random Forest and Support Vector are strong predictive power and great contenders in predicting the customers that are going to churn. Though the Decision Tree had an accuracy score of 79%. However, the difference was little. Samira et al (2016) suggested that tuning the models increased their scores and got the error score reduced. However, there was no difference after tuning my models.

Mujahid et al (2021) used the Weka technique along with other deep learning models to determine their result. Using Weka, helped in balancing the data and helped to reduce any form of overfitting. Using the Weka tool, I decided to use the simple K means methods as it helps scales large dataset perfectly, hence creating room for a better performance. My intention was to have an insight into customers that had similar features and the common reasons they wanted to leave or remain.

The number of clusters used were respectively 2, 3, and 5. Clustering into 2, the mean squared error was 48%, into 3, was 42%, while clustering into 5, had the lowest sum of mean squared error with 29%.

```
,1,No,No,3,Yes,No,'Fiber optic',No,No,No,No,No,No,Month-to-month,Yes,'Electronic check',70.3,235.5
,0,No,No,1,Yes,No,'Fiber optic',No,Yes,No,No,No,No,Month-to-month,Yes,'Electronic check',75.7,75.7
,No,Yes,69,Yes,Yes,DSL,Yes,Yes,Yes,Yes,Yes,No,'Two year',Yes,'Credit card (automatic)',79.45,5502.55
,Yes,Yes,30,Yes,No,'Fiber optic',No,Yes,Yes,No,No,No,Month-to-month,No,'Bank transfer (automatic)',82.05,2570.2
,0,Yes,No,26,Yes,No,DSL,No,Yes,Yes,No,Yes,Yes,'Two year',Yes,'Credit card (automatic)',73.5,1905.7

obally replaced with mean/mode

troids:
                                 Cluster#
             Full Data                  0                  1                  2                  3                  4
              (7043.0)           (2043.0)            (853.0)           (1264.0)           (1528.0)           (1355.0)
=====================================================================================================================
                  Male               Male             Female               Male               Male             Female
                0.1621             0.2012             0.1758             0.0736              0.034             0.3218
                    No                 No                 No                Yes                 No                Yes
                    No                 No                 No                Yes                 No                 No
               32.3711            12.6995            33.2802            56.8687            30.5681            40.6399
                   Yes                Yes                Yes                Yes                Yes                Yes
                    No                 No                 No                Yes                 No                Yes
           Fiber optic        Fiber optic        Fiber optic                DSL                 No        Fiber optic
                    No                 No                 No                Yes No internet service                 No
                    No                 No                Yes                Yes No internet service                Yes
                    No                 No                 No                Yes No internet service                Yes
                    No                 No                 No                Yes No internet service                 No
                    No                 No                 No                Yes No internet service                Yes
                    No                 No                 No                Yes No internet service                Yes
        Month-to-month     Month-to-month     Month-to-month           Two year           Two year     Month-to-month
                   Yes                Yes                Yes                Yes                 No                Yes
      Electronic check   Electronic check   Electronic check Credit card (automatic)  Mailed check   Electronic check
               64.7617            65.2802            69.9506            81.1382            21.1041            94.6684
             2283.3004           856.6249          2335.9661          4744.9265           673.0794           3920.715
```

Figure 37: Clustering the data into 5

```
======================================================================================================
customerID                          7590-VHVEG           7590-VHVEG          7469-LKBCI          7795-CFOCW
gender                                    Male               Female                Male                Male
SeniorCitizen                           0.1621               0.2171              0.0341               0.166
Partner                                     No                   No                  No                 Yes
Dependents                                  No                   No                  No                  No
tenure                                 32.3711              19.9877             30.5429              53.796
PhoneService                               Yes                  Yes                 Yes                 Yes
MultipleLines                               No                   No                  No                 Yes
InternetService                    Fiber optic          Fiber optic                  No                 DSL
OnlineSecurity                              No                   No No internet service                 Yes
OnlineBackup                                No                   No No internet service                 Yes
DeviceProtection                            No                   No No internet service                 Yes
TechSupport                                 No                   No No internet service                 Yes
StreamingTV                                 No                   No No internet service                 Yes
StreamingMovies                             No                   No No internet service                 Yes
Contract                        Month-to-month       Month-to-month            Two year            Two year
PaperlessBilling                           Yes                  Yes                  No                 Yes
PaymentMethod                 Electronic check     Electronic check        Mailed check Credit card (automatic)
MonthlyCharges                         64.7617              71.3403             21.0888             85.7963
TotalCharges                         2283.3004            1507.3564            671.6866           4712.7969
Churn                                       No                   No                  No                  No


Time taken to build model (full training data) : 0.04 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      3413 ( 48%)
1      1527 ( 22%)
2      2103 ( 30%)
```

Figure 38: Clustering into 3

```
=================================================================================
customerID                       7590-VHVEG         7590-VHVEG         7469-LKBCI
gender                                 Male               Male               Male
SeniorCitizen                        0.1621             0.2069             0.0389
Partner                                  No                 No                Yes
Dependents                               No                 No                 No
tenure                              32.3711            31.2571            35.4374
PhoneService                            Yes                Yes                Yes
MultipleLines                            No                Yes                 No
InternetService                 Fiber optic        Fiber optic                 No
OnlineSecurity                           No                 No No internet service
OnlineBackup                             No                 No No internet service
DeviceProtection                         No                 No No internet service
TechSupport                              No                 No No internet service
StreamingTV                              No                 No No internet service
StreamingMovies                          No                 No No internet service
Contract                     Month-to-month     Month-to-month           Two year
PaperlessBilling                        Yes                Yes                 No
PaymentMethod              Electronic check   Electronic check       Mailed check
MonthlyCharges                      64.7617            76.8612            31.4605
TotalCharges                      2283.3004          2611.5491          1379.8735


Time taken to build model (full training data) : 0.05 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      5166 ( 73%)
1      1877 ( 27%)
```

Figure 39: Clustering into 2.

From the above, Cluster 5, gave a better breakdown of why a customer would churn based on their monthly subscriptions and the internet service they used. We can see that customers that had no internet service subscription did not experience many charges, hence it was easy for them to remain, while the model evaluated that those who subscribed for the FiberOptics, and the monthly charges were going to leave.

The deep learning model was used as well to determine how accurate their performance was based on the sentiment analysis review. Mujahid.,et al (2021), applied the Vader model in their work "Sentiment Analysis and Topic Modelling on Tweets about Online Education during covid-19, the score he had was 89%, Magukh et al (2021), used the Roberta model in predicting neutralities in offensive language identification dataset, and had a 90% score, while Rohan et al (2021), used the Aspect Based Model on restaurant and produced 75% score.

While working on this research, the Vader model, Roberta model, and Aspect Based Method were used to determine how well they performed. After the application of the models, the Roberta model had a better performance score than the Vader model and Aspect based sentiment Analysis.

While the Roberta and Vader Model were based on numbers, the Aspect Based Method is based on the classification of specific words. With the use of the spacy library, the spacy model, "en_encore_web_sm" was downloaded. This model has been trained with text and labelled data. It helps to identify distinct parts of speech that are assigned to the text to pull out vital information from the context of the words.

The significant feature of the ASBA is that it specifically helps draw out meanings from sentences based on the contest of the words. With the use of the token library and attention mechanism, the aspect-based analysis can break down text, and capture keywords that are important especially when the text is mixed with positive and negative words.

```
[{'aspect': '', 'description': 'costly'}, {'aspect': '', 'description': 'remote'}, {'aspect': '',
'description': 'Cheap'}, {'aspect': '', 'description': 'expensive'}, {'aspect': '', 'description':
'international'}, {'aspect': 'network', 'description': 'fast'}]
```

Figure 40 Aspect-based analysis when not adjusted

```
[{'aspect': 'food', 'description': 'delicious'}, {'aspect': 'time', 'description': 'very enjoyable'},
{'aspect': 'meal', 'description': 'tasty'}, {'aspect': 'internet', 'description': 'slow'}, {'aspect':
'experience', 'description': 'suboptimal'}]
```

Figure 41 Aspect-based analysis after adjustment

Without an adjustment in Figure 41Figure 26, there were some words that the model could not capture despite using several words which is seen as a limitation compared to the Vader and Roberta Model.

Analysis of the three machine learning models: Support Vector Machine, Random Forest, and Decision tree were on the Secondary data, and the strongest predictor was the Random Forest Algorithm with a performance of 87% for the F1Score, precision (81%), recall (94%), and Accuracy being (80%). Using the principal Component Analysis was to avoid any form of overfitting even though the scores were not high before applying the PCA. There was a slight difference in the score. Also, I believe that the low score due to the PCA could be linked to the data loss as there is a chance that the number of components I chose did not capture the complete dataset.

For the deep learning model, the Roberta model performed better than the Vader model and Aspect based Analysis. The Vader model takes all the words in the sentence and splits the words. It gives value to each word, specifically telling the user how positive, negative, or neutral the statement is based on the words. The limitation attached to Vader is that it does not account for relationships between words which is important when interacting with humans. Unlike the Vader model, the Roberta model picks on the context of words which is important when interacting with humans. The Roberta model had the best accuracy score by outperforming the Vader model. The goal of this research was to determine the best model that can predict human emotion better. The aim was not to look out for a 100% accuracy score but to source for a better score higher than 80%. For the Aspect based, its performance level was not as high as the Roberta fully as there were certain words that the model could not capture except done manually. Without manually adjusting the data there were some words that the model could not capture despite using several words which can be seen as a limitation.

CHAPTER FIVE

Conclusion

5.1. Recommendation

Mark Zuckerberg ones said “Nothing comes out fully formed. You have to keep iterating till you get to a point where you are comfortable that your system works as expected”. The essence of this research was to give insight to companies on ways to satisfy customers through the application of Artificial Intelligence. With what has been done in this research, there are more things that can be done in the future.

First, I recommend an expansion of the research by creating an app that would serve as a recommender system for organizations, displaying the percentage of customers that are satisfied, not satisfied, or likely going to leave based on the customer’s data for the week, top features why customers want to leave and why they are satisfied will also come up, which can help organizations able to specifically understand if they are doing something appropriate to keep their customers or if there are any adjustments that needs to be made.

Second, In the future, the plan is to use a larger dataset that would consist of different network providers and reviews coming from customers to maximize and bring out the statistical importance coming from the data.

Third, for the secondary data, to extract and get vital and effective variables affecting customers behavioral reasons for churning, the discriminant analysis techniques can be implemented because they are seen as a powerful technique for predicting different classes of customers.

Lastly, better models for predicting correctly can be used to understand the human language correctly when there are twists to text words.

5.2. Limitation

There were some challenges faced during this thesis:

First, the secondary data was only focused on one Network provider by the company, which cannot be used to draw conclusions for others network providers. Also, since the data released was owned by the company, there was a bit of adjustment made to the datasets. The Weka tool helped me realise that out of 50% of the customers using the DSL internet service plan, 30% were not comfortable with the plan, but still decided to remain with the company.

Second, there was an intention to use APIs (Application Programming Interfaces) like the trust pilot that has one of the largest reviews from customers on mobile network providers. However, based on the meeting I had with their data team, the reviews were for organizations to access their own reviews and improve on their organizations. I needed to have a registered business; it was not for educational purposes. The only way students could have access is to directly contact the owners of these reviews. Also, 90% of customers were based in Southampton and were students, which limited the scope of this project.

Another significant limitation was the unavailability of similar research topics that comprised the primary dataset obtained through surveys.

In addition, there was the limitation of providing large data from the survey. To add to this, on the survey that was conducted, an omission of the specific age of customers was made which led to a non-critical analysis of customers in different age groups.

Lastly, Time was a great constraint to this research which also led to boycotting various areas that would have made the research move to the next level.

5.3. Conclusion

Customer satisfaction is an integral component in a company. While retaining customers, companies have gone through several methods to retain their customers especially for customers that fall under the telecommunication sector. It is vital to note that being proactive will not only help in retaining customers but also serve as a threat to other companies. However, some organizations find it difficult to improve their AI techniques in such a way that they can be used to determine customers that are going to leave or remain. This research was to determine customers satisfaction and churn rate through the application of Artificial Intelligence techniques. It explored the strategies to retain customers by surveying the causes of customer dissatisfaction. The significance of this research is to help companies in the telecom world to grow and make profit while satisfying their customers. Predicting customers satisfaction and recognizing the reasons why they want to leave is very vital. Hence, this research went deeper by building a user interface that can determine customer's level of satisfaction through analysing reviews coming from different people. As explained in the body of this work, different software tools for data visualisations were used. The PowerBi tool did great in telling stories about the data while the Weka software worked well with segmenting the customers. This work was also able to find suitable patterns using deep learning models in drawing conclusions on the best technique that organizations can use. This pattern also helped to gain better insights into the limitations and recommendations for future research.

REFERENCES

1) Abdelrahim K., A., et al., 2019. Customer Churn Prediction in telecom using machine learning in big data platform, faculty of Information Technology, highest Institute for Applied Sciences and Technology, [viewed 2 April 2022] pp.9-20.
2) Abdulkadir. H., and Omer Farouk. S., 2014. Modelling customer churn using segmentation and data Mining . Accessed in Home Feed | ResearchGate . viewed 19 July 2022. Pp.2-5
3) Adnan., A. et al., 2017. Customer Churn Prediction in the telecommunication sector using a rough set approach, Neurocomputing, vol 237, [Accessed in Customer churn prediction in the telecommunication sector using a rough set approach - ScienceDirect [viewed 3 August 2022]. pp.242-254.
4) Ammar., A.Q. et al., 2017. Churn Prediction on a Huge Telecom Data using hybrid firefly based classification", Egyptian Informatics journal 18, [viewed July 1 2022], pp. 215-220.
5) Anupreet Kaur Mokha, Pushpender Kumar. (2022) Examining the Interconnections Between E-CRM, Customer Experience, Customer Satisfaction and Customer Loyalty. Journal of Electronic Commerce in Organizations 20:1,[viewed 31 July 2022] pages 1-21.
6) Baboo Kureemum and Robert Fantina(2011), Your customer'w percerption of quality: What it means to your bottom line and how to control it in ProQuest Ebook Central - Reader , CRC press, Taylor and Francis Group, 6000 broken sound parkway NW., [viewed 2 August 2022].
7) Baier, D., et al, M. 2018. *Conversational user interfaces for online shops? A categorization of use cases*. Paper presented at the 39th International Conference on Information Systems (ICIS), San Francisco, USA in *Thomas M. Brill(2019), Siri, Alexa, and other digital assistants: a study of customer satisfaction with artificial intelligence applications* , Journal of Marketing Management, Vol 35,2019- Issue 15-16: [Accessed in Citations: Siri, Alexa, and other digital assistants: a study of customer satisfaction with artificial intelligence applications (tandfonline.com) [viewed 1 August 2022] .
8) Bhatt, A. 2014. Consumer attitude towards online shopping in selected regions of Gujarat. in *Thomas M. Brill(2019), Siri, Alexa, and other digital assistants: a study of customer satisfaction with artificial intelligence applications* , Journal of Marketing Management, Vol 35,2019- Issue 15-16: [Accessed in Citations: Siri, Alexa, and other digital assistants: a study of customer satisfaction with artificial intelligence applications (tandfonline.com) [viewed 2August 2022].
9) Bittner, E., Oeste-Reiß, S., & Leimeister, J. M. 2019. *Where is the bot in our team? Toward a taxonomy of design option combinations for conversational agents in*

collaborative work. Paper presented at the Proceedings of the 52nd Hawaii International Conference on System Sciences, Maui, Hawaii, USA.[Accessed in Citations: Siri, Alexa, and other digital assistants: a study of customer satisfaction with artificial intelligence applications (tandfonline.com) [viewed 1 August 2022].

10) Brandusoiu I, Toderean G, Ha B. 2016. Methods for churn prediction in the prepaid mobile telecommunications industry. In: International conference on communications. [viewed 2April 2022] p. 97–100.
11) Bruce. C., et al ., 2007 Approaches to Customer Segmentation in Journal of Relationship Marketing, published by Taylor& Francis,.[viewed 30 July 2022]. Pp.6-15.
12) Chartered Management Institute, 2008, "Origin of Customer Relationship Management" 1SO 9001-2015 Certified Education Provider.
13) E.B. Khedkar 2015 Effect of Customer Relationship Management on Customer Satisfaction and Loyalty in International Journal of Management , Vol 6. [Viewed 30 July 2022]. Pp. 2-4.
14) El Naqa, I., Murphy, M.J. 2015. What Is Machine Learning?. In: El Naqa, I., Li, R., Murphy, M. (eds) Machine Learning in Radiation Oncology. Springer, Cham. https://doi.org/10.1007/978-3-319-18305-3_1 [viewed 10 June 2022]. Pp. 3- 11. Accessed in What Is Machine Learning? | SpringerLink
15) Emilia Huong Xuan Nguyen, 2011. "Customer Churn Prediction for Icelandi Mobile Telephoney Market", Haskolaprent, falkagate 2, R, 107, in *Churn Predcition in Telecommunication using logistic Regression and Logot Boost.* [Accessed 1 July 2022], p.102.
16) Engel, J ., F. et al., 1995. Customer Behavior, 8th Edition, *Fort Worth: Dryden Press, TX, USA*,[viewed 3rd June, 2022]. Available in *Modelling customer churn using segmentation and data mining.*
17) Essam. A.,K., et al 2020. Customer Churn Prediction Model and identifying Features to increase Customer Retention based on User Generated Content. [Accessed in https://www.researchgate.net/publication/351743611_Customer_Churn_Prediction_Model_and_Identifying_Features_to_Increase_Customer_Retention_based_on_User_Generated_Content [viewed 2 April 2022], pp. 40-47.
18) F. Wangenheim, et al , 2017. Renew or Cancel? Drivers of customer renewal decisions for IT based service contracts, Journal of business Research, vol.79. pp. 181-188.
19) G.A Churchill Jr., C Surprenant (1982) **An investigation into the determinants of customer satisfaction** in Journal of Marketing Research, 19 (1982), pp. 491-504. Available from Google Scholar. viewed 18 July 2022.
20) Garcia , D.,L., et al., 2017. Intelligent Data Analysis Approaches to Churn a Business problem: A survey. Knowledge and Information Systems , 51(3), pp. 719- 774. [Accessed in Intelligent data analysis approaches to churn as a business problem: a survey - ProQuest [viewed 1 August 2022].
21) Gerpott.,T.J. et al 2015, "Who is (not) convinced to withdraw a contract termination announcement?- a discriminant analysis of mobile communications customers in Germany", Telecomunication Policy vol 39, No1, pp.38-52. [viewed 25 july 2022].

22) Global Graduate Network—Research Methods Handbook Vol. 10, [Viewed 6 July 2020]. Pp. 66-70
23) Goldstein, Sheldon D., 2009. Superior Customer Satisfaction and Loyalty : Engaging Customers to Drive Performance, ASQ Quality Press, ProQuest Ebook Central, [Accessed in http://ebookcentral.proquest.com/lib/solent-ebooks/detail.action?docID=3002665. Created from solent-ebooks on 2022-07-15 14:37:31 pp1-3.
24) Grigoroudis, E., Siskos, Y., 2002 . A survey of customer's satisfaction barometer: some results from the transportation- communication sector in European Journal of Operational Research Vol. 152, issue 2. Accessed in A survey of customer satisfaction barometers: Some results from the transportation-communications sector - ScienceDirect [viewed 19 July 2022]. Pp. 334-353.
25) Hemlata Jaun., et al, 2019. Churn Prediction in Telecommunication using Logistics Regression and Logit Boost" *in International Conference on Computational Intelligence and Data Science, School of Basic and Applied Sciences, Poornima University Jaipur- India. Published by, Elsevier.B.V.*
26) He .Y., et al., 2009 A study on prediction of customer churn in fxed communication network based on data mining. In: Sixth international conference on fuzzy systems and knowledge discovery, vol. 1. 20 [viewed 2 April 2022]. pp. 92–94.
27) Hiziroglu, and Omer F.,S., 2014. Modelling Customer Churn Segmentation and data mining, Frontiers in Artificial Intelligence and Applications 270:259-271 [Accessed in researchgate.net/publication/281655061_Modelling_customer_churn_using_segmentation_and_data_mining/stats [Viewed 3April 2022].
28) I.B. Brandusoui and Gavril Tuderean, 2013. Predicting Churn in Mobile Telecommunications Industry, Technical University of cluj-Napoca. Available from: PREDICTING CHURN IN MOBILE TELECOMMUNICATIONS INDUSTRY - ProQuest [viewed 30 July 2022].
29) Idris A., et al. 2012 Genetic programming and adaboosting based churn prediction for telecom. In: IEEE international conference on systems, man, and cybernetics (SMC). [Viewed 3 April 2022] pp. 1328–32.
30) Irfan, U., et al, 2019. A churn prediction model using Random forest: Analysis of Machine learning Technique for churn prediction and factor identification in the Telecom Sector.[viewed 6 June 2022]. Available from: (PDF) A Churn Prediction Model Using Random Forest: Analysis of Machine Learning Techniques for Churn Prediction and Factor Identification in Telecom Sector (researchgate.net). pp. 10-15.
31) J. SU. Y.Yang, T. Yang. Measuring Knowledge diffusion efficiency in R&D networks: Knowledge Management Research and practice, vol.16, no 2. [viewed 25july 2022]. Pp.208-219
32) J.Creswell, V. Piano. 2011. Designing and Conducting mixed method Research. Sage Publications. [Accessed in Designing and Conducting Mixed Methods Research -

2007 - Australian and New Zealand Journal of Public Health - Wiley Online Library[Viewed 6 August 2022] pp. 40-75.

33) Kahlil P., et al., 2016. "Twitter Sentiment Analysis : Capturing Sentiment from Integrated resort tweets in [Accessed in International Journal of Hospitality Management | ScienceDirect.com by Elsevier Vol 55. [Viewed 7 June 2022]. Pp. 16-24.
34) Kandampully,J and Subartanto,D. 2003. The role of customer satisfaction and image in gaining customer loyalty in *the hotel industry. Journal of Hospitality and Leisure Marketing*, 10(1,2), 3-25.
35) Kandampully, J. and Suhartanto, D., 2013. The role of customer satisfaction and image in gaining customer loyalty in the hotel industry. "*Journal of Hospitality and leisure Marketing* " Available from Home Feed | ResearchGate [Accesseed 19 July 2022]. Pp,5-12.
36) Koehler, J.(2016) Business process innovation with artificial intelligence- benefits and operational risks. European Business and Management, 4(2), pp.55-66. [[Accessed in Citations: Siri, Alexa, and other digital assistants: a study of customer satisfaction with artificial intelligence applications (tandfonline.com)
37) Kolis, K., & Jirinova .,K (2013) The impact of the Customer Relationship Management on a Company's Financial Performance . Proceedings of the European Conference on Management, Leadership& Governance.[viewed 30 July 2022] p.33. 527.
38) Kortler.P., K.L., Keller, 2006. Marketing Management , 12th edition, Precentise Hall, Upper Saddle River , New York Kotler: Marketing management: A south Asian perspective - Google Scholar [viewed 20 July] 2022]pp. 3-45.
39) Kothari, C.R., 2014. Research Methodology: Methods and Techniques. 2^{nd} Edition, New Age International Publishers, New Delhi, Vol4, No3. [viewed 20 June 2022]. pp. 30-35.
40) Li Chen. C, et al., 2019. Behaviour Analysis of Customer Churn for a Customer Relationship System: An Empirical Case Study in Journal of Global information Management [Accessed in : https://www.researchgate.net/publication/330047401 [viewed 5 July 2022]. Pp. 3-39.
41) Lubaba R., et al., 2022. Analysis of Dimensionality Reduction Techniques on Internet of Things Data Using Machine Learning, in [Accessed in Analysis of dimensionality reduction techniques on Internet of Things data using machine learning - ScienceDirect, Vol 52. Part D, https://doi.org/10.1016/j.seta.2022.102304 [Viewed 18 August 2022]. Pp.10-20

42) M., Holmlund, and S., Kock 1996. Relationship Marketing: The importance of customer-perceived service quality in retail banking in Kandampully J., and Suhartanto "*The role of customer satisfaction and image in gaining customer loyalty in the hotel industry*". Available from Home Feed | ResearchGate [Accessed 19 July 2022]. Pp. 52-53.

43) Mark R.M., et al., 2015. Analytics Defined in[Accessed in Information Security Analytics | ScienceDirect [Viewed 7 June 2022]. Pp. 1-12.
44) Martins Gonçalves, H. and Sampaio, P. (2012), "The customer satisfaction-customer loyalty relationship: Reassessing customer and relational characteristics moderating effects", Management Decision, Vol. 50 No. 9, pp. 1509-1526. https://doi.org/10.1108/00251741211266660
45) Marr, Bernard. Artificial Intelligence in Practice: How 50 Successful Companies Used AI and Machine Learning to Solve Problems, John Wiley& Sons, Incorporated , 2019. ProQuest Ebook Central, Accessed in ProQuest Ebook Central - Detail page pp. 3- 323.
46) Mayukh. S., et al., 2021. Comparison of Neutrosophic Approach to Various Deep Learning Models For Sentiment Analysis, in[Accessed in Comparison of neutrosophic approach to various deep learning models for sentiment analysis - ScienceDirectVol 223, https://doi.org/10.1016/j.knosys.2021.107058 [viewed 19 August 2022]. Pp 22- 30.
47) Mayukh S., et al., 2021. Deep Learning for predicting neutralities in Offensive Language Identification Dataset [Accessed in Deep Learning for predicting neutralities in Offensive Language Identification Dataset - ScienceDirect vol185, [viewed 30 December 2021] pp. 6-9.

48) McCann, J.M and Gallagher., J.P., 1990. Expert systems for a scanner data environment : The marketing workbench experience , kluwe Academic Publisher, Nowell, MA, USA, Available from: Home Feed | ResearchGate [viewed 20 July 2022].p30.
49) Michael, B., and D.V. Poel, 2012. Customer event history for churn prediction: How long is long enough, Expert Syst. Appl Journal, Vol. 39, [Accessed in [PDF] Customer event history for churn prediction: How long is long enough? | Semantic Scholar [viewed 17 August 2022].pp45-47.
50) Mohammed. H., et al., 2019. Churn Prediction : A Comparative Study Using KNN and Decision Trees , [Accessed in : https://www.researchgate.net/publication/340890936 [viewed 5 July 2022].
51) Moshan, F., et al.. 2011 "Impact of Customer Satisfaction on Customer Loyalty and Intentions of Switch: Evidence from Banking Sector of Pakistan": International Journal of Business and Social Science, in Paul Guruwo(2015) Impact of Customer Satisfaction on Customer Loyalty in the Banking.p.90.
52) Mustafa, N., et al., 2021 Customer churn prediction for telecommunication industry: A Malaysian Case Study [version 1; peer review: 2approved]. *F1000Research* 2021, 10:1274 (https://doi.org/10.12688/f1000research.73597.1). [viewed 5 July 2022].
53) Nguyen, N., et al., 2022. Churn Prediction in telecommunication industry using kernel support vector machines, [Accessed in https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0267935] [viewed 4 August 2022. Pp.10-13.

54) Nhi, N., Y., et al., 2021. Leveraging unstructured call log data for customer churn prediction, Nhi N.Y. Vo, Shaowu Liu, Xitong Li, Guandong Xu, Volume 212,
55) [Accessed Leveraging unstructured call log data for customer churn prediction - ScienceDirect [Viewed 2 August 2022] . pp. 43-40.
56) Nils J. W., 1998. Artificial Intelligence: A New Synthesis in Science Direct in Artificial Intelligence: A New Synthesis | ScienceDirect, ISBN [viewed 4 August 2022]. pp. 3-5.
57) Nilson, Nils. J., 2010, The quest for Artificial Intelligence: A history of ideas and achievement. Cambridge University Press: New York, USA,[viewed 10 June 2022], pp. 3-30.
58) OGuerra- Santin et al., 2016. Mixed Method Approach to determine Occupants behaviour Analysis of two case studies. Vol. 130,[Accessed in Mixed methods approach to determine occupants' behaviour – Analysis of two case studies - ScienceDirect [Viewed 5 August 2022] pp.546-566.
59) Oliver, R. L., 1980. A cognitive model of the antecedents and consequences of satisfaction decisions. Journal of Marketing Research, 17(4), https://doi.org/10.1177%2F002224378001700405 in Ishani Patharia and Anjana Pandey "A Systematic Literature Review on Factors Affecting Customer's Loyalty Towards Mobile Network Service Providers".[viewed 5 August 2022]. Pp. 460-469.
60) Pang and lee., 2008. Opinion mining and sentiment analysis in "Twitter Sentiment Analysis: Capturing Sentiment from Integrated resort tweets in International Journal of Hospitality Management | ScienceDirect.com by Elsevier Vol 55. [Accessed 7 June 2022]. Pp. 1-135.
61) Payne, A., & Frow.P.(2016) Customer Relationship Management: from Strategy to implementation. Journal of Marketing Management,[viewed 30 July 2022] pp.135-168.
62) Phillip, K., 2002, Marketing in the twenty-first century in *A framework for marketing management*, Tenth Edition., [Accessed 5 June 2022].pp.14-15.
63) Pooja. M., et al., 2020. Customer Churn Prediction for Fintech Companies Using Artificial Neural Networks in i-manager's Journal on Computer Science, Vol 7, [viewed 4 july 2022]. Pp. 10- 20.
64) Quang Li , 2022. Designing of Customer Churn Early Warning System based on Mobile Communication Technology Based on Data Mining in *Journal of Electrical and Computer Engineering ,Volume 2022, Article ID 9701349,* [Accessed in https://www.hindawi.com/journals/jece/2022/9701349/ [viewed 4 August 2022]. Pp. 3-19.
65) R.L Oliver. 1980. **A cognitive model of the antecedents and consequences of satisfaction decisions** in Journal of Marketing Research, 17. Accessed in Google Scholar. Viewed 18 july 2022).pp. pp. 460-469.
66) R.L Oliver 1996 **Satisfaction: A Behavioral Perspective on the Customer**
67) McGraw-Hill, New York. [Accessed in Google Scholar. viewed 19 July 2022)
68) Rizwan Ullah., K., et al., 2020. The impact of Customer Relationship Management and Company Reputation on Customer Loyalty: The Mediating Role of Customer Satisfaction. [Accessed in The Impact of Customer Relationship Management and

Company Reputation on Customer Loyalty: The Mediating Role of Customer Satisfaction: Journal of Relationship Marketing: Vol 21, No 1 (tandfonline.com) pp.1-26.

69) Rohan K., et al., 2021. Position Aspect Based sSentiment Analysis Using Attention Mechanism[Accessed in Positionless aspect based sentiment analysis using attention mechanism - ScienceDirect
70) Vol 226, https://doi.org/10.1016/j.knosys.2021.107136 [Viewed 80 August 2022], pp.18-22.
71) Ryals, L. (2005). Making Customer Relationship Management Work: The Measurement and Profitable Management of Customer Relationships. Journal of Marketing, 69(4), [Accessed 30 July 2022] pp.252-261.
72) Samira.K.,and M., Zivari-Rahman, 2017. Comparison of supervised machine learning techniques for customer churn prediction based on analysis of customer behaviour., Vol.19.[Accessed in (1) (PDF) Comparison of supervised machine learning techniques for customer churn prediction based on analysis of customer behavior (researchgate.net) [viewed 15 August 2022].pp. 20-26.
73) Sarah cook., 2002. The essential guide to employee engagement :Business performance throughnstaff satisfaction,. Available from,www.loganpage.com, London N19JN, United Kingdom,
74) Tianyuan, Z., et al, 2022. A Data -Driveb Approach to improve customer churn prediction based in Telecom customer segmentation. Available from MDPI - Publisher of Open Access Journals [viewed 19 July 2022].pp.3-6.
75) T.Schwartz., et al.,2014. What People do with Consumption feedback: A long-term living tab study of a home energy management system [Accessed in (1) (PDF) What People Do with Consumption Feedback: A Long-Term Living Lab Study of a Home Energy Management System (researchgate.net)
76) Takeuchi. H. A .,1983. Quality Is More Than Making a Good Product. Havard Business Review *in The Impact of Service Quality, Customer Satisfaction and Loyalty Programs on Customer's Loyalty: An Evidence from Telecommunication Sector. 2013. Journal of Asian Business Strategy*, 3(11),[viewed1 August 2022]. pp. 306-314.
77) Thomas M. Brill.,2019., Siri, Alexa, and other digital assistants: a study of customer satisfaction with artificial intelligence applications , Journal of Marketing Management, Vol 35,2019- Issue 15-16: [Accessed in Citations: Siri, Alexa, and other digital assistants: a study of customer satisfaction with artificial intelligence applications (tandfonline.com) [viewed 2August 2022]. Pp. 32.
78) Tianyuan, Z., et al, 2022. A Data -Driveb Approach to improve customer churn prediction based in Telecom customer segmentation. Available from MDPI - Publisher of Open Access Journals [viewed 19 July 2022].pp.3-6.
79) Vavra, T.G., 1997. Improving Your Measurement of Customer Satisfaction. ASQC Quality Press, Milwaukee. Accesses in Google Scholar [viewed 19 july 2022].pp. 23-26.
80) Verbeke, W., et al., 2012. New Insight into Churn Prediction in the telecommunication sector: A Profit-driven data mining approach, European Journal of Operational Research 218(I), [viewed 8 June 2022].pp. 211-229.

81) Venkatesh, V., Thong, J. Y. L., Chan, F. K. Y., Hu, P. J. H., & Brown, S. A. (2011). Extending the two-stage information systems continuance model: Incorporating UTAUT predictors and the role of context. *Information Systems Journal*, 21(6), 527–555. [Accessed in Extending the two-stage information systems continuance model: incorporating UTAUT predictors and the role of context - Venkatesh - 2011 - Information Systems Journal - Wiley Online Library [viewed 1 August 2022]. Pp.1-10.
82) Wedel., M, Kamakura, W.A.,(2000) Market segmentation: Conceptual and methodological foundations, vol.2, Kluwer Academic Publisher Dordrecht. [viewed 25 July 2002]pp.20-24.
83) Zhou S., et al., 2018. Measuring customer agility from online reviews using big data text analytics J. Manage. Inf. Syst., 35 (2) (2018), [Accessed in Measuring Customer Agility from Online Reviews Using Big Data Text Analytics: Journal of Management Information Systems: Vol 35, No 2 (tandfonline.com) [Viewed 2 August 2022] pp. 510-539.

APPENDIX C

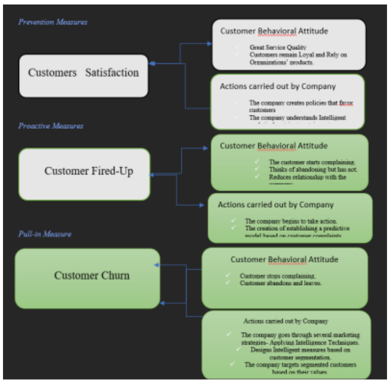


fig1

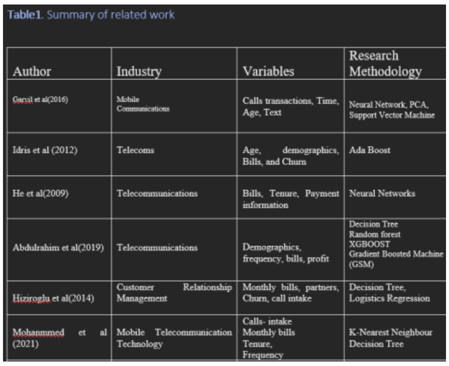
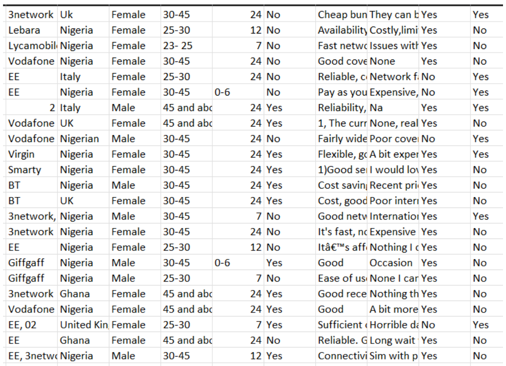


Table1. Summary of related work

Author	Industry	Variables	Research Methodology
Garsil et al(2016)	Mobile Communications	Calls transactions, Time, Age, Text	Neural Network, PCA, Support Vector Machine
Idris et al (2012)	Telecoms	Age, demographics, Bills, and Churn	Ada Boost
He et al(2009)	Telecommunications	Bills, Tenure, Payment information	Neural Networks
Abdulrahim et al(2019)	Telecommunications	Demographics, frequency, bills, profit	Decision Tree Random forest XGBOOST Gradient Boosted Machine (GSM)
Hiziroglu et al(2014)	Customer Relationship Management	Monthly bills, partners, Churn, call intake	Decision Tree, Logistics Regression
Mohammed et al (2021)	Mobile Telecommunication Technology	Calls- intake Monthly bills Tenure, Frequency	K-Nearest Neighbour Decision Tree

fig2



3network	Uk	Female	30-45		24	No	Cheap bun	They can b	Yes	Yes
Lebara	Nigeria	Female	25-30		12	No	Availability	Costly,limi	Yes	No
Lycamobil	Nigeria	Female	23- 25		7	No	Fast netwo	Issues witl	Yes	No
Vodafone	Nigeria	Female	30-45		24	No	Good cove	None	Yes	No
EE	Italy	Female	25-30		24	No	Reliable, c	Network f	No	Yes
EE	Nigeria	Female	30-45	0-6		No	Pay as you	Expensive,	No	Yes
2	Italy	Male	45 and abc		24	Yes	Reliability,	Na	Yes	Yes
Vodafone	UK	Female	45 and abc		24	Yes	1, The curr	None, real	Yes	No
Vodafone	Nigerian	Male	30-45		24	No	Fairly wide	Poor cove	No	Yes
Virgin	Nigeria	Female	30-45		24	Yes	Flexible, gc	A bit exper	Yes	Yes
Smarty	Nigeria	Female	30-45		24	Yes	1)Good se	I would lov	Yes	No
BT	Nigeria	Male	30-45		24	Yes	Cost savin	Recent pri	Yes	No
BT	UK	Female	30-45		24	Yes	Cost, good	Poor interi	Yes	No
3network,	Nigeria	Male	30-45		7	No	Good netw	Internatioi	Yes	Yes
3network	Nigeria	Female	30-45		24	No	It's fast, nc	Expensive	Yes	No
EE	Nigeria	Female	25-30		12	No	Itâ€™s affi	Nothing I c	Yes	No
Giffgaff	Nigeria	Male	30-45	0-6		Yes	Good	Occasion	Yes	No
Giffgaff	Nigeria	Male	25-30		7	No	Ease of us	None I car	Yes	No
3network	Ghana	Female	45 and abc		24	Yes	Good rece	Nothing th	Yes	No
Vodafone	Nigeria	Female	45 and abc		24	Yes	Good	A bit more	Yes	No
EE, 02	United Kin	Female	25-30		7	Yes	Sufficient	Horrible da	No	Yes
EE	Ghana	Female	45 and abc		24	No	Reliable. G	Long wait	Yes	No
EE, 3netw	Nigeria	Male	30-45		12	Yes	Connectivi	Sim with p	Yes	No

Fig3

Network S	Review
EE	Connectivity, network strength, internet
Giffgaff	Cheap bundles
EE	Easy to contact support. Easy to use app. Fair prices
EE	Gifting data to family
EE	Cost , sometimes singal goes off
Giffgaff	Network signal not everywhere
EE	Lack of signal sometimes.
EE	High priced contracts, not enough network coverage in remote areas, comes across as
Giffgaff	Affordable price, easy connections, Good app interface
Giffgaff	Reliable
Lebara	Bonuses, WiFi availability, cost
Telco	Cheap subscription, network speed and network coverage
Telco	I got better deal, good customer satisfaction
Lebara	Steady, fast and reliable
EE	Speed, coverage
Lebara	Coverage, strong network,
Lebara	No credit check necessary, decent cost, ideal for international calls
Telco	Reliable, Cheap and wide coverage
Giffgaff	None
Giffgaff	No 5g
Lebara	International bundle cost
Telco	Network stability, distance to recharge, broadband availability in my area.
Telco	Poor service in some area
Lebara	Non
EE	Nothing
Lebara	High data tariff,
Lebara	Long Wait time to reach customer service

fig4

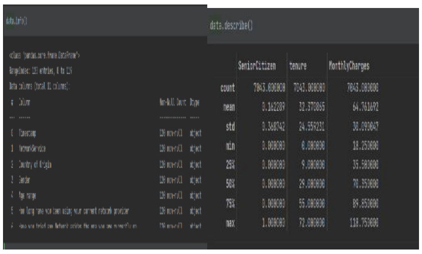
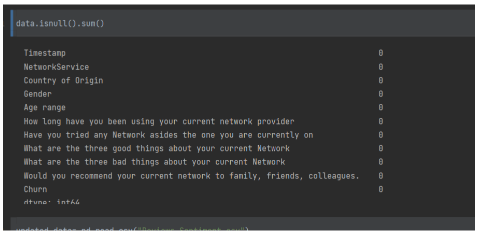


Fig5



```
data.isnull().sum()

Timestamp                                                    0
NetworkService                                               0
Country of Origin                                            0
Gender                                                       0
Age range                                                    0
How long have you been using your current network provider   0
Have you tried any Network asides the one you are currently on 0
What are the three good things about your current Network    0
What are the three bad things about your current Network     0
Would you recommend your current network to family, friends, colleagues. 0
Churn                                                        0
dtype: int64
```

fig6

```
#checking for missing values
data.isnull().sum()

customerID          0
gender              0
SeniorCitizen       0
Partner             0
Dependents          0
tenure              0
PhoneService        0
MultipleLines       0
InternetService     0
OnlineSecurity      0
OnlineBackup        0
DeviceProtection    0
```

Fig7

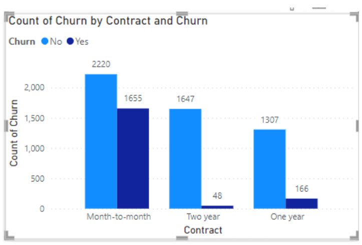


fig8

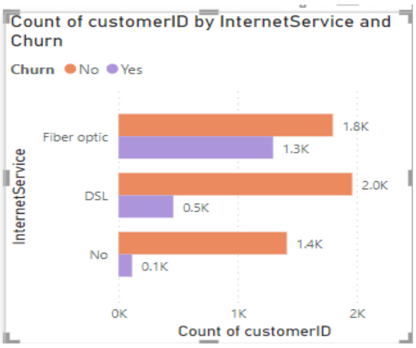


Fig9

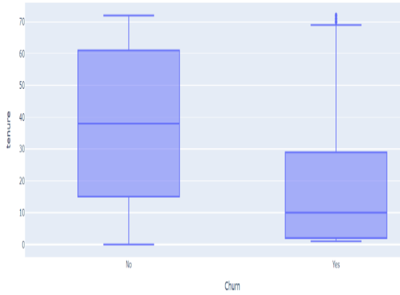


fig 10

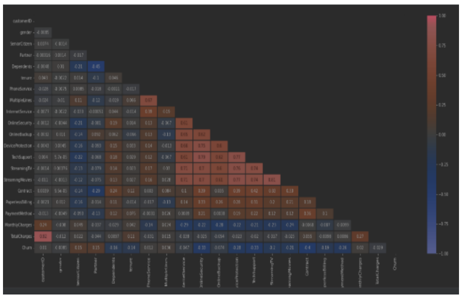


Fig 11

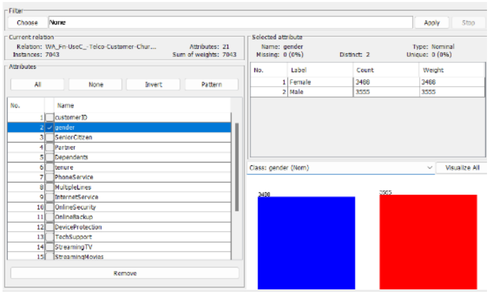


fig12

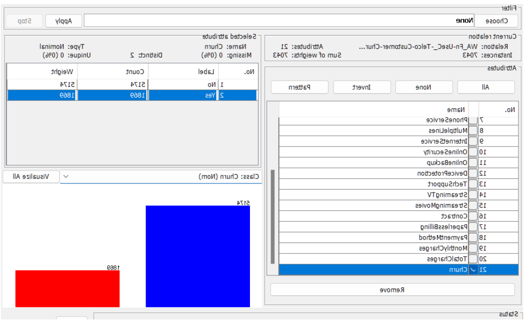


Fig 13

fig 14

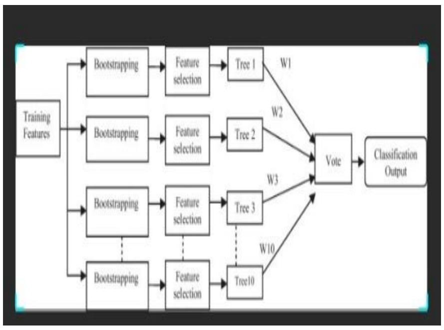


Fig15

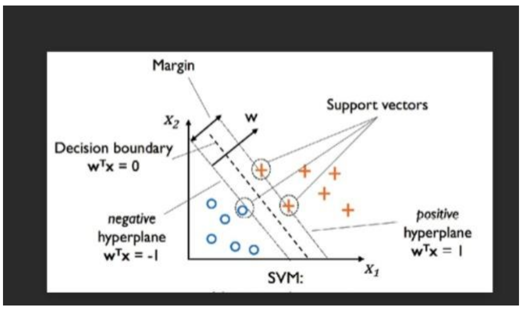


fig16

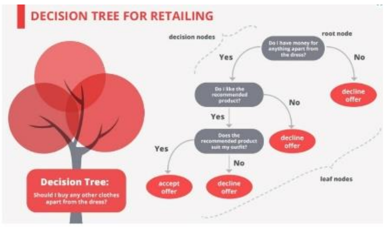


Fig17

```
components = None
pca = PCA(n_components=components)
pca.fit(scaled_x_train)
print('Variances(Percentage):')
print(pca.explained_variance_ratio_ * 100)
```

fig 18

```
scaler = preprocessing.MinMaxScaler(feature_range=(0,1))
scaled_x_train = np.array(X_train).reshape(len(X_train),45)
scaled_x_train = scaler.fit_transform(scaled_x_train)
scaled_x_test = np.array(X_test).reshape(len(X_test),45)
scaled_x_test = scaler.transform(scaled_x_test)
```

Fig19

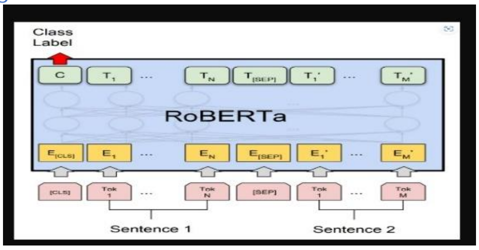


fig20

```
import nltk
nltk.download('vader_lexicon')
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
sia= SentimentIntensityAnalyzer()
```

Fig21

```
aspects = []
for sentence in reviews:
    doc = nlp(sentence)
    descriptive_term = ''
    target = ''
    for token in doc:
        if token.dep_ == 'nsubj' and token.pos_ == 'NOUN':
            target = token.text
        if token.pos_ == 'ADJ':
            prepend = ''
```

fig22

```
#each of the words have been given their part of speech
tokens= nltk.word_tokenize(words)
tagged= nltk.pos_tag(tokens)
tagged[:10]

[('I', 'PRP'),
 ('got', 'VBD'),
 ('better', 'RBR'),
 ('deal', 'NN'),
 (',', ','),
 ('good', 'JJ'),
 ('customer', 'NN'),
 ('satisfaction', 'NN')]
```

Fig23

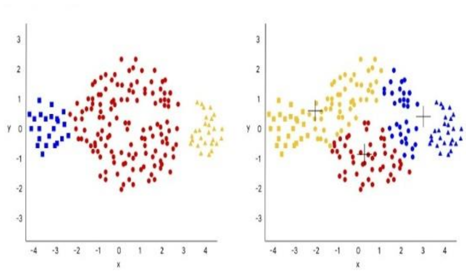


fig24

```
input_text = st.text_input("Get Predictions", value="", max_chars=None, key=None, type="defaul
if input_text:
    reason: list = []
    key_list = reasons.keys()
    for key in key_list:
        for k in reasons[key]:
            if input_text.__contains__(k):
                if reason.__contains__(key):
                    continue
                else:
                    reason.append(key)
    for n in networks:
        if input_text.__contains__(n):
            network = n
```

Fig25

The vodafone network is lovely, because it is afforadable


```
[
  0 : {
    "label" : "POSITIVE"
    "score" : 0.9998268485069275
    "Network" : "vodafone"
    "Reason" : [
      0 : "good"
    ]
  }
]
```


fig26

```python
porter = PorterStemmer()
def tokenizer(words):#breaking down the text with the tokenizer
    return words.split()

def tokenizer_porter(words):
    return[porter.stem(words)for words in words.split()]
# tokens= nltk.word_tokenize(words)
```

Fig27

```python
#seeing how well the sentiment are distributed
words= ' '.join( [reviews for reviews in updated_data['Reviews']])
word_cloud = WordCloud(width = 500, height=300, random_state =21, max_font_size= 119).generate(words)
plt.imshow(word_cloud, interpolation = 'bilinear')
plt.axis('off')
plt.show()
```

fig28

```python
class_report = classification_report(y_test, rfc_y_prediction)
print('Report of Classification;\n', class_report)
```

Report of Classification;

	precision	recall	f1-score	support
0	0.81	0.94	0.87	1555
1	0.70	0.40	0.51	555
accuracy			0.80	2110
macro avg	0.76	0.67	0.69	2110
weighted avg	0.78	0.80	0.78	2110

Fig29

```python
confusion_matrix(y_test, rfc_y_prediction)
class_report = classification_report(y_test, rfc_y_prediction)
print('Report of Classification;\n', class_report)
```

Report of Classification;

	precision	recall	f1-score	support
0	0.78	0.96	0.86	1555
1	0.67	0.25	0.37	555
accuracy			0.77	2110
macro avg	0.73	0.60	0.61	2110
weighted avg	0.75	0.77	0.73	2110

fig30

Report of Classification;

	precision	recall	f1-score	support
0	0.84	0.90	0.87	1555
1	0.64	0.51	0.57	555
accuracy			0.80	2110
macro avg	0.74	0.70	0.72	2110
weighted avg	0.79	0.80	0.79	2110

Fig31

```
[0 1 0 ... 0 1 0]

77.77251184834122
```

```python
class_report = classification_report(y_test, svc_y_prediction)
print('Report of Classification;\n', class_report)
```

Report of Classification;

	precision	recall	f1-score	support
0	0.81	0.91	0.86	1555
1	0.62	0.40	0.49	555
accuracy			0.78	2110
macro avg	0.72	0.66	0.67	2110
weighted avg	0.76	0.78	0.76	2110

fig32

```python
class_report = classification_report(y_test, dt_y_prediction)
print('Report of Classification;\n', class_report)
```

Report of Classification;

	precision	recall	f1-score	support
0	0.83	0.91	0.87	1555
1	0.66	0.46	0.54	555
accuracy			0.79	2110
macro avg	0.74	0.69	0.71	2110
weighted avg	0.78	0.79	0.78	2110

Fig33

```python
class_report = classification_report(y_test, dt_y_prediction)
print('Report of Classification;\n', class_report)
```

Report of Classification;

	precision	recall	f1-score	support
0	0.83	0.91	0.87	1555
1	0.66	0.46	0.54	555
accuracy			0.79	2110
macro avg	0.74	0.69	0.71	2110
weighted avg	0.78	0.79	0.78	2110

fig34

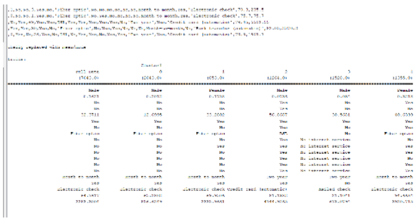


Fig36

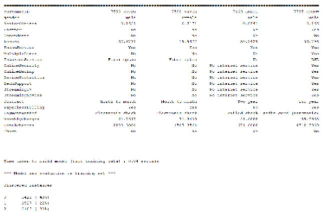


fig37

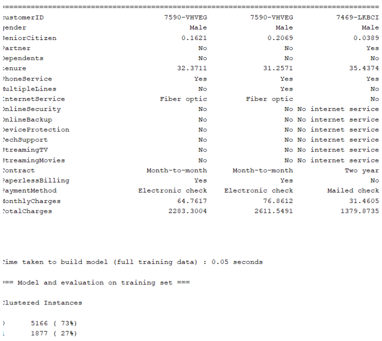


Fig38

fig39

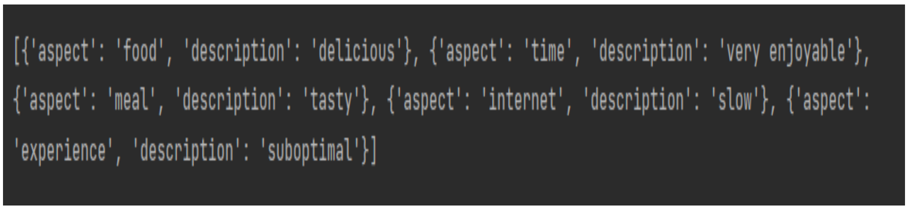


Fig41

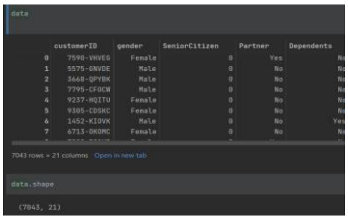


Figure 42: secondary data

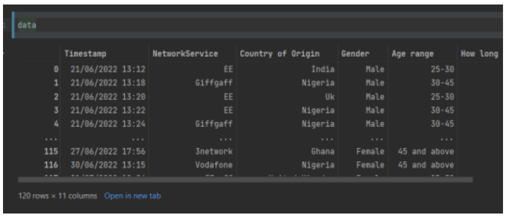


Figure 43: display of the rows and columns in the primary data

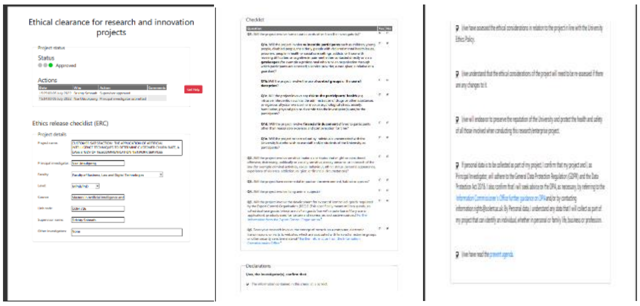


Figure 44: Ethics Form

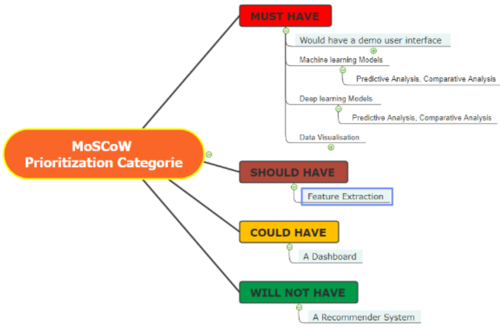


Figure 45: Moscow Prioritization Template

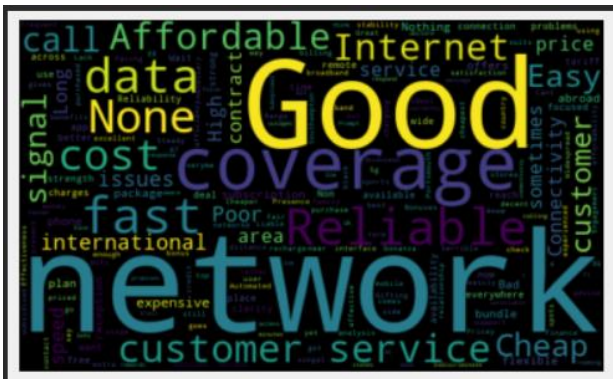


Figure 46: Display word cloud