A HYBRID MACHINE LEARNING PROCESS FRAMEWORK FOR DATA-DRIVEN DIRECT PRODUCT MARKETING

by

RUSSEL MELVIN PETERSEN
Student Number: 203132173

Thesis submitted in fulfilment of the requirements for the degree

Master: Information Technology

in the Faculty of Informatics and Design

at the Cape Peninsula University of Technology

Supervisor: Professor Justine Olawande Daramola

District Six
December 2020

DECLARATION

I, Russel Melvin Petersen, declare that the contents of this thesis represent my own unaided work, and that the thesis has not previously been submitted for academic examination towards any qualification. Furthermore, it represents my own opinions and not necessarily those of the Cape Peninsula University of Technology.

	14-Dec-2020
Signed	**Date**

ABSTRACT

The ability to propose the right product, to the right customer, at the right time is valuable in marketing. Direct marketing strategy entails the analysis of customer patterns to predict the best product offerings that are relevant to a customer per time. Machine learning (ML) has proven to be an improved, more effective approach for direct marketing compared to using more traditional statistical methodologies. However, the disadvantage of ML models is that they normally perform in a black box manner, where the predictions are produced without an explanation. In many cases an expert is required to interpret the outputted result. Studies have shown that this lack of understanding and explanation leads to reduction in the user's confidence in the results produced. This situation has resulted in a reluctance of the managements of different companies to implement ML-based solutions. The need for highly skilled personnel to interpret outputted results also increases a company's operational and capital expenditure. Thus, the need for ML-based solutions for direct marketing that produces accurate results that are accompanied with relevant explanations, is critical.

The aim of this study was to develop a hybrid machine learning process framework that could facilitate data-driven direct product marketing. The hybrid ML framework leveraged the statistical processing strength of ML models and addressed the lack of explanation by adding intelligent reasoning. The intelligent reasoning component was used to generate relevant explanations that accompanied the predicted outputs that are generated by a selected ML model. The hybrid framework has the capability to utilize multiple ML models. The applicability of the hybrid framework was demonstrated by using an anonymized and de-identified dataset from a South African-based telecommunication company.

The study adopted a design science, research strategy as the main methodological approach for the formulation of the research design. To achieve this, an analysis of existing literature, and interactive sessions with domain experts were conducted to gain an accurate understanding of requirements. This then formed the basis for evolving the architecture and the design of the hybrid ML framework. The implementation and evaluation of the prototype system was done within the framework of Design Science Research (DSR).

The hybridized ML framework consists of three ML models, which are Support Vector Machine (SVM), Random Forest (RF) and Artificial Neural Network (ANN). The intelligent reasoning component of the hybrid framework is composed of Case-Based Reasoning (CBR) and Rule-Based Reasoning (RBR). It also has a user-friendly interface that allows a user to select a specific operation/function that is desired, through which results of operations and explanations are also presented to the user. The hybrid ML framework was evaluated using the Goal Question Metric (GQM) approach in order to determine its performance and usability.

The evaluation results obtained from feedback of participants show that the results are adjudged to be very accurate, while the quality of explanations is also satisfactory. The study reveals that data-driven, direct marketing through the application of explainable AI methods is possible and is valuable to the telecommunication industry.

ACKNOWLEDGEMENTS

I would like to thank:

- GOD for HIS unconditional love upon life, it is this love which resulted in JESUS CHRIST the best of heaven to be sent to earth and lay down his life willingly for the redemption of all of the earth. It is also this love which resulted in the HOLY SPIRIT, the great comforter to live inside all who believe. I am truly grateful for this and the resurrection of JESUS CHRIST. He truly lives!
- My wife, Nicola Petersen for her love, support, prayers and encouraging words.
- My parents, Edmund & Elaine Petersen who sacrificed a lot for me throughout my life.
- Alfred & Olga Jacobs and James & Helen Esterhuizen who played instrumental roles in my life.
- Bradley Clark, Frederick Mitchell and Moegamat Solomons for all their support throughout this remarkable journey.
- Professor Justine Olawande Daramola for all his knowledge sharing, mentoring, guidance and support throughout the project from start to finish.
- Pastor Solomon & Wendy Klaasen along with the entire Kaleo Ministries International family for their spiritual guidance and teaching of the word of GOD.
- Lastly, I would like to thank all my colleagues especially Guillaume Olivier, Danie Oosthuizen, Jerome Brand and Guy Werner for their support on this amazing journey.

All mentioned above are not mentioned as an act of formality but are mentioned with true gratitude and thankfulness. My prayer is that GOD bless you all in abundance, and I know you will access all that were made available through the finished work of JESUS CHRIST. This was indeed an amazing journey and GOD placed only the best in my life to support me in completing it. Thank you.

DEDICATION

This study is dedicated to GOD the father, JESUS CHRIST the son and HOLY SPIRIT, the great comforter for HIS unconditional love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

LIST OF ABBREVIATIONS

ABBREVIATIONS	DEFINITIONS
AI	Artificial Intelligence
ANN	Artificial Neural Network
CBR	Case-Based Reasoning
CNN	Convolutional Neural Network
DBN	Deep Belief Net
DNN	Deep Neural Network
DT	Decision Tree
GQM	Goal Question Metric
IR	Intelligent Reasoning
ML	Machine Learning
MLP	Multi-Layer Perception
RBR	Rule-Based Reasoning
RF	Random Forest
SA	South Africa
SL	Supervised Learning
SVM	Support Vector Machine
UCD	Use-Case Diagram
UI	User Interface
UL	Unsupervised Learning
UML	Unified Modelling Language
XAI	Explainable Artificial Intelligence

CHAPTER 1
INTRODUCTION

1.1 Background

Today customers are more demanding than before, and the customary generic marketing strategies applied to all customers are not effective in getting products sold. Direct marketing is a method that is customer focussed whereby customer data are analyzed and product offerings are initiated based on the data analysis (Flici, 2011; Lezama, 2020). It is important for a direct marketing strategy to be data-driven as customer data will be used to establish customer characteristics and predict the probability that the customer would respond to a product offering (Alanen, 2016; Siregar et al., 2020).

The advancements of technology have enabled companies to store large sets of customer data from a variety of sources (Gordon et al., 2015; Ayyaz et al., 2019). The data sources have enabled for advanced data analytics by use of complex software. The data sets enable customer characteristics to be built around customer satisfaction, customer spending behaviour, demographics, and geographical location (Muhammedrisaevna, 2020).

Machine learning (ML) is a significantly improved way of executing predictions utilizing complex datasets from various data sources. It uses specific designed algorithmic models to recognize customer patterns from large data sources, which would have been challenging to identify by using traditional statistical methods or manual approaches (Erel et al., 2018; Zhang et al., 2020).

The drawback of ML algorithmic models is that usually after applying the ML algorithmic models, results are without explanation (Muller et al., 2014; Jiang and Nachum, 2020; Holzinger et al., 2019). It is this lack of explanation and understanding which causes management within companies to lack confidence in the results of machine learning algorithms (Lubeniqi, 2020). This complexity establishes a need for highly skilled personnel to setup machine learning frameworks and interprets the results. The effect of this need for highly skilled personnel increases capital and operational expenditure of an organization significantly (Thomas et al., 2020; Gao et al., 2019).

The commonly poor explanation characteristic of Artificial Intelligence (AI) systems has given birth to a significant interest in the field of explainable AI (XAI) (Bonacina, 2017). The data obtainable on product review websites such as capterra (www.capterra.com), softwareadvice (www.softwareadvice.com), and goto crowd (www.g2crowd.com) indicated that most ML-based direct marketing tools cannot provide a detailed explanation (Petersen and Daramola, 2020).

A detailed explanation is of great assistance to users in understanding customer behaviour better and increasing trust in machine learning frameworks. An explanation of why specific actions are taken, and why it fails or succeeds would increase trust in machine learning frameworks (Gunning, 2017).

The current landscape of the implementation and maintenance of machine learning frameworks is very much overshadowed by a significant negative sentiment (Bonacina, 2017; Sokol and Flach, 2019). This negative sentiment exists due to a lack of trust which derives from a lack of understanding of the reasoning of the results. However, it has been shown that users are more trusting and comfortable if an easy-to-understand explanation is provided (Gordon et al., 2015; Jan et al., 2020; Pintelas et al., 2020).

This research aimed to explore a hybridized framework that will enable predictive analytics using a framework that combines machine learning, and an intelligent reasoning model (IRM) to explain results. This type of framework would make it easy for users to get a better understanding of customer behaviour (Prentzas et al., 2007; Xu et al., 2020).

The objective of combining an IRM and ML model to facilitate explanation is to enable the use of domain knowledge and experience from previous use case scenarios, to generate the basis for rich explanations. To do this, the concepts of rule-based reasoning and case-based reasoning will be used (Olsson et al., 2014; Tang et al., 2019). Rule-based reasoning will enable the use of domain-specific knowledge represented in the form of if-then-rules, to make deductions that can provide a basis for good explanations. Rule-based reasoning makes the explanation provided very accurate and effective (Wojciechowski and Wrembel, 2020).

However, building a knowledge rule-base can be time-consuming, and inferencing is limited to only the set of pre-defined rules of the domain (Eisenstadt et al., 2019; Tang et al., 2019). The use of case-based reasoning (CBR) offers a way to overcome this limitation because it enables the reuse of knowledge from past occurrences to solve a problem, therefore its ability is not limited to predefined rules. A combination of RBR and CBR will be suitable for the design of the IRM (Vantara et al., 2018).

A data-driven direct marketing tool that is based on this type of hybridized framework will increase management's trust and confidence in the recommendations of direct marketing tools, make them easier to use, understand, and acceptable as a form of decision support and decrease capital and operational expenditure significantly (Koh and Tan, 2011; Slik and Bhulai 2019). This ease of use and understanding of the data-driven direct marketing of this nature would enable users with a lower technical experience to participate in predictive analytics. A detailed explanation would also reduce reliance on the need to have many highly skilled machine learning experts (Siau et al., 2018).

The telecommunication domain has noted a continued drop in revenue in essential areas such as text messaging and voice (Vantara et al., 2018; Wimmo, 2020). In conjunction with this, the domain has also seen a variety of new technologies coming into existence which includes over-the-top streaming services and message services. These influences have resulted in a constant need to upsurge product sales within the telecommunication domain. Having rich customer data at its disposal, a data-driven direct marketing strategy allows for the accurate product to be offered to the precise customer and at the right time (Amoaka, 2019; Stone et al., 2019). This makes this domain ideal for this research.

A study by Chen (2016) has shown that the retrieval of data from the service, resource, and customer layers within a telecommunication company can offer an insightful and accurate customer profile. In conjunction with this, the billing information consisting of spend limit, consumption, and usage data has also proven to be of value in predicting product offerings (Iyengar et al., 2007; Li et al., 2019). It is evident that there is a connection between current consumption and a customer's past spending pattern. This connection can possibly be exploited to predict meaningful products to offer to a customer at the right time (Prasljivic and Ramic, 2019; Quinnery, 2019).

Another study has also identified billing data as being an important component in building an accurate prediction of customer usage patterns, current and future needs. This type of data can be used to predict products that a customer is likely to purchase in the future (Gemechu, 2020). The study aims to add the billing data as a billing layer to the framework recommended.

1.2 Research Problem

Most data-driven direct marketing tools are based on ML. The lack of easy-to-use ML models and the lack of explanation of results have reduced the level of trust and reliability that users have in the results of ML models (Muller et al., 2014; Tseng et al., 2019). Also, highly skilled human resources are needed to compile a meaningful dataset, apply an algorithmic model to process the data, and interpret the output of ML-based tools (Siau et al., 2018).

This has an impact on users of the ML results as it reduces confidence in basing decision making on results. It also has an impact on companies as highly qualified experts are always in high demand (Muller et al., 2014; Pintelas et al., 2020).

The consequence of this is a low uptake of data-driven direct marketing tools and companies spending a lot more on ML experts for interpretation and use (Thomas et al., 2020).

Thus far, most ML-based tools that are used for data-driven direct marketing, lack good explanation capability and are not easy to use by non-AI experts.

1.3 Objective of the Study and Research Questions

1.3.1 Aim and Objectives of the Study

The research aims to develop an easy to use and explainable machine learning (ML) process framework for data-driven direct product marketing.

Based on the aim, the objectives of the research are:

i) To determine the requirements for a hybrid ML framework for data-driven direct product marketing.
ii) To design a hybrid ML framework for data-driven direct product marketing.
iii) To implement a hybrid ML framework for data-driven direct product marketing.
iv) To evaluate the hybrid ML data-driven direct product marketing framework in terms of performance, usability, and explanation of output.

1.3.2 Research Questions

In this research, the main research question is:

How can a hybrid ML framework for data-driven direct product marketing be created that is easy to use and has explainable results?

To answer the research question, sub-questions are created to substantiate the research. The sub-questions for this research are:

i) What are the attributes of an easy to use and explainable ML framework for data-driven direct product marketing?

ii) How can ML and intelligent reasoning be combined for data-driven direct product marketing?
iii) How can a hybrid ML framework for data-driven direct product marketing be implemented?
iv) How can a hybrid ML framework for data-driven direct product marketing be evaluated in terms of performance, usability and explanation of output?

1.4 Ethics

Ethics in research consists of principles on how a researcher and a research organization ought to conduct themselves when interacting with research participants, other researchers, colleagues, the users of their research and society in general.

This research adhered to Cape Peninsula University of Technology's (CPUT) research Code of Practice on Ethical Standards.

The research focused on groups, individuals and it did not cause any harm to any group or individual. All confidential information was kept confidential to prevent harm to an individual or group.

The data collected was in the trust and ownership of the researcher and none of the data was up for sale in any shape or form.

Anyone who participated in the research was free to withdraw at any time and all participation was voluntary. This was in adherence to the research code of conduct and in compliance to the CPUT policy.

1.5 Delineation of the Study

For the purposes of the study a fully functional hybrid machine learning process framework for data-driven direct product marketing was developed using data from a single South African-based telecommunication company.

The data used was for customers who purchase data products over a period of six months. The framework also focussed on only three algorithmic models which are Random Forest (RF), Support Vector Machine (SVM) and Artificial Neural Networks (ANN), but the concept of the framework is that it is not only limited to these three algorithmic models.

1.6 Significance of the Study

The significance of the study is that it will increase trust and confidence in a data-driven direct product marketing solution using machine learning. This type of solution would benefit the telecommunication industry as it enables businesses to get a better understanding of customer behaviour and increase sales.

The research provides a solution which is a template for a company to simply load a large amount of data and generate an easy to understand output. This output includes an accurate probabilistic prediction of which products to sell to which customers. This ease of use and explanation of results require a low-skill level resource to operate and maintain the tool. This, in turn, reduces a company's capital and operational expenditure.

The study was derived from a theoretical and practical perspective. Theoretically, the study contributes by combining an existing framework used to extract data within the telecommunication domain with a billing layer to build an accurate and insightful customer profile.

The theoretical contribution is also in the form of designing a hybridized framework consisting of machine learning and intelligent reasoning for an explainable output to users. The practical contribution will be in the form of a functional product which could be used for data-driven direct product marketing within the telecommunication domain.

1.7 Thesis Outline

The remainder of this thesis is organised as follows. Chapter 2 presents the literature review which covers background and related work. Chapter 3 describes the methodology used to achieve the research objectives. Chapter 4 presents the requirements analysis and design of the process framework. Chapter 5 discusses the details of the implementation and evaluation, while Chapter 6 presents the summary, conclusion, and recommendations for further research work.

CHAPTER 2
LITERATURE REVIEW

In this chapter, the context of the telecommunication industry in South Africa, along with the theoretical background on direct marketing and data-driven direct marketing are discussed.

The theoretical description of machine learning, the different types of machine learning algorithmic models, case-based and rule-based reasoning are also presented. In making the presentation the functionality, advantages and disadvantages are discussed to provide a full view of each concept. The evaluation of the framework will be an important step and therefore the concept of goal question metric is also presented.

Lastly, the chapter presents a review of related work in the area of applying machine learning, rule-based and case-based reasoning to direct product marketing. The review of related work includes studies within and outside of the telecommunication industry.

2.1 Telecommunication Industry in South Africa

The telecommunication industry in South Africa has become very competitive over the years with new entrants entering the market and Internet Service Providers also competing for certain products (Morgan and Govender, 2017; Wimmo, 2020). In addition, over-the-top application providers such as WhatsApp, Facebook, and Google also entered the market creating a new form of competition (Ngwnya, 2017).

With the three big telecom market holders in South Africa namely Vodacom, MTN, and Cell C, the market has seen several smaller entrants who entered the market with low-priced products to gain an initial market share (Mpwanya and Letsoalo, 2019). Global companies such as Starlink who is constructing a mobile network spanning right across the world poses a great threat to South African telecommunication companies as Starlink will be able to offer low-priced products based on the economy of scale (Mironova, 2018; Babu et al., 2020).

Telecommunication companies also face an additional challenge of strict regulatory rules within South Africa. Additional spectrum (wireless radio frequencies) allows for better quality of service, cheaper and new products (Somdyala et al., 2017; Pintelas et al., 2020). Obtaining new spectrum is normally a challenging task. Therefore telecom providers have to maintain current customers and get new customers with the technology that is currently available (Hinson et al., 2016). However, the South African government is in the process of allocating additional spectrum. This spectrum would allow for companies to offer an increased variety of products at a more competitive rate (Ngwnya, 2017; Walker and Brown, 2019).

South Africa is classified as an emerging market and almost everyone requires access to a telecommunication service in some form or the other. The market is very diverse, ranging from customers with lots of money and other customers who are below the poverty line (Pau, 2011; Wimmo, 2020). It has large enterprises and small businesses whose sole purpose of trading is to provide for a household. It then becomes important to understand these customers on completely opposite sides and offer products directly (Isa, 2019). The market has been known to become saturated over the years with operators contending aggressively for customers. It has been increasingly visible that a number of companies have been increasing investment into machine learning solutions to increase revenue (Kau and Kogeda, 2019; Tseng et al., 2019).

It has become crucial to provide quality products and really understand which customer to market certain products to, as customers expect almost tailor-made products to address customer needs (Hadzic, 2019; Walker and Brown, 2019). A study that focused on telecommunication customer loyalty, collected data from 227 subscribers from various mobile network providers. It concluded that customer loyalty and reducing customer churn rates are a key factor in the success of telecommunication companies going into the future (Robb and Paelo, 2020). Customer communication and the way products are marketed to customers have a significant contribution to overall customer satisfaction (Mpwanya and Letsoalo, 2019).

Due to the challenging regulatory environment, increased competition, and high customer expectations, direct marketing is an effective strategy to sell the right product to the right customer at the right time (Iyengar et al., 2007; Babu et al., 2020). The success of any direct marketing strategy is dependent on obtaining the right customer data at the right time. Within a telecommunication company, data can be collected from the resource, service, and customer layers to build an accurate and insightful customer profile (Chen, 2016; Iwuchukwu et al., 2019).

The literature also indicates that customer billing information which includes spending limit, consumption, and usage data, has also been of value in predicting product offerings as shown in Table 2.1. Using features such as customer spending and consumption patterns, yielded positive results in executing effective product offering predictions. The key to these types of predictions is executing the predictions at the right time (Petersen and Daramola, 2020).

Billing data is an important component in building an accurate prediction of customer usage patterns, current and future needs (Koh and Tan, 2011; Li et al., 2019). This type of data can be used to predict products that a customer is likely to purchase in the future which is the main purpose of the direct marketing strategy (Putit and Abdullah, 2019). A number of organizations within the telecommunication, retail, and marketing industries are including billing data to determine which customers are potential purchasers of certain products (Aung et al., 2019).

Table 2.1 Layers within a telecommunication company (Chen, 2016)

DATA LAYERS WITH A TELECOMMUNICATION COMPANY	
LAYER NAME	**DESCRIPTION**
RESOURCE	*Network devices, Switches, Mobile Devices*
SERVICE	*Voice, Data, Video Service*
CUSTOMER	*Incidents, Orders, Service Provisioning*
BILLING	*User Consumption, Monthly Spend and Product Offerings*

2.2 Direct Marketing

Direct marketing is an effective way to target customers in a unique manner in order to increase sales of products (Low et al., 2011; Brezak et al., 2019). The main purpose of direct marketing strategy is to increase awareness of products, to establish a demand for specific products and to increase the purchase of products with minimal contact with the customer (Santucci et al., 2012; Yiyou, 2019).

Previously, companies would use customer contact details such as phone number, email address or physical address to send marketing material directly to customers (Hollon, 2004; Gatzke, 2019). This method has seen an increase in product sales, but with increased competitiveness and greater consumer power, a more direct approach was needed (Alanen, 2016; Barnett et al., 2020).

In current times, companies build customer profiles using historical purchasing, location, online browsing and a number of other personal data to determine if a customer is most likely interested in a specific product. This is applied by using complex data processing mechanisms and technologies (Lezama, 2020).

With the development of technology, a more technological approach to execute a direct marketing strategy is needed. Customers can be reached using various platforms such as Short Messaging Service (SMS), social media, downloaded applications, instant messaging application and many other methods (Alanen, 2016; Yiyou, 2019). Irrespective of the means of communicating with the customer, the development of a technology-based direct marketing strategy requires an in-depth understanding of the customer, to identify the customer's behavioural patterns (Slik and Bhulai, 2019).

To establish an accurate customer behavioural pattern requires vast amounts of data and will allow for customer segmentation (Santucci et al., 2012; Khan, 2020). Customer segmentation will include children, youth, adults, female, male or the type of monthly spend. Customer segmentation will allow for greater, tailored product development and decreasing, unnecessary communication with the customer (Renton and Simmonds, 2019).

Communicating with customers on multiple technological platforms can also increase sales significantly. However, the right product needs to be presented at the right time (Tkachenko et al., 2016; Firman, 2019). In many cases, smaller businesses rely on face to face conversations with customers whereas the larger businesses rely on marketing exclusively to its customer base (Wright, 2019).

With South Africa having a large number of both these groups, a direct marketing strategy requires to be personalised to complement the face to face communication and communicating in a meaningful way to the customers of large enterprises (Daugherty and Biocca, 2008; Mkhize and Ellis, 2020).

The pharmaceutical industry in New Zealand has identified a similar outcome where customer data was used to market prescription drugs. A direct marketing campaign was used to address undertreated conditions and resulted in customers receiving the right treatment for conditions by simply analysing the data. A significant increase in sales was realised throughout the country (Vandevijvere et al., 2017; Renton and Simmonds, 2019).

Within the banking sector, it is critical to be the first in offering credit to qualifying customers. The reason for this is that the customer could possibly exhaust all credit with a competitor which will result in the company having to wait for the customer to first settle credit before becoming a potential customer again (Migueis et al., 2017). As a result, banks include algorithmic models in direct marketing strategies to classify if a customer would be likely to take on new credit or not. These strategies include a number of customer parameters, large volumes of data and would be very difficult to produce similar results using traditional statistical methodologies (Ladyzynki et al., 2019; Sadq et al., 2019).

In the age of information, more detailed sets of data have become key in getting to the customers first and with the right products. Studies have shown that to achieve this the right data set and data processing mechanisms are required; the right data permits for customers to be grouped accurately; and products can be marketed to these customer segments (Alanen, 2016; Khan, 2020).

A direct marketing strategy can also be categorized into inbound or outbound marketing. Inbound marketing is a strategy where potential customers are contacted through social media platforms, blogs or entertainment platforms (Patrutiu-Baltes, 2016; Babu et al., 2020). On the other hand, outbound marketing is defined as a more intrusive strategy where potential customers are contacted through the more traditional methods of print advertisements, call centre operators or direct mailing (Hawik, 2018). Whichever category an organization implements, both are heavily reliant on the available customer data (Dakouan et al., 2019; Brezak et al., 2019).

A direct marketing strategy utilizes a number of key steps across the targeted customer base. The key steps are to identify and understand customer behaviour, select customers to be targeted, select the products which are to be sold to the targeted customers, select the time to propose the product to customers and selecting the medium to communicate with the customer (Ayyaz and Hussain, 2019).

Although looking at a customer's behaviour as an individual, it is also important to look at what the behaviour is of a group of customers in a specific location (Al-Rifaie and Alhakbani, 2016; Pittz et al., 2019). This would allow for the direct marketing strategy to be structured for a specific location and in turn, more potential customers could be targeted with a single strategy (Deffner et al., 2020).

2.3 Data-Driven Direct Marketing

With the technological advancements over time, the introduction of data-driven direct marketing has allowed for an even more effective direct marketing strategy (Sweers et al., 2014; Slik and Bhulai, 2019). Various technologies enable companies to store large data sets which can be utilized to construct customer characteristics and identify customer patterns. This data can be used to execute the prediction of the probability that a customer would potentially respond to a specific product offering (Flici, 2011; Lezama, 2020).

A number of companies have noted a significant shift in analysing customer behaviour, distribution of customer data and methods where each transaction initiated by a customer is analysed (Daugherty and Biocca, 2008; Grandhi et al., 2020; Khan, 2020). Data-driven technologies such as machine learning, blockchain, and programming advertising have shown product sales increases when implemented in the correct manner (Camilleri, 2020).

To build customer characteristics and identify patterns require the right data set consisting of the right features (Gardan et al., 2015; Stone et al., 2019). Selecting the right features can be time-consuming as features need to be included and the contribution to the prediction needs to be assessed. Stronger customer patterns are identified by also excluding features containing ambiguous values (Li et al., 2020).

Historical data and the quality of this data is an important factor in predicting demand for products (Reutterer et al., 2006; Ghasemaghaei and Calic, 2019). As a company's system and data ages, the quality of this data normally deteriorate. Proper data cleaning and wrangling processes need to be implemented before executing any type of forecasting. With good forecasting of profits, maximized orders can be placed to manufacturers on time which in turn would reduce storage or holding cost (Kumar et al., 2020; Pittz et al., 2019).

A data-driven approach yields better results to know when and how to construct a push or pull strategy. The data on customers usage on the network will inform the direction of a pull strategy whereby a customer can be encouraged to pull certain product offerings which the customer might be interested in at that specific time (Hossain et al., 2020).

A study within the sports marketing domain also echoed the sentiment that the dataset is of utmost importance in selling products directly to customers. However, it highlighted the importance of customer behaviour always changing hence the model needs to change (Kumar et al., 2020; Walker and Brown, 2019). A number of data-driven direct marketing tools exists and are used by some companies. Some tools rely on data retrieved from emails and social media, but other tools gather large amounts of data from various sources such as customer-relationship management systems (Coita et al., 2019).

Over the years retailers have relied on data extracted from the customer relationship management database. This database provides a consolidated view of a customer's behaviour (Lawi et al., 2017). By identifying customer behavioural patterns, it allows for retailers to increase customer basket size by offering products when the customer needs them most (Avery and Israeli, 2018; Stone et al., 2019). Being aware of the customer behaviour also allows for an accurate customer segmentation for retailers to have the right products available at the right time which in turn reduces overhead costs (Li et al., 2020; Firman, 2019).

Data-driven direct marketing strategies in some cases have been known to use data relating to the customer interaction with products, services and the actual brand to visually create a customer journey mapping (Troisi et al., 2020). These types of visual methods enable companies to take complex data sets and present them to management in a simplified manner for improved decision making (Micheaux and Bosio, 2019; Babu et al., 2020).

2.4 Rule-Based Reasoning

Rule-based reasoning is an approach where rules are defined as normal standard behaviour and if a rule was fulfilled or not, an action is taken accordingly (Golding and Rosenbloom, 1991; Seydoux et al., 2019). A rule can be described as an absolute in a specific environment, whereby if a rule applies then the action taken should be exactly as it was configured, no more or no less (Brożek, 2017; Ducamp et al., 2020).

Rule-based reasoning is domain-specific and building the knowledge base requires interviewing domain experts. It is also seen as a representation of general knowledge relating to a specific domain and the basic structure consists of if-then-rules (Jetinai, 2018). The condition is a set of logical sequential rules which form a logical function within a system. If the required conditions are met a conclusion is established and a trigger is executed (Prentzas et al., 2007; Sheng-Sheng and Fan-Liang, 2019).

A rule can be inserted, updated, or deleted easily (if there is no dependency on another rule) from the knowledge base without disrupting the functionality, making it very flexible throughout the development. Normally, the system contains a rule base, inference engine, working memory and explanation mechanism (Bialek et al., 2017). The facts concerning a specific problem are stored in the rule base, and these facts are used by the inference engine. The working memory stores all the facts for processing and the explanation mechanism explains the conclusion (Berka, 2020).

If rules are maintained and structured correctly the rules could simulate a small part of the domain expert's rules which is normally applied when the domain expert is assessing a situation (Abu-Nasser et al., 2018). The rules created by the domain experts are normally generalized trends according to the knowledge of the domain expert which possibly could result in small number of exceptions being missed (Delgoshaei et al., 2017). Where rules miss a small amount of exceptions this can be hybridized with case-based reasoning to mitigate this risk (Zhang et al., 2019; Seeger et al., 2019).

2.4.1 The Structure of Rule-Based Reasoning

As rules are the natural perception of a domain expert's knowledge and a rule looks very much like the expression of a natural language, the rule attempts to replicate a scenario where a domain expert would be looking at the scenario in reality and answering the question (Quinn et al., 2017; Carral et al., 2019). The "IF" statement relates to certain information given which will be led to an action by the "THEN" part of the statement. A rule can consist of a number of rules joined with the inclusion of an "AND" or an "OR" as below (Saul and Wuttke, 2013; Rakib and Uddin, 2019).

IF part: This part is called *antecedent, premise or condition*

THEN part: This part is called *consequent, conclusion or action*

A rule consisting of two parts, the object and the value. These parts are joined by an operator and includes not only symbolic operators such as "is", "is not" but includes mathematical operators such as "<", ">", and "=" as shown below (Berka, 2020; Carral et al., 2019).

IF "*answer_1*" is wrong **THEN** "*action*" is "*retry*"

IF "*score_1*"< 5 **THEN** "action" is "*move_to_next_question"*

2.4.2 Advantages and Disadvantages of Rule-Based Reasoning

The advantages of rule-based reasoning are that it is flexible as rules can easily be added, updated, or removed and the author can provide the exact justification for the existence of the rule. It is this justification that contributes significantly to trust in the performance (Khandelwal and Sharma, 2013; Thangaramya et al., 2020).

Rule-based reasoning has been known to produce results speedily as it relies on a group of predefined rules configured by domain experts (Quinn et al., 2017). These experts apply these configurations based on past experiences and normally vast industry experience. This also results in increased accuracy and decreased errors (Thike et al., 2019; Kourtis et al., 2019).

Rule-based reasoning also allows for new authors to understand the rules easily which is great for continuity (Arndt et al., 2017; Angilica et al., 2019). As rules are amended mechanisms can be put in place for increased managerial and audit control, which is a great feature compared to machine learning techniques which are often applied in a black-box manner (Schwabe et al., 2019).

The disadvantages of rule-based reasoning are that if the rules are not managed correctly and updated regularly it might not handle new situations well and it has to be accepted that there is not a rule for every possible situation (Krupitzer et al., 2018; Seydoux et al., 2019). Adding, updating, or deleting a rule can be time-consuming as the change needs to be implemented correctly and documented to maintain change control. For effective rules to be created there is a reliance on a very knowledgeable and experienced author (Song et al., 2020; Ducamp et al., 2020).

2.4.3 Applications of Rule-Based Reasoning

Rule-based reasoning is very useful and effective in the application of a traffic monitoring system. The rules allowed for distinguishing between high and low-level image processing. The image processing extracted visual information using spatial and morphological analysis. For scenarios where rules were not available, an artificial intelligence technique was used (Al-Sharif et al., 2017; Petersen and Daramola, 2020).

In implementing a handwriting recognition solution that focused on identifying handwritten digits, the rule-based reasoning module enabled easy management of rejection criteria. This improved the reliability and speed of the system in producing results. The rule-based reasoning module was used in conjunction with a Support Vector Machine classifier algorithmic model (John and John, 2010; Delgoshaei et al., 2017).

In more recent times rule-based reasoning has been combined with artificial intelligence methods as a way to optimize manufacturing processes and making them more efficient. The rule-based component allows for early fault detection in production processes as its configuration encapsulates past occurrences. The rapid response to these fault occurrences results in production continuity and increased production (Ghahramani et al., 2020; Khosravani and Nasiri, 2020).

2.5 Case-Based Reasoning

Case-based reasoning (CBR) is a process that uses past occurrences of a situation to provide a solution to new occurrences. The process has been proven to have strong explanation mechanisms as it bases its explanations on similar previous cases (Aamodt and Plaza, 1994; Daramola et al., 2013; Keane and Kenny, 2019). It is an experience-based system and formed as the basis for intelligent computer models designed to adapt and solve new problems. The CBR system accepts a new case, then stores it in its knowledge, and if a similar case occurs, the stored knowledge is used (Shen et al., 2017).

A knowledge base is constructed in the form of a database consisting of previous cases. When a new case is presented a query is executed to determine if a similar case exists (Richter et al., 2016; Wan et al., 2019). This similarity is assessed by comparing the weighting of each feature by the use of an algorithm. When compared to algorithmic models such as deep learning, with case-based reasoning it is easier to understand why a specific outcome was presented (Lamy et al., 2019; Nakhjiri et al., 2020).

Case-based reasoning can be utilized to address the lack of explanation and simplify the interpretation of results. Case-based reasoning substantiates the predictions made and it would complement the machine learning model in this way (Choudhury et al., 2016; Berka, 2020). It will examine past occurrences of similar output and provide a comprehensive explanation of the current output. Accurate evidence is used from a set of applicable cases to explain the task at hand. Explanations are normally simplistic regardless of the complexity (Fei and Feng, 2020).

2.5.1 Functionality of Case-Based Reasoning

Aamodt and Plaza proposed that the case-based reasoning cycle consists of 4 sequential steps, which are retrieve, reuse, revise and retain as shown in Figure 2.2 (Aamodt and Plaza, 1994; Zhao et al., 2019).

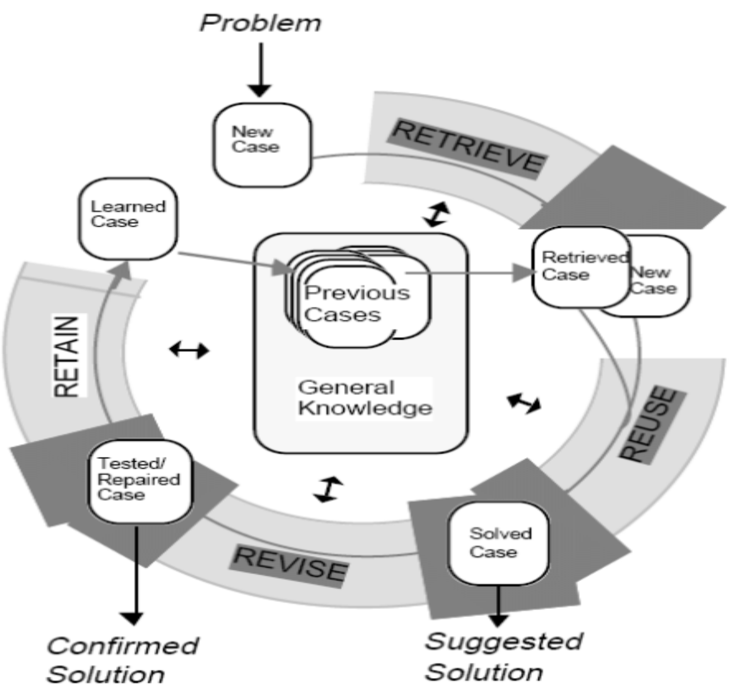

Figure 2.2: Case-based reasoning cycle (Aamodt and Plaza, 1994)

2.5.1.1 Retrieve

When a new case enters the retrieval phase, it looks at analysis cases that occurred previously and retrieve a list of cases that most likely match the new case. The retrieval of these past cases is done by utilizing the features of the new case and using these features as indexes (Ketler, 1993; Smiti and Elouedi, 2020). In many cases, the CBR process has been found to use superficial indexes. However, other CBR processes have been configured to use more detailed and in-depth descriptors as indexes (El-Sappagh et al., 2015; Mabkhot et al., 2019).

In the event where the indexes of the new and previous case match, the previous case is retrieved. From the list of retrieved cases, a subset of cases is selected where only the most relevant cases are selected and the least relevant cases are excluded (Lamy et al., 2019).

The element makeup of the CBR structure is the knowledge-base which consists of all the previously occurred cases and the similarity measure which assesses the new case to determine the similarity of the case (El-Sappagh et al., 2015; Flores et al., 2019). How a case is represented in the knowledge-base is of utmost importance. The attribute definition, description of case content, and how the case is organized within the knowledge-base are significant factors for successful representation (Aamodt, 1995; Sheng et al., 2019). Although the case-base is not limited to a specific structure to physically store cases, in most cases it is an object-oriented or relational database, plaintext, or XML structure. This structure is also accompanied by indexes to ensure a timeous retrieval of cases (Batzias and Kopsidas, 2019).

The case retrieval and comparison are conducted in a fairly straightforward way where the features of the new case are compared to the features of the already existing case. The previous and current occurrence of the case must have the same features for a comparison to be initiated (Lee et al., 2020). The system does not build any type of constraint or relationship between features. Each feature of the new occurrence of the case is assigned a weighting. Once the weightings have been assigned, the similarity between the previous and new occurrence of the case is computed by using an equation as shown in Figure 2.3 (Zhao et al., 2019; Nakhjiri et al., 2020).

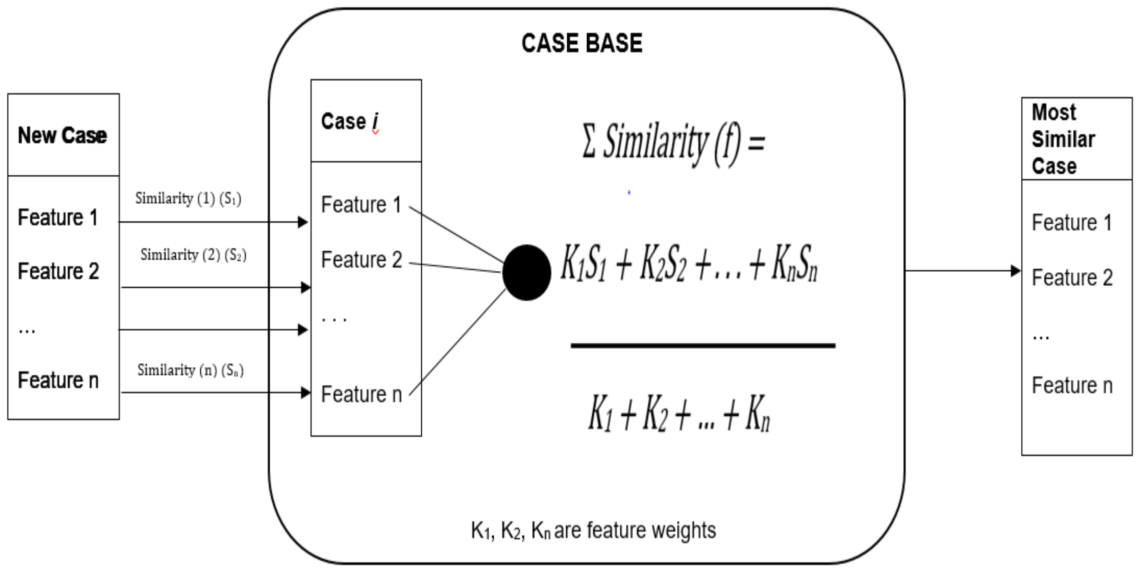

Figure 2.3: Case retrieval mechanism (Zhao et al., 2019)

2.5.1.2 Reuse

The reuse phase takes the list of most relevant retrieved cases and selects a solution that was applied to a previous occurrence of the problem. It selects a solution by focusing on the difference between past and current occurrences of the problem (Lamy et al., 2019). Another part focuses on which part of the solution applied from the past occurrence can be applied to the current occurrence (Aamodt, 1995; Sheng et al., 2019).

Depending on the complexity of the cases being compared the case can be divided into multiple parts and indexes assessed to determine if all or only parts of the previous case will be used (Batzias and Kopsidas, 2019; Berka, 2020).

2.5.1.3 Revise

The revised phase assesses the solution selected in the reuse phase to determine if the solution is the correct or incorrect solution. In the event that a case selected in the reuse phase is not the correct solution for the problem, the revised phase is where the failure is assessed and corrected (Arditi and Tokdemi, 1999; Lee et al., 2020).

2.5.1.4 Retain

The retain phase is where the new knowledge which was deemed as useful in the previous phase is added to the existing knowledge-base (Rahman et al., 2018). It consists of selecting which information from the case to retain in the knowledge-base, decides on the form in which it would be retained and applies the index for case retrieval in the future (Goel et al., 2017).

Where a failure has identified a repairing of the case takes place, as it updates the knowledge-base to improve the future retrieval, and reuse phases. Learning from the case success and failure is important to the future performance of the CBR process (Berka, 2020).

2.5.2 Advantages of Case-Based Reasoning

The process in which CBR compares previous and current cases is quite efficient due to its use of indexes. It also enables to handle of large amounts of cases in several seconds making it favourable for large data sets (Fei et al., 2020). CBR's ability to evaluate its successful, failed case selection and repairing the failed cases within the knowledge-base, enables for improved case selection over time. A common finding is that solutions are produced speedily, and it highlights the importance of features used (Gigih et al., 2020; Smiti and Elouedi, 2020).

2.5.3 Disadvantages of Case-Based Reasoning

The initial build of the CBR knowledge base can require lots of effort and time. This is dependent on the domain. In the event that the initial knowledge base is set up poorly, previous cases will be insufficient to compare to current cases and it could take a lot of time to correct. There is a risk that previous cases within the knowledge-base might be biased resulting in selected solutions being skewed (Relich and Pawlewski, 2018; Lee et al., 2020).

2.5.4 Applications of Case-Based Reasoning

The manufacturing industry has seen case-based reasoning solutions being implemented where complex geometrics were produced as output in constructing plastic products (Relich and Pawlewski, 2018). Later, these case-based reasoning solutions were used in conjunction with artificial intelligence functionality. The case-based reasoning functionality has proven to produce output accurately and speedily (Wan, 2019; Khosravani and Nasiri, 2020).

CBR has previously been applied in the medical field by offering credible diagnosis in identifying disorders. The doctor would identify the novel diagnosis combination of disorders and this knowledge would be used in assessing future occurrences. This resulted in time being saved as the doctors did not have to go through all the data to identify a disorder that was previously diagnosed (Faia et al., 2017; Silva et al., 2020)

In the milling process which is normally an extremely high-speed process, accuracy in predicting surface residual stress and roughness is critical in producing a quality product. Case-based reasoning was proven to produce an 80% prediction accuracy with a low error rate. In addition to this, the solution was fairly low cost to implement throughout an organization (Silva et al., 2020; Xu et al., 2020).

Similarly, a computer program also found CBR to be useful as it was used to produce computer troubleshooting reports. The reports would list the possible solutions were was found to be successful. Solutions were produced efficiently and speedily, and computer technicians had more time to spend on other tasks (Petersen and Daramola, 2020).

2.6 Machine Learning

Machine Learning (ML) is classified as a subfield of Artificial Intelligence with many computer scientists looking at the field of study since the 1950s. ML enables computers to identify patterns in vast amounts of data and also learn these patterns (Nilson, et al., 1998; Krems, 2019). The ability to learn the patterns enables the computer to predict a future occurrence without being explicitly programmed. This allows for identifying crucial relationships and correlations that exist in big data sets (Simeone, 2018; Choubin et al., 2019).

This computer-based process analyses existing data as input and produces a prediction as output by using an algorithmic model (Alpaydin, 2009; Qu et al., 2019). An algorithmic model is a set of instructions utilized to transform the input into the output. At a high level, machine learning consists of two phases. The first phase applies an algorithmic model to a data set to train the model on a dataset. In the second phase, the model is used to make predictions on new data (Carleo et al., 2019; Tang et al., 2019).

The application of machine learning varies across domains. With cancer research, machine learning has been used to look at large complex datasets which include proteomic and genomic measurements for cancer diagnosis and detection. It would be extremely difficult to achieve this with traditional computing methods (Goldenberg et al., 2019; Belkin et al., 2019).

Machine learning has also been applied to find missing persons which may be suffering from Alzheimer's or dementia. A person is fitted with a wearable camera that includes a GPS sensor. As the person moves through an area a neural network identifies geographic locations derived from training images and maps the images to an exact name of a location. This information is used to inform relatives where the person is at a specific point in time (Cheng et al., 2018; Himanen et al., 2020)

Generally, ML is divided into supervised and unsupervised learning. Supervised learning is the ability of an algorithm to identify patterns within a dataset by using labelled features as an input and predicting a targeted outcome (Cruz et al., 2006; Wiens et al., 2019). Unsupervised learning is the ability of an algorithm to process data into a cluster by identifying underlying features from unclassified or uncategorized data (Berry et al., 2018; Bhatt et al., 2020)

Supervised learning must have input data labelled. This process is normally time-consuming and one needs to know which value is being predicted. It has been deemed to be more structured and easier to understand the performance than unsupervised learning. With supervised learning, it is important to fit the model correctly, with the aim being to accurately predict future occurrences. Caution is needed to not over fit or under fit the model. Overfitting normally occurs when the algorithmic model becomes too familiarized with the training data. Underfitting normally occurs when the algorithmic model has not learnt enough from the training data (James et al., 2013; Sathya and Abraham, 2018). Supervised learning also consists of classification and regression problem-solving. In solving a classification problem, the algorithm predicts a category the items belong to. Regression on the other hand attempts to predict a specific value for a time period going into the future using inputs (James et al., 2013; Wiens et al., 2019). Unsupervised learning does not require initial labelling of data and as a result, it normally makes connections that humans normally would not (Sathya and Abraham, 2018; Qu et al., 2019; Krems, 2019).

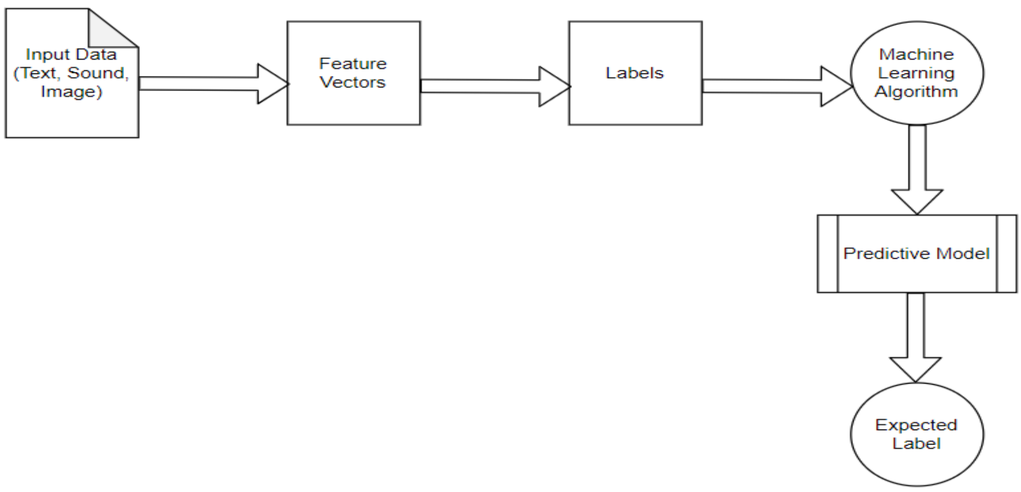

Figure 2.4: Supervised learning structure (Sathya and Abraham, 2018)

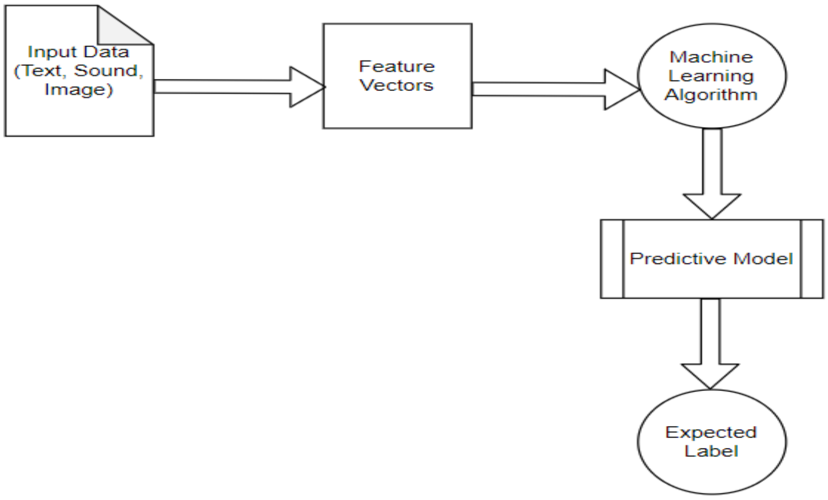

Figure 2.5: Unsupervised learning structure (Sathya and Abraham, 2018)

Numerous efforts that investigated the utilization of machine learning processes to advance direct marketing have been noted in the literature. These efforts included four machine learning models being compared to analyse the success in predicting potential customers for banking products (Beam and Kohane, 2018). The results indicated that the utilization of machine learning models enables the processing of large data sets, allowing for a greater historical view of customer spending patterns. The focus was on the accuracy of results and not the establishment of an explanation or lack thereof (Nachev, 2015; Belkin et al., 2019).

It was noted that the utilization of machine learning models advances the overall marketing considerably as this was tested by the application of machine learning models to customer datasets comprising of a number of customer features to produce predictions (Bayoude et al., 2013; Himanen et al., 2020). In a study to categorize customers for potential product offerings within the telecommunication industry, the use of machine learning models proved to reduce unwanted customer communication as customers were predicted more accurately. Unwanted communication normally has a negative impact on customer satisfaction (Lian-Ying et al., 2019; Nguyen et al., 2019).

2.7 Random Forest

Decision Trees is a supervised ML algorithmic model, can be used for classification and regression but in many cases, it’s used to solve classification problems. The model is the basis of the Random Forrest (RF) algorithmic model and utilizes a tree-like structure to solve problems. Using Attribute Selection Techniques, an attribute is selected as a root node. This attribute is used to divide the data into two or more datasets. This results in the creation of decision nodes that are branched out of the root node. A decision node can be split into two or more sub-nodes by applying a process known as splitting. This splitting process uses if-else conditions to split the node. The decision tree concludes with a terminal node. This is the point where the tree cannot be split into further sub-nodes (Sekhar and Madhu, 2016; Ao et al., 2019).

The Random Forest (RF) algorithmic model consists of a vast number of trees that are constructed in a random manner and each tree places a weighting on a specific class of the trees. These trees comprise of a set of binary rules used to compute a specific value (Liaw and Wiener, 2002; Chen et al., 2019). RF uses an algorithm to split a node by using a set of predictors for each node. RF also has the ability to handle a large set of output as these trees grow as more input classes are supplied (Sekhar and Madhu, 2016; Stafoggia et al., 2019).

The RF algorithmic model performs speedily and efficiently with large complex data sets. This is mainly due to the model using a tree structure to understand the data and making decisions based on the nodes and branches within the tree structure. The model performance output lacks an explanation of results which leads to a lack of understanding of how the model came to a specific prediction (Belgiu and Dragut, 2016; Brieuc et al., 2018).

2.7.1 The Functionality of Random Forest

The RF algorithmic model accepts the input data and a decision tree is started as the root (root node) of the tree moves in a downward manner. The data is analysed at the root node and a decision is made to branch into additional nodes (Roguel et al., 2018). In the event where the chain comes to an end as no further decision can be made, this is known as a leaf node. A node within the decision tree structure forms part of a characteristic and the branch is normally a range of values as shown in Figure 2.6 (Sekhar and Madhu, 2016; Wang et al., 2019).

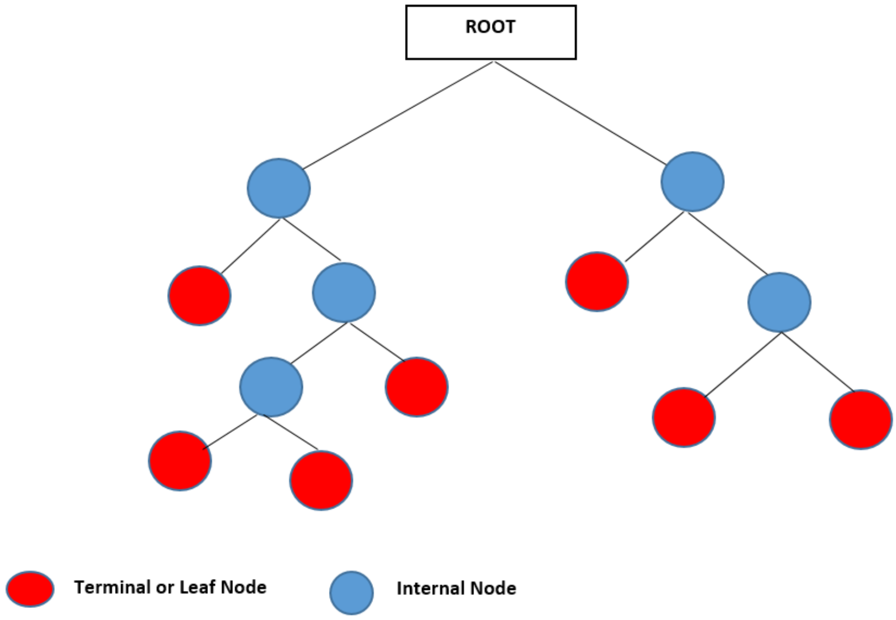

Figure 2.6: Decision tree structure

In the event of solving a classification problem the RF model will firstly establish the number of trees to be created to perform a prediction. A bootstrap method is then utilized to select data and for each tree the same sample size is created (Chen et al., 2019). The tree grows by applying splitting rules at each node and the tree is not pruned in any way. At each node random features are selected and as no pruning is required, all the created trees branching into different nodes are now part of a forest (Xuan et al., 2018). The forest consist of all the trees and each tree has produced a classification result. The forest produces a final classification result by selecting a classification with the highest weighting (Sekhar and Madhu, 2016; Pouladi et al., 2019).

When Random Forest is used to solve a classification problem, in most cases it is done by utilizing the Gini index to establish the branching of nodes within the decision tree by using the below equation (Chen et al., 2019; Zamani et al., 2019).

$$Gini = 1 - \sum_{i=1}^{C} (P_i)2$$

Where P_i is used to calculate the probability of class *C*. The sum is then calculated over the class.

2.7.2 Advantages and Disadvantages of Random Forest

The advantages of the RF model are that it does not require pruning and overfitting is not a problem as with other decision tree-based models. The model normally performs with a high level of accuracy, feature importance is produced automatically and not too sensitive with outliers when processing the training dataset (Shi et al., 2019; Ao et al., 2019).

The disadvantage of the RF model is that it has been found to perform with a lower level of accuracy when solving regression problems. The RF model often underestimates high values and overestimates low values (Dogru and Subasi, 2018). The RF model also produces predictions without an explanation which makes it difficult to understand why it produced a certain prediction (Singh et al., 2017; Tan et al., 2019).

2.7.3 Applications of Random Forest

The RF model has been found to be applied in predicting potential customers in several direct marketing campaigns. In doing so the RF model produced a potential customer list with reasonably high sensitivity, specificity, and accuracy which in turn resulted in successful direct marketing campaigns (Ladysynski et al., 2019).

Several banks have found it challenging to classify from various customer channels which customers to market long term products to. RF resolved this classification problem by analysing data from across channels and classifying which customers are likely to purchase long term products. For the first two quarters, long-term products were increased by 60% (Lawi, et al., 2017; Wang et al., 2019).

RF was also used in the banking sector to classify customers who potentially would subscribe to long term deposits. A banking dataset was labelled with the necessary attributes and customers segmented by region. RF was found to place customers in a classification group with a higher level of accuracy (Appiahene et al., 2020).

2.8 Support Vector Machine

The Support Vector Machine (SVM) algorithmic model was first introduced in the 1960s by Vapnik. SVM was mostly only popular within the Neural Information Processing System community but over the years it has become an important part of ML research throughout the world (Lessman, et al., 2006; Feng et al., 2019). The model became popular by performing handwriting analysis with high levels of accuracy (Jakkula, 2006; Niu et al., 2019).

Support Vector Machine (SVM) has been found to be a reliable and accurate classifier technique for identifying patterns across domains. The model is known as a supervised learning algorithm which uses the means of maximal, marginal hyperplane to execute a prediction (Aoyagi et al., 2019).

While Neural Networks formulation are known to use the Empirical Risk Minimization principal, SVM formulation uses the Structural Risk Minimization principle (Richhariya and Tanveer, 2020). SVM uses kernel functions to search a whole range of hypothesis spaces and instead of using non-convex optimisation to solve problems SVM uses a quadratic linear equality and inequality problem (Niu et al., 2019; Fu et al., 2019).

2.8.1 The Functionality of Support Vector Machine

To explain the functionality of SVM, let's take the following:

$$Classes: P, N$$
$$for\ y_i = +1, -1$$

The SVM algorithm outlines the linear separable case and searches for the most optimal hyperplane that is equidistant from either class (Cervantes et al., 2020). The model then introduces Kernel Functions to produce non-linear decision surfaces. The model then addresses any form of noisy data slack variables to permit for handling of training errors (Burbidge and Buxton, 2001; Tang et al., 2019).

Also, when calculating the SVM for classification the following can be considered:

$$[a]\ If\ Y_i = +1; wx_i + b \geq 1$$
$$[b]\ If\ Y_i = -1; wx_i + b \leq 1$$
$$[c]\ For\ all\ i; y_i(w_i + b) \geq 1$$

Where x is the vector point with w representing the weight, [a] should always be greater than 0 in order to separate the data. SVM will look at all the available hyperplanes, and then select the one where the distance of the hyper plane is the largest (Niu et al., 2019; Adams et al., 2020). Each of the test vectors will be located in radius from the training vector. In the event that the selected hyperplane is at the furthest point from the data, the hyperplane which maximises the margin will intersect between the closest points on the convex hull of the two datasets as shown in Figure 2.8 (Ling et al., 2019; Rezvani et al., 2019).

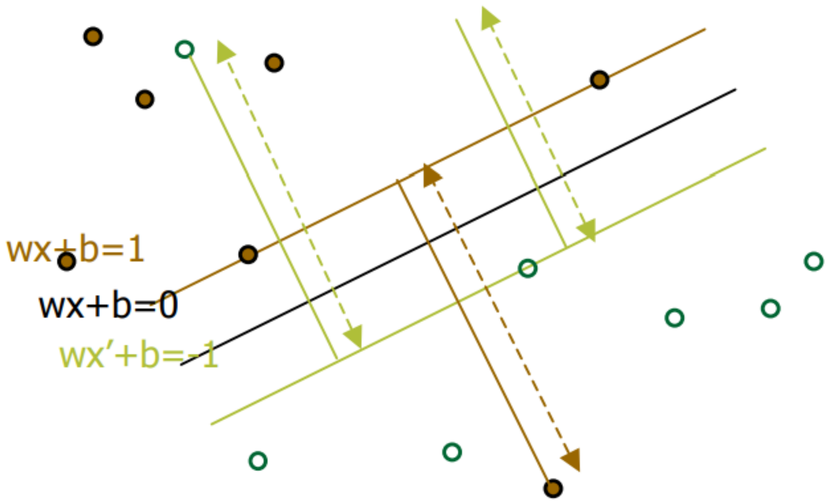

Figure 2.8: Representation of classification hyperplanes

The maximum margin hyperplane defines the separation amongst the different decision classes. The inputs which are the nearest to the maximum margin hyperplane are known as support vectors (Bi et al., 2018). The remaining furthest inputs which are then known to be irrelevant in the process of defining the binary class boundaries in a linear separation case, whereby support vectors are separated from binary decision classes, the following three attributed cases can be used and represented as follows:

$$y = w_0 + w_1\, x_1 + w_2\, x_2 + w_3\, x_3$$

y is calculated by computing the three attributes x with a weighting w learned and assigned by the algorithm. As the weightings are the parameters which establish the hyperplane, the maximum margin hyperplane can be represented as follows:

$$y = b + \, \Sigma\, \alpha_i y_i x\, (i), x$$

y is the class value for the inputs and x the vectors. $x(i)$ Is the support vector, "." the product b and $\propto i$ are the parameters which will establish the hyperplane (Karimi et al., 2019; Adams et al., 2020).

2.8.1 Advantages and Disadvantages of Support Vector Machines

The advantage to the SVM model is that the kernel function allows for great flexibility as the classification threshold need not to be linear or the functional form for all the data need not be the same as the function is non-parametric and also operates locally. In addition, the kernel retains implicitly non-linear transformation making data linearly separable, making the model robust and does not require human intervention beforehand. SVM also provides good sample generalization even though the training data contains some bias contributing to the robustness of the model (Tharwat et al., 2017; Shukla, 2020).

The disadvantage of SVM is that lack of transparency making it difficult to understand results. This is common for non-parametric models (Lin, 2019; Feng et al., 2019).

2.8.2 Application of Support Vector Machine

A customer dataset was utilized with an SVM model to segment customers into churners and non-churners. The segmentation was then utilized to establish the probability to upsell or not. The prediction showed a high level of accuracy with minimal errors. Predictions were produced with no explanation to justify the customer segmentation (Lian-Ying, et al., 2019; Jozdani, 2019).

In a cancer study using a large set of data containing various prognostic and diagnostic features SVM perform very well in identifying certain genetic relationships found in subjects diagnosed with cancer. The model was able to classify patient's molecular signatures quite accurately and speedily (Wang and Chen, 2020; Wadkar et al., 2020).

Within the service aspect of a business, customer responses to services received gives companies invaluable insights as to what the customer perceives from the business. Forming part of the direct marketing campaign, the SVM model was used for the customer profile data to predict which customers will provide a low response rate. The dataset included monetary, service frequency and schedule data (Che et al., 2020; Song and Alkhalifah, 2020)

2.9 Artificial Neural Network

Artificial Neural Network (ANN) is an algorithmic model that is based on the way the human brain works using inputs and transforming this input into output signals. The input signal is received by a neuron and weighting is assigned. The weighting is then calculated and processed to formulate the prediction as an output (Zurada, 1992; Bayat et al., 2019; Akhgar et al., 2019). Artificial Neural Network are also described as a physical cellular network with the ability to store, acquire and use experimental data (Zhang, 2016; Nasser et al., 2019).

The Artificial Neural Network processing architecture consists of a vast number of highly and simple interconnected neurons that were inspired by the makeup of the cerebral cortex of the human brain which also consists of neurons as in Figure 2.9 (Abu Dalffa et al., 2019). Over time, research has shown that the traditional computer system did not do certain things well which was found in humans and animals. The development of the ANN aimed to do what humans and animals do well which computers do poorly. Humans and animals adapt well to experimental data, adapt easily to complex data and are able to store complex patterns for future reference (Asteris and Mokos, 2019; Cavaleri et al., 2019).

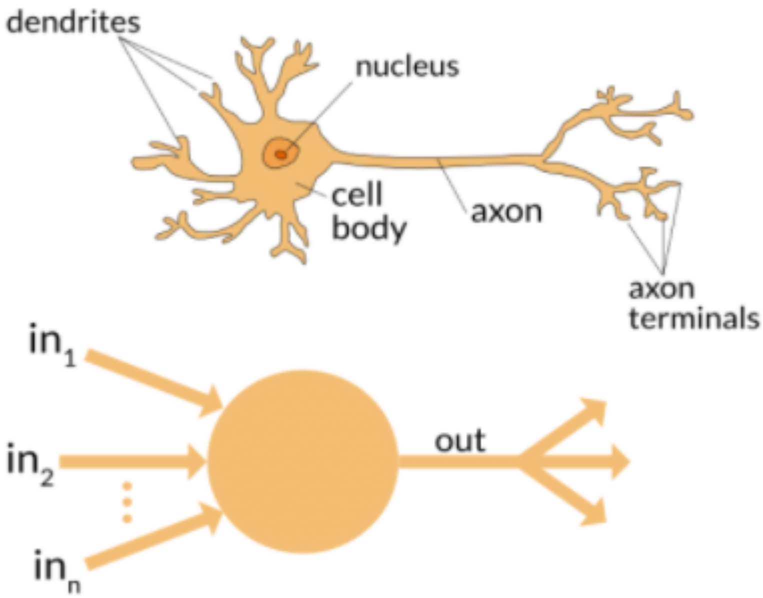

Figure 2.9: Structure of a neuron network (Abu Dalffa et al., 2019)

The ANN model in general computes the output by processing multiple input vectors X. The net input function comprises of input variables x_1, x_2, x_3……x_i which are multiplied with the coefficients w_{ji} which is known as weights. The transferring of the net input transfers the signals through the neurons (Baroni, 2020; Kimaev and Ricardez-Sandoval, 2019).

The following shows how the output y_n, on the n_{th} neuron is calculated:

$$Net_n = m\,\Sigma\, w_{ni}x_i$$

Net_n is then transferred as an input argument into the transferring function

$$yn = out_n = \ 1/\{1 + \exp[-a_n(Net_n +\ \theta_n)]\}$$

The commonly used ANN algorithms include the Multi-Layer Perception (MLP) and the Radial Basis Function Network (RBFN). The RBFN is commonly utilized in biological systems and uses radial basis functions to solve interpolation problems. The model was first used to illustrate a non-linear relationship and show how a solution to an interpolation problem can be implemented (Cai and Liu, 2018; Baroni, 2020).

The RBFN model consists of an input, hidden and output layer. However, the model only has one hidden layer unlike the MLP which can consist of many hidden layers. This hidden layer consists of nodes and the radial symmetric basis function is utilized to activate the function of the nodes (Maind and Wankar, 2014; Niedbala, 2019).

The Gaussian function is the most commonly used function to activate the nodes with the network. It has been noted the choice of function does not necessarily improve the performance of the network. The Gaussian function can be expressed as follows:

$$\varphi\, j = [\, x - cj2 \,/\, \sigma\, 2\mathrm{j}]$$

Where $\varphi\, j$ represents the output the j node with the hidden layer, x represents the input vector and c is the centre (Maind and Wankar, 2014; Cavaleri et al., 2019).

2.9.1 Functionality of the Artificial Neural Network

The Artificial Neural Network consists of hundreds of processing units all interconnected in a complex network. A unit or node is a simplified version of a neuron, which receives an input signal from a connecting node and sends it off to another node as shown in Figure 2.10 (Al-Shawwa et al., 2019; Bui et al., 2020).

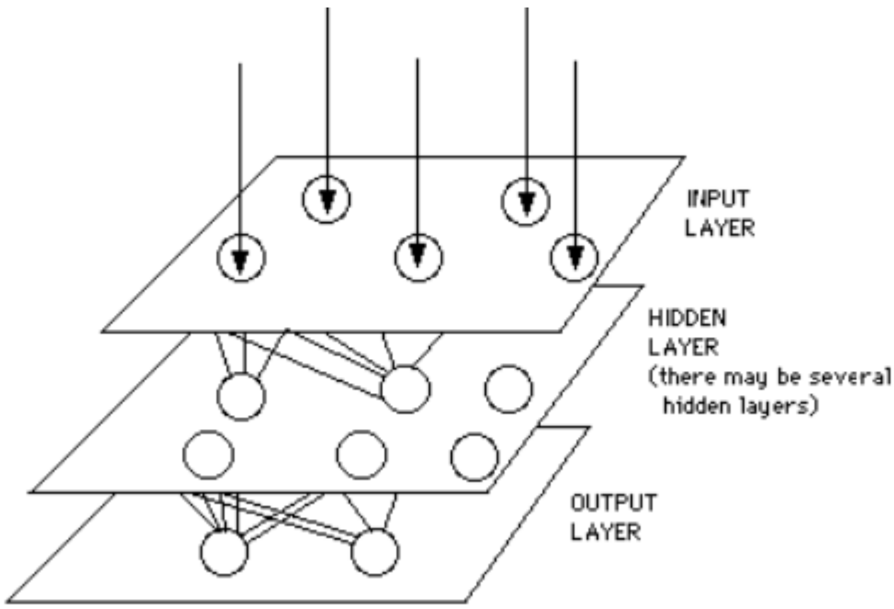

Figure 2.10: The basic structure of a node

The ANN model basically has three layers, the input, hidden and output layer. The input layer neurons receive data from a data source such as an input file or in some real-time applications it receives data directly from a device. The hidden layer neurons are where a weighting is placed on a decision which was based on the input data and sent onto the output layer. The output layer sends data to another computational process or another device (Naderpour et al., 2018; Pham et al., 2020)

The crucial part of the ANN is the way the neurons communicate with one another. The model relies on connections to communicate between neurons. The first connection is responsible for the ability sum from one neuron to another. The other connection is responsible for the subtract function between the neurons (Castanedo, et al., 2014; Rem et al., 2019).

The ANN model also allows for a neuron to restrict another neuron in the same layer which is normally commonly utilized in the output layer and is known as lateral inhibition or competition (Asteris et al., 2019; Sharifzadeh et al.,. 2019). As an example, this works as follows, when taking a text recognition problem and the probability of the character being an “A” is 80% and the probability of the character being an “F” is 60%. The model will select the one with the highest probability and inhibit the rest. It also allows for a connection known as feedback, where a neuron from one layer is able to communicate with a neuron in a previous layer as in Figure 2.11 (Dongare et al., 2012; Coutinho et al., 2019).

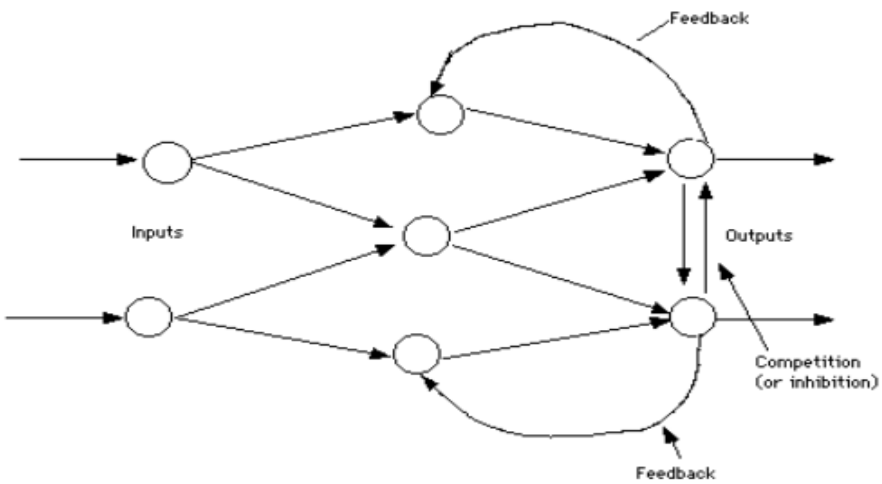

Figure 2.11: ANN communication structure

2.9.2 Advantages and Disadvantages of the Artificial Neural Network

The advantage of the ANN model is flexibility to data consisting of drastic change as it has greater data pattern recognition when compared to other popular models. With the ability to execute parallel computational processes the model is a great fit for real-time processing and the model performs well with complex data sets. Fault tolerance is high as one or more corrupt neurons will not prevent output and the model has been known to be able to work with incomplete data in certain circumstances (Ahmad et al., 2014; Kuo and Huang, 2018).

The disadvantage of the ANN model is that it has a high dependency on hardware as it requires processors that allows for parallel processing and its behaviour is unexplainable as it does not give reason for the output (Mijwel, 2018; Toghraie et al., 2019).

2.9.3 Applications of the Artificial Neural Network

The ANN model proved to be of great value in the application of implementing an automated call answer solution where the model was able to identify the waving of a hand to answer the call. Another hand gesture application which used the ANN model was an application which identified left and right-hand gestures to page through an electronic book and scroll web pages (Keane et al., 2019; Bayat et al., 2019).

In a study where an Artificial Neural Network (ANN) model was utilized to establish if a customer was a potential churner. The results showed that accurate predictions were produced, but an absence of explanation prohibited a better understanding of customer behaviour to decrease the churning rate in the long term, proactively (Molchanov et al., 2017; Al-Shawwa et al., 2019).

2.9.4 Deep Neural Networks

A greater capitalization of the Artificial Neural Network exists, and it is known as Deep Neural Networks. With an Artificial Neural Network each node constructs a relationship with another node, and it rates that specific relationship. This all occurs within a single hidden layer, but with Deep Neural Networks this occurs within multiple hidden layers as shown in Figure 2.12 (Convington et al., 2016; Bashar, 2019).

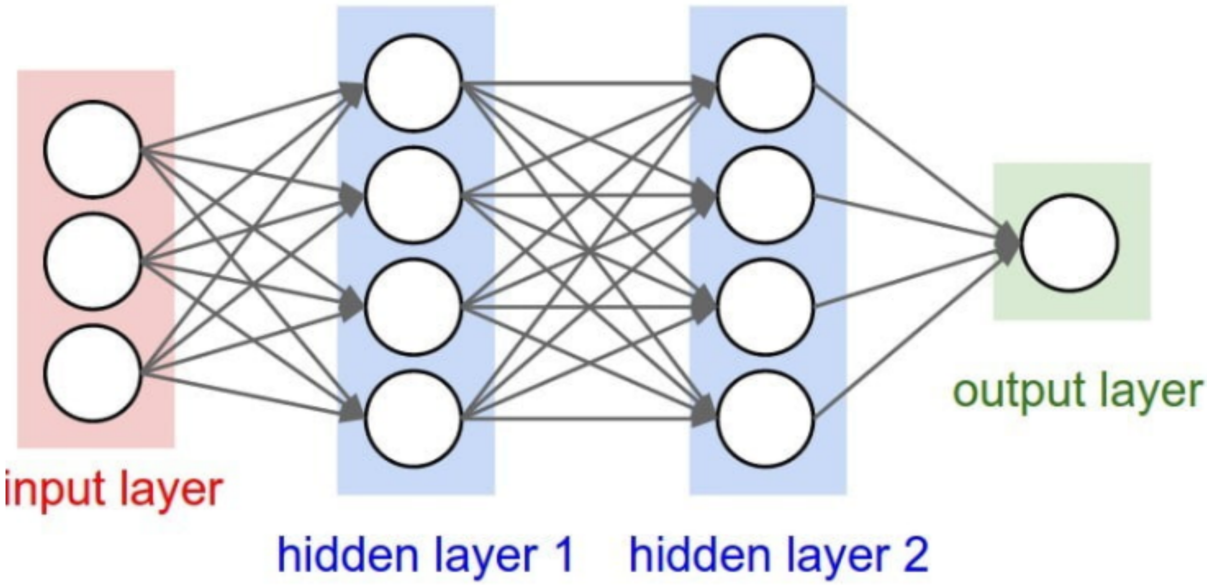

Figure 2.12: Deep neural network structure with more than one hidden layer (Bashar, 2019)

The Deep Neural Network computation are executed by each hidden layer. The output of hidden layers is $h^{(i)}(x)$, the computational process for each network with *L* hidden layers is as follows:

$$f(x) = f[a^{(L+1)}\ (h^{(L)}\left(a^{(L)}\left(\dots\left(h^{(2)}\left(a^{(2)}\left(h^{(1)}\left(a^{(1)}(x)\right)\right)\right)\right)\right)\right)$$

The multiple hidden layers allow for complex data-driven predictions such as self-driving cars, optical recognition systems and advanced web searching processes (Wen et al., 2016; Koopialipoor et al., 2019).

2.9.5 Different Types of Deep Neural Networks

The Auto-encoder is a type of deep neural network that allows for the learning and predicting of coding patterns (Su et al., 2019). The Auto-encoder algorithmic model is made up of an input layer, one or more hidden layers and an output layer (Montavon et al., 2018). The significant functionality of the Auto-encoder is that output layer has the same number of nodes as the input layer. The Auto-encoder functionality is to utilize the output layer to predict the input layer (Too et al., 2019; Li et al., 2019).

The Deep Belief Net is a type of deep neural network that allows for non-convex problem handling and local minima by utilizing the multilayer perceptron (MLP) (Ruder, 2017; Barragan-Montero et al., 2019). The algorithmic model consists of multiple layers with connected latent variables in between the layers. Subnetworks are constructed with hidden layers which function as input layers and the lowest hidden layers function as a training set to the next layer (Durstewitz et al., 2019). This allows for each layer within the network to be trained independently and has proven to be very effective in acoustic modelling (Wongsuphasawat et al., 2017; Ancona et al., 2019).

The Convolutional Neural Network is also another type of deep neural network and functions by assembling small pieces of data and moving then deeper into the network of multiple layers (Sun et al., 2018). The model individual neurons are assembled in a way to react to overlapping when solving visual regional area problems. This response to overlapping is solved by the first layer, identifies the boundary of a region and builds a template for all the boundaries (Shwartz-Ziv and Tishby, 2017; Thulasidasan et al., 2019). The following layer combines the boundary templates and places it into position and the last layer matches the input image to the constructed template for a final prediction (Wang et al., 2020; Hollon et al., 2020).

2.9.6 Applications of the Deep Neural Network

Deep Neural Networks are utilized in a number of really interesting fields of research. Scientists have deployed Deep Neural Networks in the field of human neural and cognitive understanding. The research aims to explain a human’s cognitive decision making and understanding. The research has sparked a number of healthy and constructive debates from the view of philosophical science, but the main objective is to understand human thinking and reasoning (Cichy and Kaiser, 2019; Li et al., 2019).

With the stock market buying and selling being extremely competitive, a number of top stock market investing companies have applied Deep Neural Networks in predicting stock market prices. Data is collected from a wide variety of data sources across industries and supplied into a Deep Neural Network. This Deep Neural Network functions with common evaluation metrics and other predictive methods. The Deep Neural Network allows for the processing of complexed data structures (Hollon et al., 2020; Jiang, 2020).

2.10 Goal Question Metric

As with many engineering fields, the development of software requires a method to evaluate the effectiveness of what was developed and implemented (Basili, 1994; Solano et al., 2019). Many studies have shown that to measure a developed product or services effectively the metric used needs to be goal specific, applied to the product or service throughout the development lifecycle, and interpreted in the context of the specified environment (Van Solingen et al., 2002; Godin et al., 2019).

The Goal Question Metric (GQM) is an evaluation approach in which the detailed goals are defined by an organization that resides within a specific project (Hussain and Ferneley, 2008; Gallina, 2020). Once the goals are clearly defined within the project, a clear path needs to be defined to trace the goal to the data which will establish the goal. The final step is to provide a context to interpret the data which is related to the goal (Boehm et al., 2005; Kim et al., 2019). The general outline of the GQM approach is a hierarchy structure that starts with the setting of the goal, the questions to assess the goal, and metric to measure the effectiveness of the goal being satisfied as in Figure 2.13 (Gallina, 2020; Ehrlacher et al., 2020).

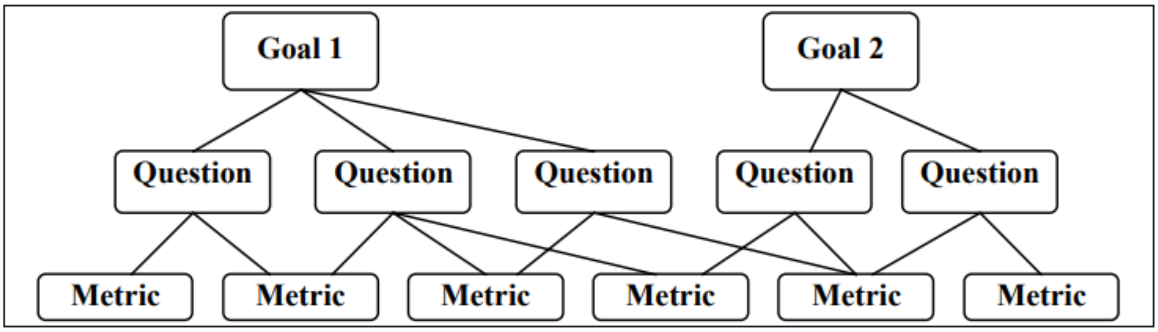

Figure 2.13: Goal question metric structure (Gallina, 2020)

The GQM approach was initiated with the purpose of evaluating defects for a group of projects for the National Aeronautics and Space Administration agency's Goddard Space Flight Centre. The agency had a number of projects in which the goals were defined, the quality of defining the goals improved over time and though the GQM approach was limited to a single environment it quickly expanded into other areas of work (Yahya et al., 2015; Movilla-Pateiro, 2020).

In many cases it is challenging to measure the development of software as the environment includes the supplied data, the project needs and the context is different for many stakeholders (Differding et al., 1996; Franco et al., 2019). The GQM approach allows for a Goal to be setup for each phase of the project, Questions to be answered to ensure that the Goal is met and a Metric to assess how effectively the Goal was met (Rajput et al., 2011; Godin et al., 2019). These Goals, Questions, and Metrics can be defined with all stakeholders collectively to ensure that the different contextual views are covered. The approach also allows for a great ability for project monitoring throughout the software development process (Chicco and Jurman, 2020).

2.10.1 Structure of the Goal Question Metric Approach

A detailed description of the three GQM layers can be defined as follows:

Goal: Within the project, a product, service, or process would be defined as an output. A goal would then be attached to the output in terms of quality or effectiveness from various perspectives but relative to a specific environment (Nagappan et al., 2008; Chicco and Jurman, 2020).

Question: A list of questions is established to assess the delivery of the defined goal. The questions are formulated for a specific product, service, or process. It is important to note the perspective in which the question is being answered and also the environment (Beer and Felderer, 2018; Solano et al., 2019).

Metric: The quantitative data used to answer the question. The data can be objective (only dependent on product, service, or process being measured) or subjective (dependent on the environment, perspective, product, service, or process being measured). It is also important to be aware of the data quality throughout the GQM process (Chicco and Jurman, 2020).

Table 2.2. An example of GQM layout (Van Solingen et al., 2002)

Goal	Purpose Issue Object (process) Viewpoint	Improve the timeliness of change request processing from the project manager's viewpoint
Question		What is the current change request processing speed?
Metrics		Average cycle time Standard deviation % cases outside of the upper limit
Question		Is the performance of the process improving?
Metrics		$\frac{\text{Current average cycle time}}{\text{Baseline average cycle time}} * 100$ Subjective rating of manager's satisfaction

2.10.2 The Advantages and Disadvantages of the Goal Metric Question Approach

The advantage of the GQM approach is that it has been found to be an effective tool to gather an understanding and monitor the implemented software development practices within an organization as it questions every implemented process. It is also for software alignment in the organization as each software deliverable needs to answer a very specific goal (Rahimi, 2013; Beer and Felderer, 2018).

The disadvantage of the GQM approach is that if the establishment of goals is not managed correctly the goal and the measure thereof could become very subjective making it ineffective. The approach has been found to also lack a formal guideline and structure set out by the software development community (Rahimi, 2013; Franco et al., 2019).

2.11 Related Work

Direct marketing over the years has become a method widely used across industries to target specific groups of potential customers. The method assesses customer parameters such as geographical information, spending patterns, and previous customer purchases to construct a product offering directly to a customer. In many areas, this method has been effective in sales increases and customer retention (Alanen, 2016; McCoy, 2019).

A number of advancements in data processing technologies have been made which also allow for the storage of large amounts of data. This advancement which allows for the storage of large data sets enables companies to execute a more detailed and accurate customer analysis. The analysis in turn results in more accurate potential customer predictions. These advanced data sets enable customers to be grouped or segmented in an accurate manner and product offerings can be offered to these customers at the right time (Flici, 2011; Zhang et al., 2019).

The field of data-driven direct marketing has been found to capitalize on the data technological advancement. Complexed data sets are compiled and processed by data algorithms to identify customer patterns to predict the probability that a customer would purchase a specific product (McCoy, 2019). Well-established companies all around the world who rely on data-driven direct marketing strategies are heavily reliant on large data sets and it was found that smaller companies rely less on customer data sets (Beheshtian-Ardakani et al., 2018; Cavaleri et al., 2019).

There are a number of data-driven direct marketing software packages which is currently in use by a number of companies. A number of current software packages utilize social media and emails to gather customer data but there are also a number of other solutions that gathers data from more advanced sources such as Customer Relationship Management (CRM) platforms (McCoy, 2019; Chatterjee et al., 2019).

Other studies have shown that the banking sector has also shifted to data-driven direct marketing strategies. Large amounts of data were gathered from across internal data sources for the construction of customer characteristics. This data was then supplied as inputs to algorithmic models such as Nominal Regression, Bayesian Networks, and Decision Tree models. The output produced by these models were proven to be accurate and effective to the bank, but it was noted that highly skilled resources were needed to interpret the output and to actually configure the models (Elsalamony, 2014; Ancona et al., 2019).

Direct marketing has been proven to yield great results when using machine learning. Machine learning has allowed companies to make more accurate predictions resulting in a saving of time and money (Erel et al., 2018). The commonly used accurate machine learning models are normally structured in a nested non-linear manner. The Random Forest model which is a well-known non-linear structured model is normally applied in a black-box manner with no or very little explanation provided on how the model derived a particular prediction. It is this lack of explanation that makes it challenging to identify faults in the model and any biasness which might exist in the data (Guidotti et al., 2018).

The algorithmic models utilized in machine learning enables data processing and the removal of a lot of speculation that was evident in prior direct marketing solutions. However, these machine learning models do not deliver detailed explanations of why a specific probabilistic prediction was made. It is this absence of explanation and understanding that brings about a level of uncertainty as results are being generated as output, but few know why (Erel et al., 2018; Hollon et al., 2020).

Another study concurred with the problem of complexity surrounding machine learning models but focused on the fact that this difficulty creates a demand for greatly skilled personnel (Fischer et al., 2020).

Case-based reasoning (CBR) utilizes the stored previous occurrences to generate the solution to a current occurrence. The CBR approach includes functionality to have robust explanation mechanisms as it originates its explanations from similarly matched previous cases (Daramola et al., 2013; Lamy et al., 2019).

In many cases, CBR is utilized to address the absence of explanation and simplify the understanding of results. CBR can be utilized to support the predictions produced by the machine learning models, which would complement the machine learning models' functionality in a positive way. The CBR problem-solving process includes retrieval of a case, reuse of a case, and adaptation of a case and retention of a case. CBR examines the previous occurrences with comparable output and provides a comprehensive explanation of the current output. Realistic evidence is utilized from a group of applicable cases to explain the current occurrence of a case. The explanations are simplified irrespective of the complexity of the current problem (Vásquez-Morales et al., 2019; Fei and Feng, 2020).

(Fischer et al., 2020) made an effort to establish the most suited classification technique for data-driven direct marketing. Four commonly used classification techniques were selected and evaluated. The results were indicative that decision trees generated a reasonably high specificity, accuracy, and sensitivity. (Zhao et al., 2019) completed a comparison within the banking industry where four different machine learning models were compared to analyse the success in predicting potential customers for banking products.

The results have proven that the utilization of machine learning algorithmic models enables for the processing of vast volumes of data and retrieving a larger historical view of customer expenditure patterns. The importance was on the accuracy of results and the provision of no explanation. (Lawi et al., 2017) assessed banking data that was applied to the Random Forest Regression and SVM model to produce a prediction of potential banking customers. Both the Random Forest Regression and SVM models have produced a great performance for classification but the Random Forest Regression indicated a slightly improved result for sensitivity and accuracy. It was also noted that both algorithmic models produced results with no explanation in order for the user to get a greater understanding of how and why results were produced as output.

Another study compared the Naive Bayes and Random Forest Regression algorithmic models by utilizing a banking dataset to analyse and determine which customers would likely opt-in for long term deposit products. The Random Forest Regression algorithmic model has proven to produce better output in terms of specificity, sensitivity, and accuracy. However, these models have also produced output with no explanation (Brieuc et al., 2018).

(DeSolo et al., 2019) showed that data classification by machine learning algorithmic models is accurate and effective in establishing which customers will subscribe to long-term saving products within the banking sector. A large customer dataset was labelled with appropriate features, and customers were segmented by location. A Random Forest and Naïve Bayes classification algorithmic model was utilized to categorize the customers into likely and unlikely groups. The Random Forest algorithmic model was found to perform with a higher level of accuracy.

A study has shown that the implementation of machine learning algorithmic models improves marketing considerably. This was tested by implementing a machine learning algorithmic model to a large customer dataset containing various customer characteristics to produce predictions (Bayoude et al., 2018).

(Lian-Ying et al., 2019) made an attempt to identify potential customers to be targeted for upselling certain products within a telecommunication company by the utilization of machine learning algorithmic models. The Support Vector Machine (SVM) algorithmic model was implemented to a customer dataset. The SVM algorithmic model then classified the customers as a non-churner or churner. These segmented groups were then used to upsell products and a special focus was placed on the churner group to possibly convert them to a non-churner. The prediction showed a high level of accuracy with very minimal errors but lacked an explanation as to why a customer was classified as non-churner or churner.

A Deep Learning model was implemented with a telecommunication customer dataset to predict customer churn. It was evident from the findings that predictions were produced with a high level of accuracy with no explanation accompanied by the predictions. This lack of explanation restricted an increased understanding of customer behaviour to decrease the level of churners proactively in the long term (Al-Shawwa et al., 2019).

(Chen, 2016) suggested that the retrieval of customer data from the service, resource and customer layers within a telecommunication company would allow for the construction of an insightful customer profile. The customer billing data which includes customer consumption, usage and spend limit data are also valuable attributes in predicting which customers are potential purchasers of products. It is evident that there is a correlation between current consumption and a customer's past spending pattern. This correlation can be utilized to executing a meaningful prediction at the right time.

A study has shown that combing a case-based reasoning system with a machine learning algorithmic model improves prediction output in terms of the speed of producing a prediction as its knowledge base is built by adding domain knowledge from experts in the form of previous cases. The hybridized system was implemented to predict domain name prices and it did so with minimal hardware allocation, high level of accuracy and speed (Bayoude et al., 2018).

(Hegdal & Kofod-Petersen, 2019) assessed a hybridized CBR-ANN system and has indicated an improvement in results when compared to a standalone ANN algorithmic model. Although there was not a major difference in performance it was noted that the system was applied to a single-use case. Another study that focused on the execution of skin disease prediction by (Dabowsa et al., 2017) utilized a real-life skin disease dataset that was implemented to a system consisting of CBR-ANN has been evident to be feasible with acceptable performance.

(Bayoude et al., 2018) compared and analysed the performance of a standalone CBR and a hybridized CBR-ANN system. The results indicated that an enhanced performance was evident as the ANN model was utilized to enhance the weighting allocation process in the CBR and in many cases the hybridized approach has produced predictions in a more speedily manner due to the utilization of previous cases.

(Lian-Ying et al., 2019) experimented with a hybridized approach. It compared an ANN only, CBR only and a hybrid CBR-ANN system to execute a prediction of property value within residential areas. The hybridized CBR-ANN solutions have indicated a significantly improved performance and ease of usability.

Lastly, although the concept of post hoc explanation for explainable machine learning is currently attracting increasing attention from researchers, so far this has not been applied to the telecommunication domain, which makes the intended contribution of this study to be unique (Vásquez-Morales et al., 2019; Schwabe et al., 2019).

2.11 Summary

In this chapter, the literature was presented on the current landscape of the telecommunication industry within South Africa. This was accompanied by the current data-driven direct marketing tools and the current lack which exists in these tools according to the literature.

A detailed description of a number of machine learning algorithmic models, case-based and rule-based reasoning was presented. This was in the order of a description of the model's functionality, application, advantages and disadvantages.

Lastly, the Goal Question Metric (GQM) approach was presented detailing the functionality thereof, and the contribution it is designed to make to software development projects.

CHAPTER 3
METHODOLOGY

This chapter explains the research philosophy, methodology, and strategy applied for the study. The research strategy is described in detail with each phase being outlined which is also graphically presented by using a workflow diagram. The chapter also provides an overview of the alignment between the set of objectives and the adopted methodology mapping for the study.

3.1 Research Philosophy

Research philosophy describes the belief concerning the methods in which data will be gathered, how the data will be analysed, and how the data will be evaluated in making a conclusion. Research philosophy can be one of pragmatism, positivism, realism, or interpretivism (Moon et al., 2019).

Pragmatism is one that uses multiple research methods using qualitative or quantitative data collection methods. Positivism is normally very much structured, factual, and using large samples of data. In most cases, it applies quantitative data collection, but a qualitative approach can be applied. The idea of realism is heavily dependent on the idea that there is an absolute separation between reality and the human mind. Data collection can be quantitative or qualitative. Interpretivism is mostly applied to small data sets for an in-depth analysis or investigation. Data collection is normally qualitative (Moon et al., 2019).

Epistemology is defined as the description of how someone has come to know something or how one knows the reality of something. This knowledge can be acquired by various human interactions or by reviewing existing literature (Kivunja et al., 2017).

The epistemology of the research was informed by the knowledge and information gained from existing, published literature on the design and implementation of similar data-driven direct marketing tools, technologies, and types of products marketed on data-driven direct marketing platforms.

In gaining an understanding of machine learning algorithmic models and intelligent reasoning systems, relevant books, publications and journals were reviewed. Also, interactions with various field experts enabled the gathering of knowledge of the application of machine learning in the telecommunication domain, effective ways of building intelligent reasoning systems and direct product marketing for data-related products.

The ontology is described as the nature and essence of the truth or reality. It assists with the conceptualization of the nature of reality and concepts which would be analysed to make sense of the research data (Kivunja et al., 2017). The ontological perspective of this study was objectivism as the essence of the truth or reality, defined by the nature of the data gathered throughout the study.

Pragmatism philosophy assumes to be relatively objective or subjective and represents a quantitative or qualitative account. The philosophy allows for the reality of an object to be evaluated in a manner that facilitates a solution to the problem. This is ideal for the construction of a hybridized framework that would be operationalised in a way to allow for facts to be measured quantitatively. The pragmatism philosophy also allows for both quantitative and qualitative data collection. The researcher would be enabled to perform detailed data analysis using various data processing mechanisms and build detailed explanations (Park et al., 2020).

3.2 Methodology

Research methodology describes the specification of procedures and techniques used to identify the data to be used, how the data will be processed and how it will be evaluated. The distinguishing emphasis of the methodology is that it focuses on the techniques (Wang et al., 2019).

A quantitative research methodology was utilized for this study as the majority of the data is of a numerical nature. The data collected throughout the study was evaluated quantitatively using statistical analysis. This quantitative evaluation was done within a design science methodology.

3.3 Research Strategy

In order to meet the objectives of the research, the research strategy describes the manner in which research will be planned, implemented and monitored. In selecting a research strategy, it is important that the selected strategy complements the methodology (Huynh et al., 2018).

There are a number of commonly used research strategies such as an experimental, survey, or case study strategy. The experimental strategy basically assesses the cause and effect relationship between variables which is normally of a quantitative nature. The survey strategy is normally used to capture a group of individual’s sentiments at a specific time. This can include the behaviour or attitude of a specific group. The case study strategy involves the assessment of previous or current cases/situations and applying a theoretical concept in formulating findings related to the case (Huynh et al., 2018).

For the purpose of this research, the design science research methodology (DSR) will be used as the research strategy which entails the design, construction, and evaluation of the hybridized framework in a quantitative manner. A great deal of literature exists on design science methodology and assisted a great deal in applying the methodology. As the solution was an interaction of people, company and complex technology it was important to include a phase that was able to assess the technology's interaction with people and the solution design itself (Kogan et al., 2019).

The design science methodology was an ideal methodological fit to create an artefact (a designed solution aimed at solving a specific problem) and evaluate the artefact effectively. It has been a widely used methodology and previously used with the implementation of network systems, prototyping web applications and educational information systems (Kogan et al., 2019).

3.4 Design Science Research

The Design Science Research framework applied for this study consists of the following activities: explicate problem, requirements definition, artefact design and development, demonstration and evaluation as shown in Figure 3.1 (Johannesson and Perjons, 2014).

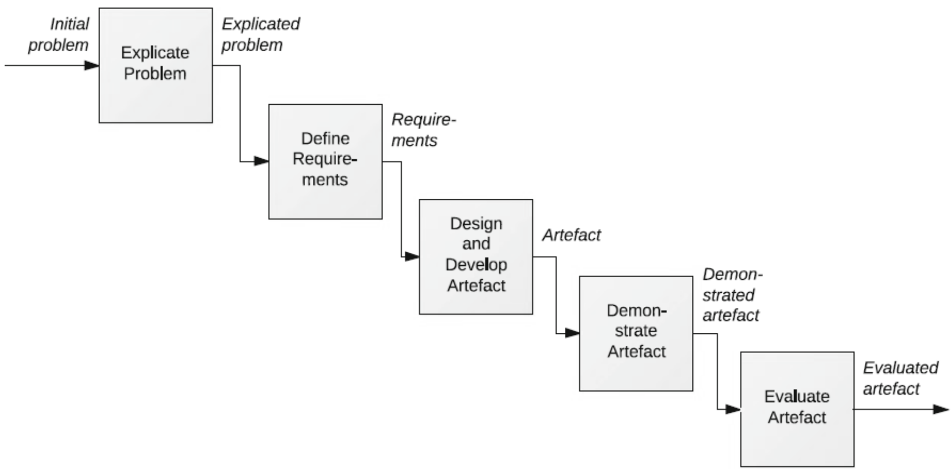

Figure 3.1: Design science research model (Johannesson and Perjons, 2014)

3.4.1 Explicate Problem

This phase of DSR was used to identify the problem surrounding the need for an explainable ML framework for data-driven product marketing systems, documents of current data-driven product marketing tools were analysed. In addition to this, further literature on the explainability of ML-based data-driven products, marketing output was reviewed (Kogan et al., 2019).

3.4.2 Requirements Definition

The requirements definition phase is where the requirements to meet the objectives are defined and an assessment is made as to what is feasible and possible (Kogan et al., 2019).

In achieving this it is good practice to involve a number of stakeholders and experts who can assist in executing a detailed analysis of what can be achieved. The requirements should be concluded rationally from the problem statement (Kogan et al., 2019).

There are a number of ways to execute the requirements definition phase such as one on one or group interviews, questionnaires, surveys or joint application design sessions (Kogan et al., 2019).

The documented problems from the initial phase were modelled by defining requirements using a functional decomposition diagram and use case diagrams. Each requirement importance was evaluated and a priority of 1 – 5 (where 1 is a high priority and 5 is a low priority) was assigned to each requirement.

These requirements addressed the problems in providing an understandable ML framework for data-driven direct product marketing. The output of this phase was a requirement specification document that was used to design and develop an artefact.

3.4.3 Artefact Design and Development

In the artefact design and development phase, an artefact/prototype was developed to satisfy the requirements specified in the requirements definition phase. In this phase, the necessary resources in terms of skilled personnel, hardware and software were used to create a fully functional artefact. The testing of functionality as per specified requirements also took place in this phase (Kogan et al., 2019).

For this study, data collection was done at the resource, service, customer and billing layer derived from the telecommunication industry. The data collected was pushed to the ML framework for data-driven direct product marketing, resulting in a prediction and explanation outputted to a user.

The front-end was developed using Java as this allows for the development of an easy to use front-end. The data was passed from the Java front-end to the Random Forest algorithmic model using the Java Worker API. The architecture included the jCaBaRe Case-Based Reasoning API and the Java Jess rule engine as the intelligent reasoning system.

The developed solution is an easy to use graphical interface for any employee within an organization to use. In turn, this was expected to reduce the skill requirement, capital and operational expenditure. There was a need to include a data processing mechanism and the hybridization consisted of a reasoning and algorithmic mechanism.

3.4.4 Demonstration of the Artefact

In the demonstration of the artefact phase the fully functional artefact was demonstrated to all stakeholders. All resources are available in this phase to ensure that the artefact is able to execute successfully (Kogan et al., 2019).

The artefact was installed on five users' desktops. Once the artefact was successfully installed the evaluation of the artefact was initiated (Kogan et al., 2019).

3.4.5 Evaluation of the Artefact

In the evaluation of the artefact phase all stakeholders were given the opportunity to observe and measure, to gauge how satisfactory were the requirements delivered within the artefact. This phase requires that the technique of analysis is defined and that the correct metrics are selected to measure the delivery of requirements. The evaluation is conducted from the perspective of each stakeholder (Kogan et al., 2019).

To determine if the artefact was successful in addressing the defined requirements, the hybrid ML framework for data-driven direct product marketing was evaluated by performance, usability and the quality of explanation.

In evaluating the hybrid ML framework for data-driven direct product marketing the Goal Question Metric (GQM) was used. GQM proposed a question to a user and a rating was allocated based on the understanding of a goal that was achieved (Achtaich et al., 2019).

The questions included within the GQM were based on metrics highlighted in a study by Gunning (2017). In measuring the effectiveness of explanation, the metric focussed on user satisfaction, mental model, correctability, trust assessment and task performance (Gunning, 2017). Figure 3.2 below shows a graphical view of the GQM and measuring the effectiveness.

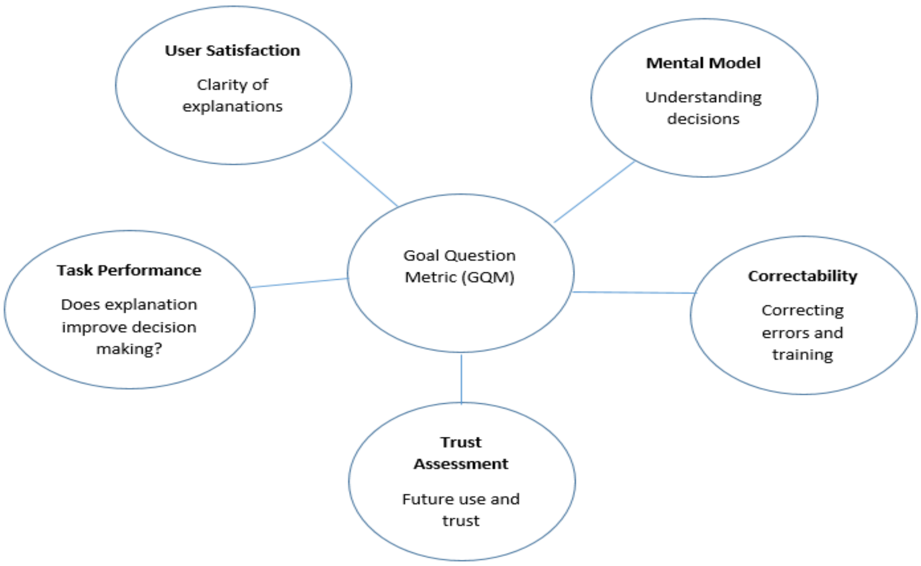

Figure 3.2: Graphical view of the GQM to be used

In Figure 3.3 the research methodology workflow and the output produced by each phase throughout the research process is presented.

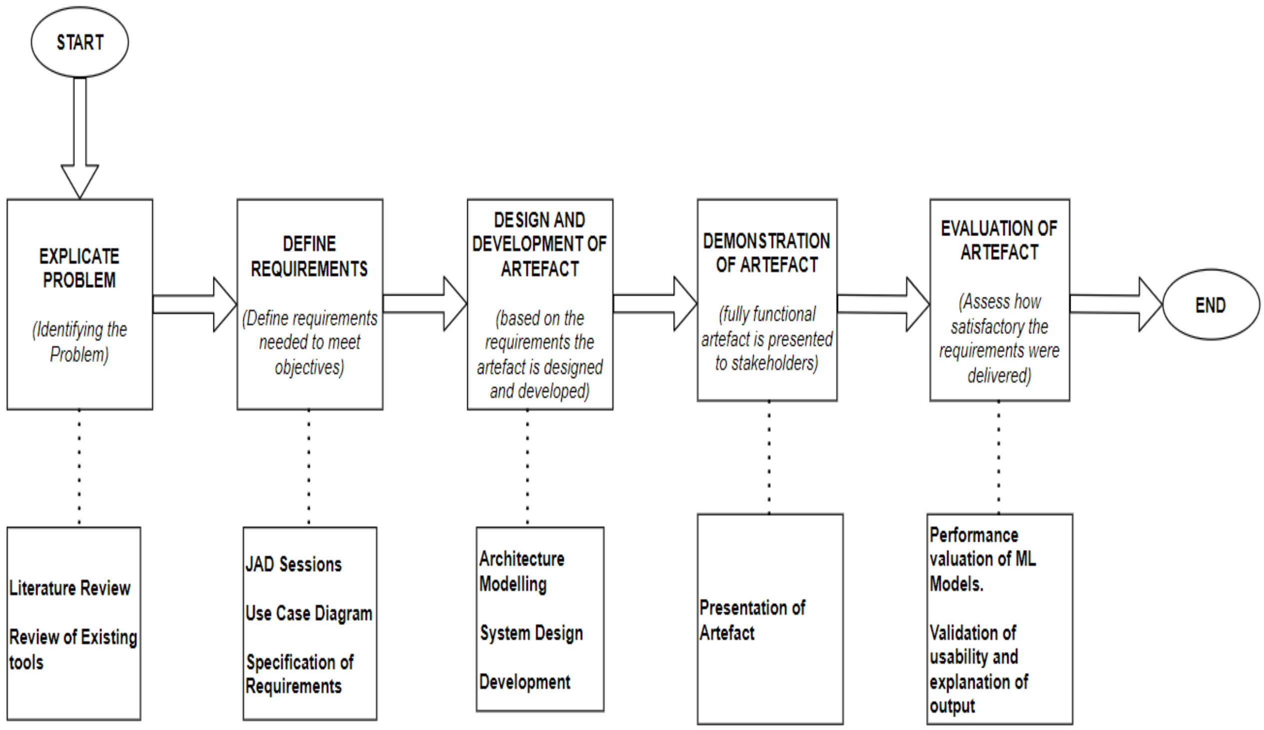

Figure 3.3: Research methodology workflow

3.5 Objectives to Methodology Mapping

Table 3.1 shows how the objectives mapped to each phase of the adopted design science research methodology framework. This provides a clear understanding of how each objective was achieved by each phase.

Table 3.1 Objective to methodology mapping

Research questions	Objectives	DSR phase	Methods/activities/tools
What are the attributes of an easy to use and explainable ML process framework for data-driven direct product marketing?	To determine the requirements for an explainable ML framework for data-driven direct product marketing.	Explicate Problem	• Literature review • Review of existing tools
		Requirement Definition	• JAD sessions • Use case diagrams • Specification of requirements
How can ML and intelligent reasoning be combined for data-driven direct product marketing? How can a hybrid ML framework for data-driven	To design a hybrid ML framework for data-driven direct product marketing.	Artefact design and development	• Architectural modelling • System design • Front-end development in Java. • The development and composition of algorithmic model which includes ANN, SVM and Random Forest • Case-based reasoning using jCaBaRe.

direct product marketing be implemented?			• Intelligent rule-based reasoning system using JESS.
		Demonstrate artefact	Presentation of artefact
How can a hybrid ML framework for data-driven direct product marketing be evaluated for performance, ease of usability and explanation of output?	To evaluate the hybrid data-driven direct product marketing ML framework for performance, ease of usability and explanation of output.	Evaluation of artefact	• Experimental validation of performance of 3 ML models: ANN, SVM and Random Forest using metrics such as accuracy, sensitivity and specificity. Validation of usability and explanation of output by using the Goal Question Metric (GQM): • User satisfaction • Mental model • Task performance • Trust assessment • Correctability

3.6 Summary

In this chapter, the research philosophy included the information which informed the epistemology of the study, an ontology which was one of objectivism and the positivism philosophy which was applied for the study. The overview of the research strategy which is a design science research strategy was also presented. A description of each phase was provided and how each phase was executed in completing the study successfully. This description ended by assessing the evaluation technique which is the goal question metric and how it was implemented in the study.

The chapter concluded by presenting the objectives to methodology mapping which showed the research questions, the objectives which will satisfy the questions, and how the specific phases of DSR correspond to specific objectives.

CHAPTER 4
REQUIREMENTS ANALYSIS AND DESIGN OF THE PROCESS FRAMEWORK

In this section, the requirements and conceptual design of the proposed process framework for direct marketing with explanations are presented. The description of how each phase of the Design Science Research (DSR) strategy was applied by using a sample dataset to design and develop the artefact was presented.

4.1 Requirements Elicitation and Analysis

This section consists of detail retrieved from the execution of the explicate problem and requirement definition phases of DSR. This consists of how the problem was understood and the actual definition of the requirements.

4.1.1 Understanding the Problem

The explicate problem phase of the DSR strategy enabled an understanding of the problem. The methods utilized to achieve this was a literature review and a review of existing tools. The detailed literature review is presented in chapter 2 and the review of existing tools are shown in Table 4.1.

Based on the product review listed in Table 4.1, only 35% of the products reviewed can handle large datasets, 71% of the products reviewed require a high level of skill to setup and interpret results.

It was also noted that 85% of the products reviewed had a rating of very difficult to difficult for ease of use and none of the models allowed for the application of multiple predictive models. Lastly, all the products reviewed showed no explanation to minimal explanation of results.

Table 4.1: Review of existing tools

S/N	Application Name	Manage Large Data sets Yes/No	Skills Level Required Low Medium High	Ease of Use Very Difficult 1 – 10 Very Easy	Explanation of Result No Detail 1 – 10 Extremely Detailed	Multiple Predictive Models Yes/No	Website	Review Platform
1	Direct mail manager	Yes	Medium	10	5	No	www.softwareadvice.com	www.softwareadvice.com
2	MailChimp	Yes	Medium	5	2	No	https://mailchimp.com/	www.capterra.com
3	Aweber	No	High	3	1	No	https://www.aweber.com/	www.capterra.com
4	Constant contact	No	High	4	2	No	https://www.constantcontact.com/global/home-page	www.capterra.com
5	Maxtra technologies	No	Medium	5	3	No	https://maxtratechnologies.com/	www.softwareadvice.com
6	Loss	No	High	1	3	No	https://ioss.in/	www.capterra.com
7	Autopilot	No	High	10	3	No	https://www.autopilothq.com/	www.g2crowd.com
8	Drip	Yes	Medium	5	2	No	https://www.drip.com/	www.softwareadvice.com
9	MoonMail	Yes	High	1	1	No	https://moonmail.io/	www.softwareadvice.com
10	ActiveCampaign	No	High	3	3	No	https://www.activecampaign.com/	www.softwareadvice.com
11	Zift Solutions	Yes	High	2	1	No	https://ziftsolutions.com/	www.g2crowd.com
12	Agency analytics	No	High	1	1	No	https://agencyanalytics.com/	www.g2crowd.com
13	Post affiliate pro	No	High	1	3	No	https://www.postaffiliatepro.com/	www.g2crowd.com
14	TargetEveryOne	No	High	1	3	No	https://www.targeteveryone.com/	www.g2crowd.com

4.1.2 Specifying the Requirements

The requirement definition phase within the DSR strategy allowed for a detailed gathering and definition of requirements to satisfy the objectives. To ascertain the requirements which would form part of a fully functional data-driven direct product marketing framework, the Joint Application Development (JAD) technique was applied.

JAD allowed for input from all stakeholders, development teams and application users. This allowed for each requirement to be analysed by participants and stakeholders. The requirements were established from all functional and non-functional sides which related to the entire system and identified use cases. Sessions were conducted over four days in one hour sessions until all concerning issues have been deliberated and data collected. The JAD session comprised of 9 participating members as below:

- Specialist in Business Intelligence (1 person)
- Specialist in Customer Experience Management (2 persons)
- Specialist in Predictive Analytics (2 persons)
- Customer Care End User (3 persons)
- Customer Care Supervisor (1 person)

The JAD sessions have also shown that different user profiles are needed within the framework. A senior user, user and user admin profile were created within the framework. A user's interaction with the framework is dependent on the profile of a user. Figure 4.1 illustrates the different types of user-profiles and how these user profiles are allowed to interact with the framework, based on the defined requirements.

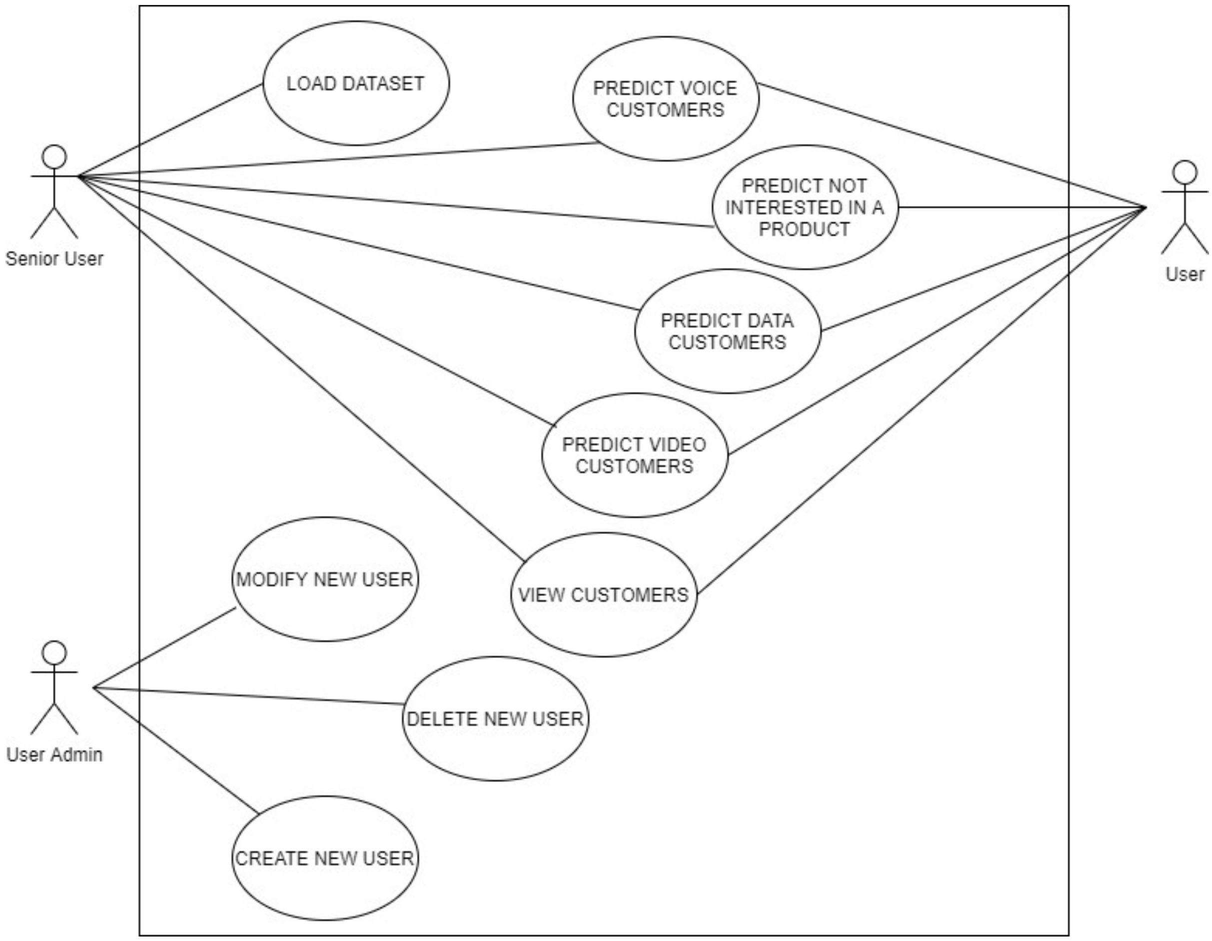

Figure 4.1 Use case diagram showing the user interaction with the framework

During the JAD sessions, participating members were given the chance to comment on the different types of interfaces the customer and user would use to interact with the system. These interfaces included both front-end and back-end resources. Participants then identified how a customer's purchasing pattern can be influenced positively. Participating members were also tasked to provide detail on any problem which might be identified as a challenge in customer interaction or any feature which has the potential of influencing a customer's purchasing pattern. By utilizing the brainstorming process which involved all participants, solutions were detailed for all the problems which were then converted to the requirements. At the end of the JAD sessions, the following requirements of the system were derived, and the requirements definition was executed with Table 4.2 as the output.

Table 4.2: List of requirements

Req. No.	Description
1	Prediction of customers most probable to be interested in a product accompanied by an explanation for the selected customer.
2	Prediction of which products to be used to upsell.
3	Prediction of which customers to exclude for a product offering. Explanation to be included.
4	Predictions of the system must be produced in a reliable and speedily manner.
5	Customers should be placed into categories of low and high spending customers based on the amount they spend per month.
6	Customers should be placed into categories to identify low and high usage customers. This should be based on the amount the usage is per month.
7	The explanation must be of a high quality and easy to understand.
8	The solution should be developed in a way that code can be used in various scenarios.
9	Due to the system containing customer data, the system has to include a login functionality to only allow authorized users to the system.
10	A product list must be available for selection before the prediction task.
11	The system must display the executed prediction and detailed explanation to the user.
12	The system should allow the loading of products to be sold to customers on the system.
13	It should be possible to create and delete user accounts.

4.2 Design and Development of the Process Framework

The Design and Development of the Artefact phase was the construction of the artefact and a data sample was used to confirm the functionality. The section presents the architectural modelling and detailed design utilized in this DSR phase.

4.2.1 Architectural Modelling

The architectural modelling section provides a high-level view, the architecture and the process workflow of the framework.

4.2.1.1 General High-Level View of the Framework

From the list of identified requirements, a hybrid ML framework that can enable data-driven direct marketing with an explanation was formulated. The proposed framework enables data retrieval from 4 layers of a telecommunication organization, which are the service, resource, customer and billing layers. The high-level view of how customers are targeted for direct marketing using the hybrid ML framework is shown in Figure 4.2.

The resource layer includes customer and device data to confirm the technologies available to the customer. The service layer contains the data, voice and video services utilized by the customer during six months. The customer layer includes any interaction the subscriber had with the organization and the service provisioning data; while the billing layer includes data on spending and usage of data. Once the data collection process was completed successfully, the data was handed to the hybridized machine learning system that executed the prediction and produced an explanation.

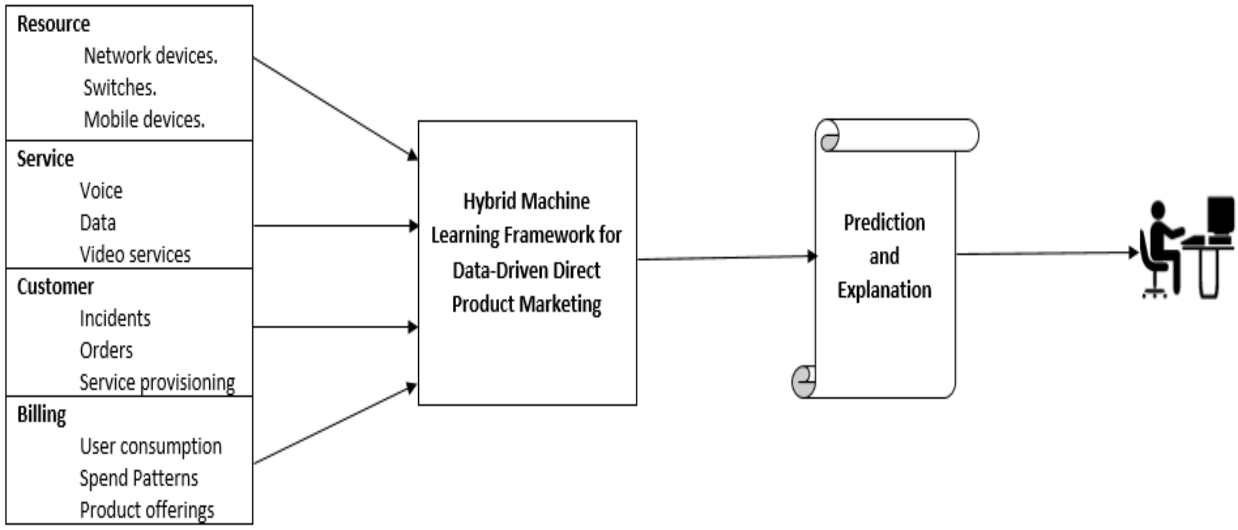

Figure 4.2: The high-level view of the hybrid ml framework for explainable direct marketing

4.2.1.2 The Architecture of the Hybrid Machine Learning Framework

The architectural structure of the Hybridized Framework for Explainable Direct Marketing is made up of ML algorithmic models, CBR and rule-based reasoning. This hybrid system provides a complete range of processes which include the graphical user interface, data processing, middleware, outputs and the CBR knowledge base.

The python-based ML algorithmic models were utilized to leverage the supervised learning abilities for prediction execution. The ML algorithmic models included Random Forest (RF), Support Vector Machines (SVM) and Artificial Neural Networks (ANN). The myCBR Restful API was utilized as the foundation to extract CBR functionalities.

The hybridized framework consists of distinctive interfaces to support all forms of user interaction. These interactions included the selection of a particular direct marketing task, executing the data processing and the particular type of ML algorithmic prediction, utilizing the various algorithmic models and retrieving explanations by utilizing the intelligent reasoning knowledge properties. The components of the integrated hybridized ML framework architecture are presented in Figure 4.3. The data and feature selection were initiated from the Scikit-learn features and the myCBR feature. These features in turn generated the prediction and explanation (Petersen and Daramola, 2020).

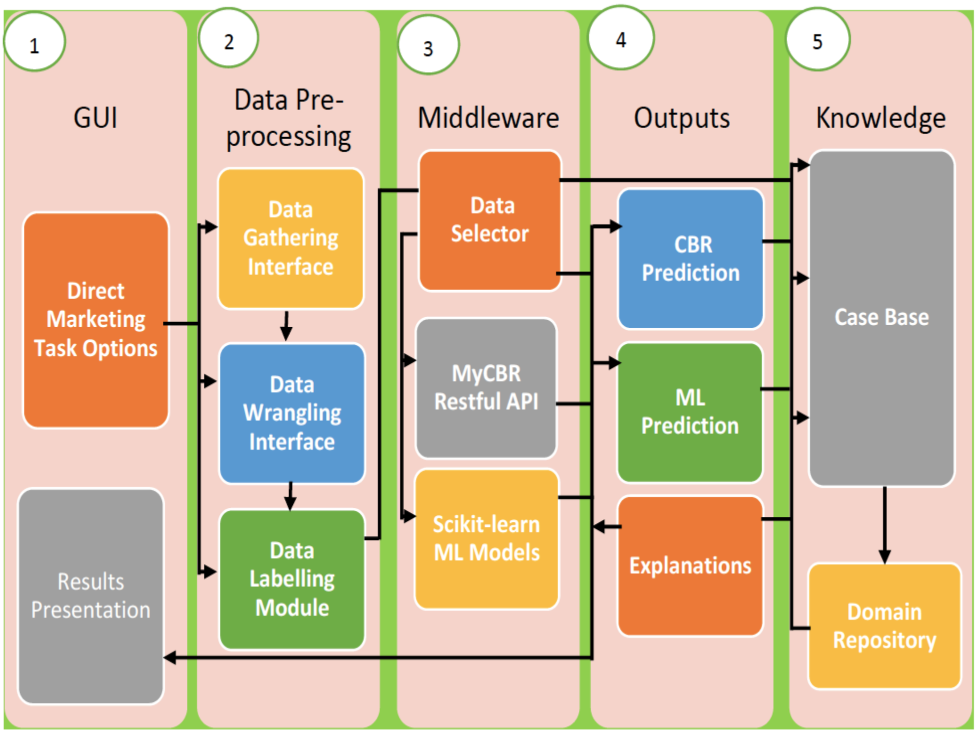

Figure 4.3. Components of the hybrid ML framework architecture

4.2.1.3 Process Workflow of the Hybrid Framework

The process framework for the hybridized framework consists of two phases, one phase is for the supervised machine learning algorithmic models and the other phase for the intelligent reasoning which is made of case-based and rule-based reasoning. The two-phased process framework is illustrated in Figure 4.4 with sequential steps from 1 to 15.

The first phase is made up of steps 1 to 4 and is mainly focussed on the preparation of the data – (1) Data gathering was the centralizing of all the data into a single dataset (2) Data wrangling which involved removing any duplicated or null values – (3) Data labelling which involved tagging the data to assist in predictions later in the process – (4) are majority automated and procedures. Data selection involved determining the features that were needed for the prediction task at hand – (5) in training a nested non-linear ML algorithmic model such as SVM, RF or ANN are applied to the dataset. The training of the algorithmic models was executed as an offline process separate from the prediction execution activity.

The second phase of the process framework includes the generating of the prediction and an explanation that is initiated by the query. For the generating of prediction, as the data from the new query or case – (6) is passed into the process framework, the feature selection is executed – (7) this occurrence is to ensure that the necessary data features are retrieved from the new case or query. Then the matching of a case and retrieval of a case from the knowledge base

– (8) occurred to establish if the new occurrence of the case is alike to previous cases. If that the new occurrence of the case is considerably alike to previous cases, then the CBR module– (9) by utilizing predefined domain rules – (10) was utilized to produce a prediction and necessary explanation – (11).

When a sufficient match with previous cases in the knowledge base is not found, the ML algorithmic prediction was executed – (12) executes the prediction. This functionality proposes three possible predictions. Firstly, which customers are potentially interested in a new product, secondly which customers are not interested in a new product, and thirdly which customers are possible candidates for upselling? Once the prediction was executed successfully by the supervised ML algorithmic model – (13) the CBR module – (15) will search the knowledge case base – (14) for case instances that might have similar features to the current incoming case, and utilize the case to build a suitable explanation for the new case – (11) by depending on the pre-existing domain knowledge rules that are stored in the repository – (11). The alternate processing paths can be explored in the framework as shown in Figure 4.5.

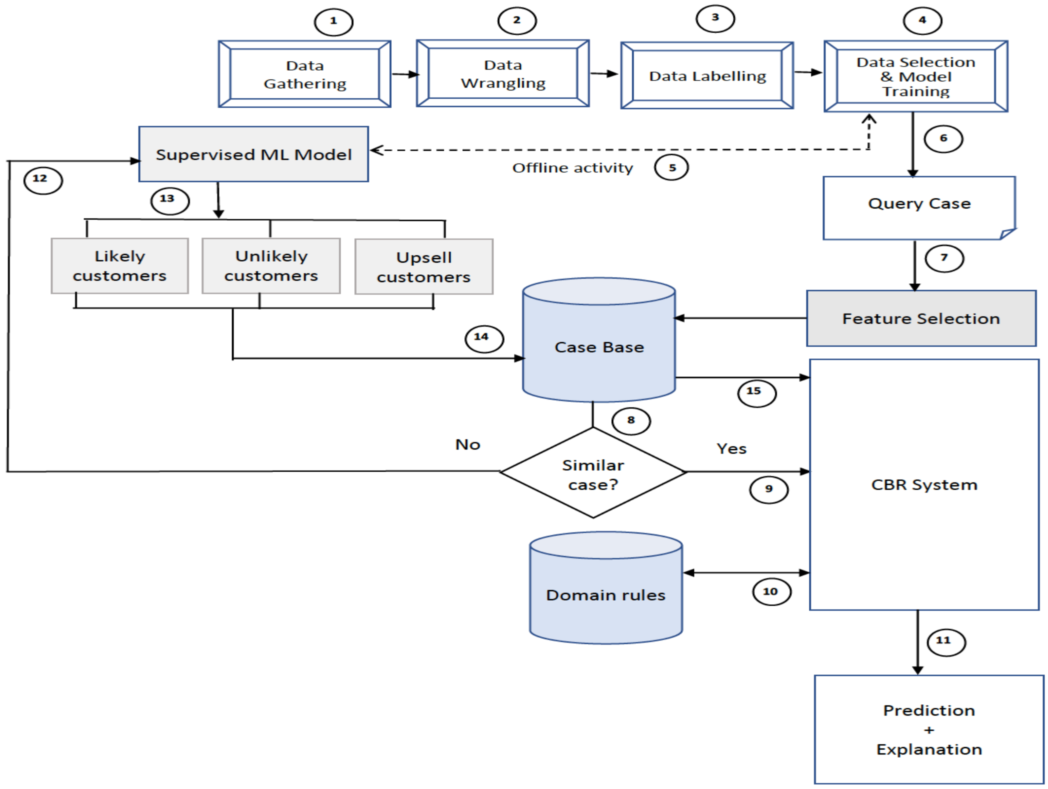

Figure 4.4. Overview of the process workflow of the hybrid ML framework

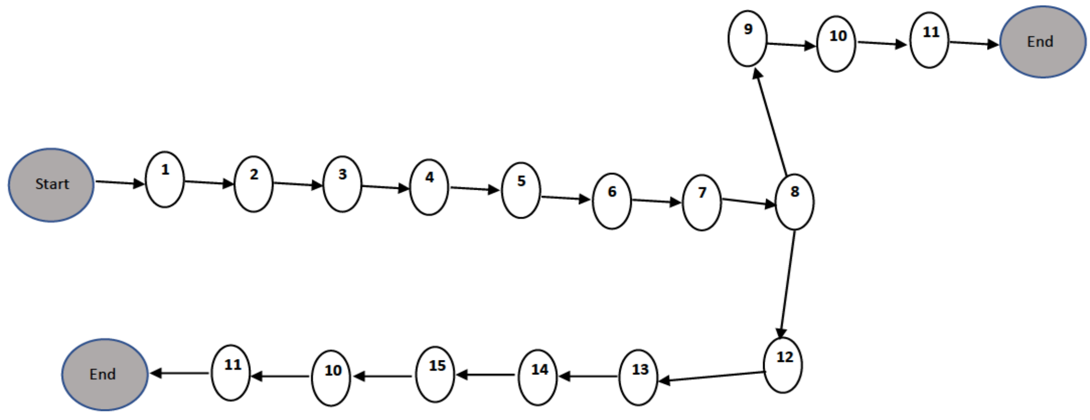

Figure 4.5. The paths communication during the process workflow

4.2.2 Detailed Design

This section details the design of the entities, classes and communication among key components of the framework to ensure successful functionality.

4.2.2.1 Entity Classes of the Process Framework

The framework was constructed for users with different user profiles assigned to a user to load data, view products and predicting which customers are potentially interested in a product. Figure 4.6 shows the design of the various classes and how these classes interact with each other.

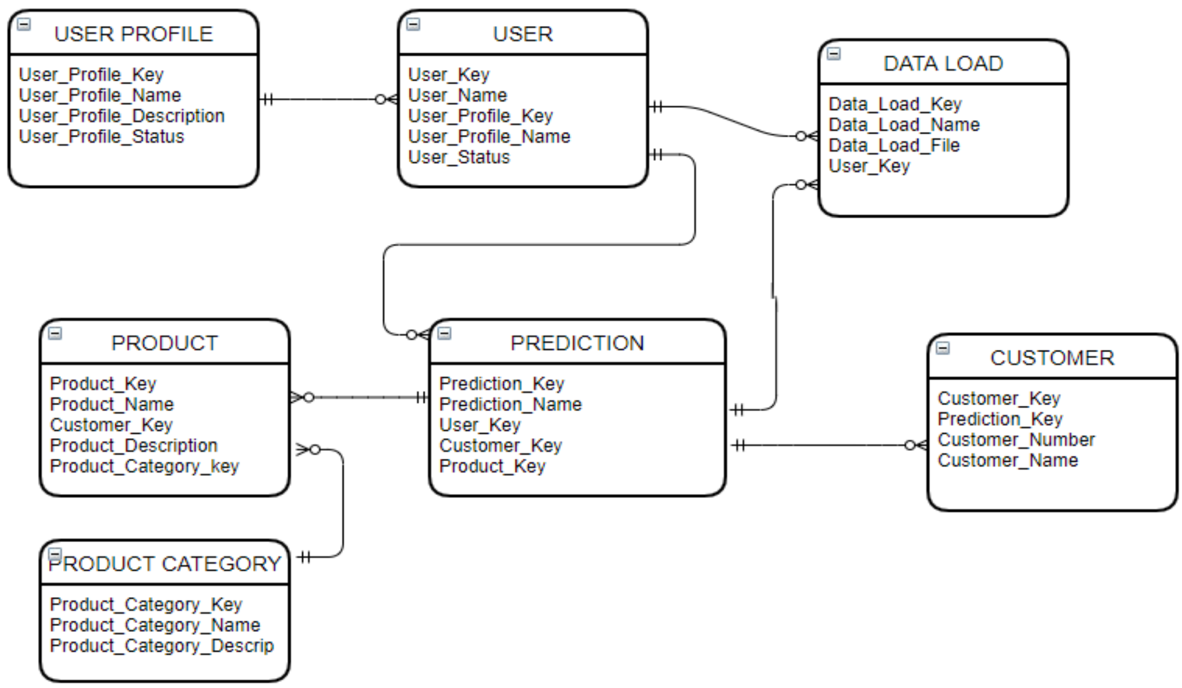

Figure 4.6. Entity class diagram of the hybridized ML framework

4.2.2.2 Communication and Interaction of Entities of the Process Framework

The framework allows various forms of communication and interaction among entities. Figure 4.7 shows the design of how the user interacts with the framework. The communication activity includes login, load data, prediction and functionality for administering users.

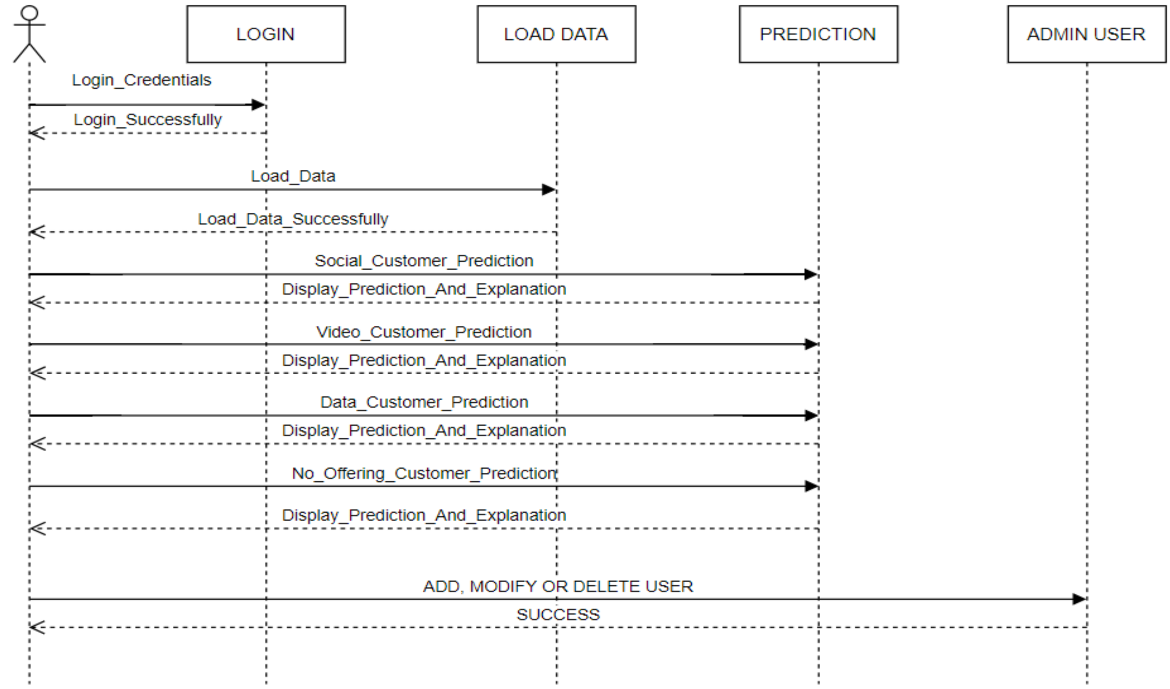

Figure 4.7. UML communication diagram showing interaction of entities of the framework

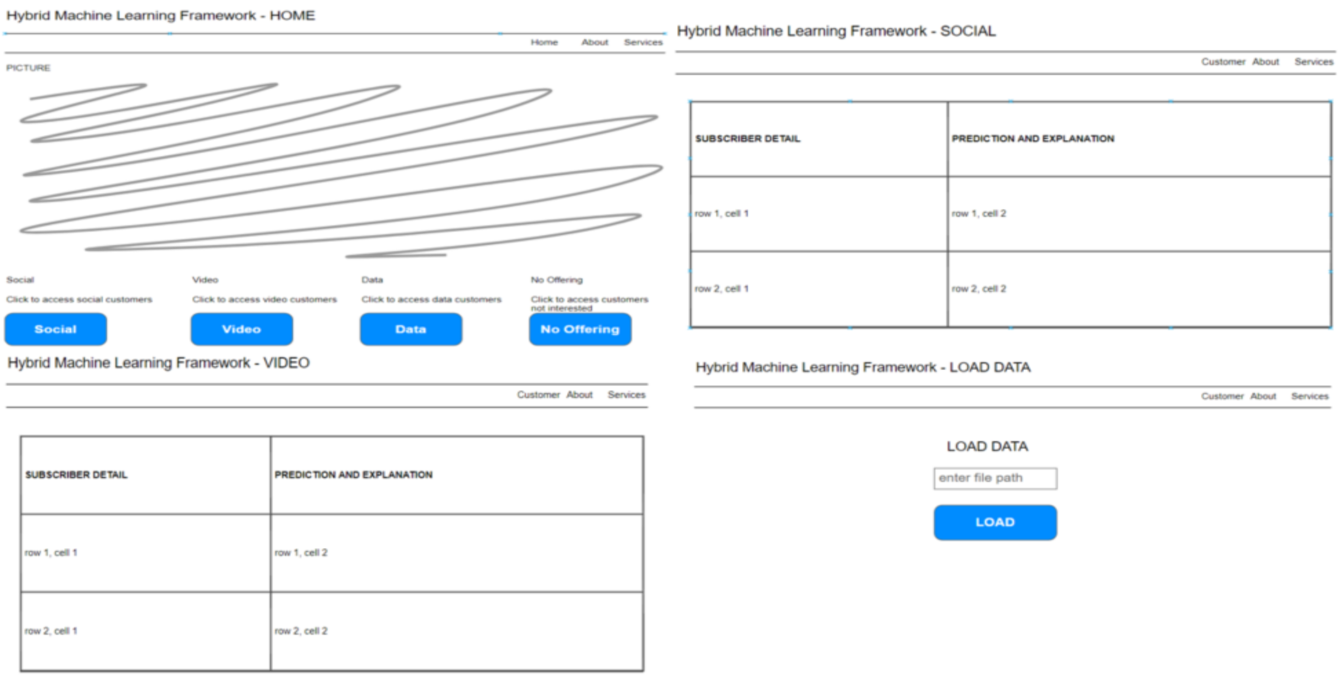

Figure 4.8. User interface design in wireframe

4.3 Prediction and Explanation via the CBR Module

The design of the case-based reasoning module specified that the component operates by executing a comparison of the features extracted from the new case and the features of the previously occurred case which was stored in the knowledge base. The extraction of cases was done in the form of a query and the comparison of features was done by the use of the K-Nearest Neighbour algorithm that uses a cosine similarity metric. The K-Nearest Neighbour algorithm assigns a weighting to each feature retrieved from the two cases (the new case and the one retrieved from the knowledge base).

Once the weightings have been assigned a comparison is done to assess the similarity between the two cases. In the study, a threshold of 0.8 or higher was adjudged to be a significant similarity. This will enable the most similar cases within the knowledge base to be selected. In addition to the cases being selected, the domain rules were used in conjunction with the attributes to generate an explanation to be accompanied by the prediction. On the other hand, if the similarity was less than the threshold of 0.8 the ML algorithmic model was used to generate a prediction. The algorithm for case selection based on the similarity between a new case and the previous, and case prediction by the CBR is represented by algorithm 4.1.

Algorithm 4.1: The CBR prediction algorithm

Function CBR_Prediction
Input: Case Base: CB; Query case: Q_c
Output: CBR_prediction: number ≥ 0;
new_features = $\{x_1, x_2 . x_n\}$ of Q_c // set of features 1. n of a query (new) case
CB= $\{C_1, C_2 C_k\}$ // set of cases contained in C
Selected_Cases ← { } // null set of cases
Case_features = $\{s_1, s_2 .s_n\}$ of each $C_i \in$ CB
W = $\{w_1, w_2 . w_n\}$ // set of weights associated with specific Case features where $w_i \Rightarrow s_i$

If not empty(CB) **then**
  for each $C_i \in$ CB
    //Compute Similarity between Querycase_features and an old _case_features
    Simscore← Cosine_Similarity (WC_i, WQc) // which is $\frac{WC_i . WQ_c}{\|WC_i\| X \|WQ_c\|}$

    if Simscore >= 0.8 **add** C_i to Selected_Cases
  end for
   CBR_prediction← CaseAdaptation(Selected_Cases)
endif
 return CBR_prediction
end function

4.4 Prediction and Explanation via the ML Algorithmic Model

Once the similarity threshold of the new case and existing case was determined and are less than 0.8 the trained algorithmic model was utilized to produce a prediction as output. After the generation of the prediction, the case-based reasoning component utilized the knowledge in the case base repository to establish an explanation for the new case.

In achieving this, the most recognized and discriminant attributes for the ML algorithmic prediction output were used as a foundation to determine the previous cases which were most similar to the new case and retrieve the features from the case base. By the use of an exploratory operation to search, separate discriminant features of multiple cases were scrutinized to establish occurrences of similar cases. These were then utilized in conjunction with the domain rules to establish an explanation to be accompanied by the prediction obtained from the ML algorithmic model. The sample set of domain rules were utilized to establish an explanation as shown in Table 4.3. By utilizing serveral predefined textual templates, significant explanations based on known imperative case features and the domain rules resulted in the construction of cases that supported the prediction made by ML algorithmic models. The algorithm showing how the prediction and explanation were established using the machine learning model is represented by algorithm 4.2.

Algorithm 4.2. The Machine Learning Prediction Algorithm

```
Function ML Prediction ( CBR_prediction: number)
Input: Predefined domain rules: Domain_Rules; Query case: Qc;
trained_RF_model; trained_SVM_Model; trained_ANN_model
Output: ML_prediction: number ≥ 0; Explanation: String

If CBR_prediction = 0 then

    If model_selected = 'Random_Forest' then
        ML _Prediction ← ComputeRF(Qc) // compute Random Forest (RF) prediction
        Key_Features ← Extract_Keyfeatures(trained_RF_model)
    end if

    If model_selected = 'SVM' Then
        ML _Prediction ← ComputeSVM(Qc) // compute SVM prediction
        Key_Features ← Extract_Keyfeatures(trained_SVM_model)
    end if

    If model_selected = 'ANN' Then

        ML _Prediction ← ComputeANN(Qc) // compute ANN prediction
        Key_Features ← Extract_Keyfeatures(trained_ANN_model)
    end if

    Explanation ← Construct_Explanation (Key_features, Domain_Rules)

    return prediction + Explanation

end if
end function
```

Table 4.3: Domain rules for explanation generation

Rule No	Condition	Offer
1	If high data and low spender receiver and streaming capability	YouTube product
2	If high data consumption and medium spender and WhatsApp usage	WhatsApp product
3	If long-time user and low spender	Data product
4	If high data consumption and high spender and streaming capability	Data product
5	If high data consumption and high spender and high streaming	Video streaming product
6	If medium data consumption and high spender and high streaming	Offer YouTube product
7	If low spender and low consumer and short-time user	No offering
8	If high data sender and high spender	Data product
9	If high streaming and high spender	Video streaming product
10	If high data consumption and low spender and streaming capability	Video streaming product

4.5 Summary

In this chapter, a presentation was made on how the DSR strategy was applied in terms of the explicate problem, the requirement definition, artefact design and development phases. The discussion started with a description of the reviewing of existing tools provided with a detailed description of how the JAD sessions were executed and the output thereof. The chapter was concluded with a detailed description of the design of the framework which included the architectural modelling, system design and the algorithms used for the prediction.

CHAPTER 5
IMPLEMENTATION AND EVALUATION

A sample scenario is presented detailing sample instances of subscribers to be utilized for direct marketing to identify potential customers for a YouTube product offered by a telecommunication organization (denoted as XC). The study selected a group of customers from the XC customer base which was identified for direct marketing from a 25 million customer base. Sample customer data are utilized and retrieved based on data features that are to be utilized for the prediction.

Data were collected from the 4 layers within a telecommunication organization (XC), which are the service, resource, billing and customer layers. The data retrieval process to identify customers that could be targeted for direct marketing is shown in Figure 5.1. The resource layer includes customer and device capability data. The service layer includes the data, voice and video usage by the customer over a six-month period. The customer layer includes service provisioning data while the billing layer includes data usage. The data retrieved was saved in a result table which was later passed onto the hybridized machine learning system.

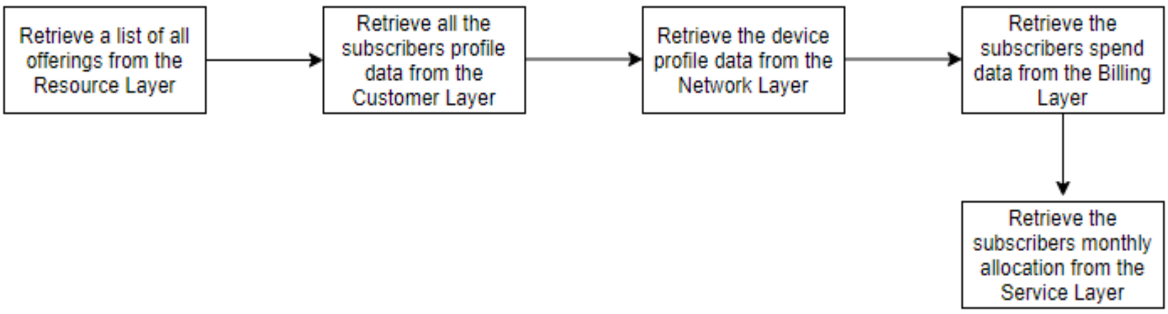

Figure 5.1. Steps within the data gathering process

The customer data was stored in a result table named subscriber_input_data with the description as shown in Table 5.1.

Table 5.1 Description of subscriber input data

S/n	Field name	Description
1	Month_of_transaction	The month when the subscriber initiated the transaction.
2	Subscriber_unique_identifier	A unique identifier to identify the subscriber on the network.
3	Subcriber_sim_card_identifier	The SIM card number used by the subscriber in the device.
4	Duration_subscriber_connected_to_network	The number of days the subscriber has been active on the network since joining.
5	Subscriber_device_identifier	Device identifier number used by the subscriber.
6	Device_manufacturer_name	Device manufacturer name.
7	Device_model_name	Device model name assigned by manufacturer.
8	Device_streaming_capability	Indicates if the device has streaming capability or not.
9	Device_technology_supported	Indicates the highest device technology capability such as 2G, 3G or 4G.
10	Subscriber_total_monthly_spend	The total amount of money the subscriber spends during the month.
11	Subscriber_device_type	The device type being used by the subscriber such as Router, Handset or Tablet.
12	Subscriber_main_offering	The identifier of the subscriber's main/primary allocation.
13	Subscriber_monthly_main_allocation_amount	The amount of the main offering data allocated to the subscriber on the first of the month.
14	Subscriber_monthly_main_allocation_balance	The current balance of the main offering data amount which is allocated on the first of the month.
15	Subscriber_total_data_used_from_main_offering	The amount used of the main offering data amount which was allocated on the first of the month.
16	Subscriber_data_package_offering	The identifier of the subscriber's monthly data-add on allocation (An extra data package which is added every month).
17	Subscriber_monthly_data_package_allocation_amount	The amount of the data add-on allocated to the subscriber on the first of the month.
18	Subscriber_data_package_allocation_type	The type of data add package as this can be an incentive (which is free of charge) or a purchased data package.
19	Subscriber_monthly_data_package_allocation_amount	The amount of data add-on allocated to the subscriber on the first of the month.
20	Subscriber_monthly_data_package_allocation_balance	The current balance of the data add-on amount which is allocated on the first of the month.
21	Subscriber_total_data_used_from_data_package	The amount used of the data add on amount which is allocated on the first of the month.
22	Subscriber_video_purchase_code	The identifier of the video product purchased during the month.
23	Subscriber_video_package_name	The name of the video product purchased during the month.
24	Subscriber_video_package_allocation_amount	The amount of video data allocated to the subscriber on the first of the month.
25	Subscriber_social_purchase_code	The identifier of the social product purchased during the month.
26	Subscriber_social_package_name	The name of the social product purchased during the month.
27	Subscriber_social_package_allocation_amount	The amount of social data allocated to the subscriber on the first of the month.
28	Subscriber_additional_data_package_code	The identifier of a once off data add on product purchased during the month.
29	Subscriber_additional_data_purchase_name	The name of a once off data add on product purchased during the month.
30	Subscriber_additional_data_package_allocation_amount	The amount of a once off data add on product allocated to the subscriber during the month.

The dataset consists of 8 380 908 rows which have 800 000 unique customers with monthly usage spanning over a 6-month period. Figure 5.2 shows fields from the output table queried from the database.

MONTH	MANUFACTURER	MARKETINGNAME	DEVICESLOT_TYPE	WDSNWCHARACT_WIRELESSTECH	TOTAL_MONTHLY_SPEND
MAY	Samsung	Galaxy S9	Mobile Handset	4G	755.31
MAY	Huawei	ELE-L09	Mobile Handset	4G	596.51
MAY	Huawei	P20 Pro	Mobile Handset	4G	651.3
MAY	Huawei	ELE-L09	Mobile Handset	4G	1725.22
MAY	Apple	iPhone XS Max	Mobile Handset	4G	1846.82
MAY	Samsung	Galaxy Note 5	Mobile Handset	4G	199.12
MAY	Huawei	Y5 Prime 2018	Mobile Handset	4G	128.69
MAY	Huawei	Mate 20 Lite	Mobile Handset	4G	700.85
MAY	(null)	(null)	(null)	(null)	325.21
MAY	Nokia	3	Mobile Handset	4G	117.39
MAY	Vodafone	R209-Zr Mobile WiFi	Wireless Router	3G	97.38
MAY	Samsung	Galaxy Note 9	Mobile Handset	4G	1260

Figure 5.2 Fields selected from the output table

5.2 Data Wrangling

The data from the stored table that was created during the data gathering process was the input for the data wrangling process. The purpose of the data wrangling process is to ensure that all duplicates are addressed, null records are removed, and to conduct any data cleaning where necessary. The overall goal is to ensure that the dataset does not contain any duplicate or null records. This procedure was executed to ensure that no duplicate or null record is included in the final dataset that was used.

Table 5.2. Description of the data fields in the dataset

Field Name	Description	Data Type
Month_of_transaction	The month when the subscriber initiated the transaction.	String
Subscriber_unique_identifier	A unique identifier to identify the subscriber on the network.	String
Subcriber_sim_card_identifier	The SIM card number used by the subscriber in the device.	String
Duration_subscriber_connected_to_network	The number of days the subscriber has been active on the network since joining.	String
Subscriber_device_identifier	Device identifier number used by the subscriber.	String
Device_manufacturer_name	Device manufacturer name.	String
Device_model_name	Device model name assigned by manufacturer.	String
Device_streaming_capability	Indicates if the device has streaming capability or not.	String
Device_technology_supported	Indicates the highest device technology capability such as 2G, 3G or 4G.	String
Subscriber_total_monthly_spend	The total amount of money the subscriber spends during the month.	String
Subscriber_device_type	The device type being used by the subscriber such as Router, Handset or Tablet.	String
Subscriber_main_offering	The identifier of the subscriber’s main/primary allocation.	String
Subscriber_monthly_main_allocation_amount	The amount of the main offering data allocated to the subscriber on the first of the month.	String
Subscriber_monthly_main_allocation_balance	The current balance of the main offering data amount which is allocated on the first of the month.	String
Subscriber_total_data_used_from_main_offering	The amount used of the main offering data amount which was allocated on the first of the month.	String
Subscriber_data_package_offering	The identifier of the subscriber’s monthly data add-on allocation (An extra data package which is added every month).	String
Subscriber_monthly_data_package_allocation_amount	The amount of the data add-on allocated to the subscriber on the first of the month.	String
Subscriber_data_package_allocation_type	The type of data add on package as this can be an incentive (which is free of charge) or a purchased data on package.	String

5.3 Data Labelling

With the data successfully gathered and cleaned the data can now be labelled. The purpose of the data labelling process is to formulate relevant data points/ features that would assist in generation predictions through the CBR component or the ML model.

The data features that were derived are as follows:

i) Spend Category – This was categorically defined as low, medium and high based on TOTAL_MONTHLY_SPEND (e.g. Low < R50, Medium ≥R50 AND ≤ R149 or High ≥R150)
ii) Data Use Category – This was categorically defined as Low, Med and High based on CCS_MON_CURRENT_AMOUNT (e.g. Low < 100 MB, Medium ≥100 AND ≤ 299 MB or High ≥300 MB).
iii) Connection Period – this was the time period in years from the day the customer connected to the network. The number of days was divided by 365 days to convert it to years. This was derived from the TIME_CONNECTED field.
iv) Spending daily rate – this was the total spend per month divided by the number of days in the month(s). This was derived from the TOTAL_MONTHLY_SPEND field.
v) Data usage daily rate – this was the total amount of data used per month divided by the number of days. This is derived from the CCS_DATA_USED field.
vi) Available Device Capability – This was grouped into 4G (High), 3G (Medium) or 2G (Low). This was derived from the WDSNWCHARACT_WIRELESSTECH field.
vii) Stream Capability – This was grouped as non-streaming (0) or streaming (1) to specify if a customer's device has the ability to stream or not. This was derived from the SERVICES_SUPPORTEDSERVICES field.
viii) Stream Usage Daily Rate – This was grouped as High, Medium, or Low using the CCS_VIDEO_PURCHASE_MEASURE field.

5.4 Data Selection and Model Training

As the research is an experiment, there is a need to first measure the outcomes, the features which are inputted and processed by the machine learning algorithm. To achieve this the data was separated into a training set and test set. The training set was used by the ML algorithm to identify existing patterns and the test set was used to execute a prediction. It is common practice to use 70 – 75% of the data as a training set and 25 – 30% of the data as a testing set (Moore et al., 2019). The division of the data into these groups is completely random and the data were kept completely separate at all times.

The numpy random function of SciKitLearn was used to divide the data into a training and testing set. The dataset utilized consisted of 2380908 rows with all duplicate and null values removed. The dataset was divided into a training set which consisted of 75% of the data and a test set which consisted of the remaining 25% of the data. The below table shows the number of observations for each dataset. The training data set consisted of 1779842 rows and the test data set consisted of 591713 rows.

Table 5.3. Number of observations

	NUMBER OF OBSERVATIONS
Testing Data	591713
Training Data	1779842

5.5 Implementation of CBR System

In predicting the cases by using case-based reasoning (CBR) and domain expert knowledge, the myCBR RESTFUL API was utilized to execute a prediction using CBR. Using Python, the input data supplied by the user was passed as an HTTP request to the RESTFUL API. The RESTFUL API then passed the input data to the myCBR component. The myCBR component then took the input data and checked if the case matches what is stored in the knowledge base. This the myCBR did by assigning a weighted score to the features and comparing the feature weighting with those of cases that exist in the general knowledge base. Where a similar case was identified, a prediction was then retrieved and passed back to the RESTFUL API which in turn passed it back as an HTTP response as shown in the figure below.

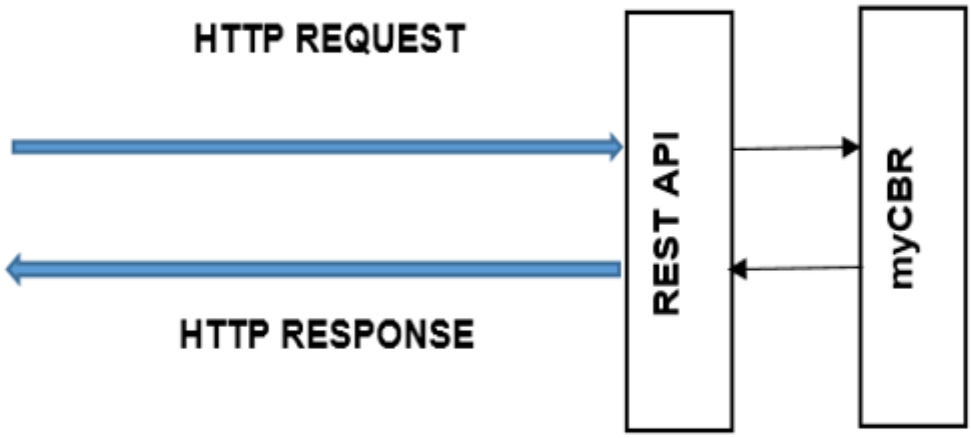

Figure 5.3. Architecture of CBR module

For the purposes of this research, the myCBR knowledge base was configured to handle two use cases. The first use case is to identify customers who were not likely to purchase a product. The second use case was to identify customers who would likely purchase a video streaming product by assessing the customer's data and streaming consumption.

The myCBR knowledge base consists of attributes with each attribute having a weighting. The incoming case is compared to the weighting to determine if the solution to the previous occurrence solves the current presented case.

Identifying customers unlikely to purchase a product, limits the number of wasteful contact with customers and also allows for a viewer to identify an increase in these types of customers. This would allow for the possibility of developing a product for this specific segment if necessary. The assessment was conducted by analysing the customer’s spending, data consumption and duration the customer has been active on the network.

Table 5.4. Weights assigned to cases (Case 1: Unlikely to purchase a product)

CASE	
Case ID:	String
Product:	String
ATTRIBUTE	**WEIGHT**
Package Name:	0.38466
Device:	0.15328
Spending category	0.10521
Data usage category	0.73427
Period of connection	0.63823
Device capability	0.23455
Streaming capability	0.66386

In assessing if the present case presented is a case that justifies a no product offering the similarity was generated using the above table.

In identifying customers likely to purchase a video streaming product a significant revenue increase in video products could be realised and an understanding of customer behaviour surrounding these products could be improved. This could also result in developing and providing more products for the specific segment if necessary.

The assessment was conducted by analysing the customers spending, data consumption and video streaming activity on the network.

Table 5.5. Weights assigned to cases (Case 2: Likely to purchase a video streaming product)

CASE	
Case ID:	String
Product:	String
ATTRIBUTE	**WEIGHT**
Package name	0.69210
Device	0.17328
Spending category	0.26453
Data usage category	0.37453
Device technology supported	0.64889
Device type	0.64387
Streaming capability	0.66929
Monthly main balance	0.63281
Video package name	0.69421
Period of connection	0.49852

In assessing if the present case presented was a case that justified a video streaming product, the similarity was generated using the above table. The rule-based component consists of the following rules, which helped generate explanation for each prediction.

Table 5.6. Predefined domain rules

Rule Number	Spending Category	Data Usage Category	Period of Connection	Device Capability	Streaming Capability	Predicted Product	Explanation
1	Low	High	High	High	Streaming	Video	As the subscriber is a low spender, and receiving a high amount of data a YouTube product is recommended.
2	Medium	High	Medium	High	Non-Streaming	Data	As the subscriber is a medium spender, and consumes majority of the allocation. A WhatsApp product is recommended.
3	Low	Medium	High	High	Non-Streaming	Data	As the subscriber is a low spender, and a long-time user of the network. A data product is recommended.
4	High	High	Low	High	Streaming	Data	As the subscriber is a high spender, and consumes majority of the allocation. A data product is recommended.
5	Medium	High	Low	High	Streaming	Video	As the subscriber is a high spender, and consumes majority of the allocation with streaming capability. A video streaming product is recommended.
6	High	High	Low	High	Streaming	Data	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A YouTube product is recommended.
7	Low	Low	Low	Low	Non-Streaming	No-Offering	As the subscriber is a low spender and a new user of the network. No product will be offered at this point in time.
8	High	High	Low	High	Streaming	Data	As the subscriber is a high spender and shares a lot of data. A data product is recommended.
9	High	High	Medium	High	Streaming	Video	As the subscriber is a high spender with streaming capability. A video streaming product is recommended.
10	Low	High	Medium	High	Streaming	Video	As the subscriber is a low spender and consumes majority of the allocation. A video streaming product is recommended.
11	High	Low	High	High	Streaming	Video	As the subscriber is a low spender and consumes majority of the allocation. A video streaming product is recommended.
12	High	Medium	High	High	Streaming	Social	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A WhatsApp Ticket is recommended.
13	High	Medium	Medium	High	Streaming	Social	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A WhatsApp Ticket is recommended.

5.6 Supervised ML Model

In executing the prediction, the following Supervised ML Models were utilized, Random Forest Classifier, Multi-Layer Perception Artificial Neural Network and Support Vector Machine (SVM) Classifier. The following section discusses the respective models and compares the model's performance during the experiment.

5.6.1 Random Forest Classifier

Figure 5.4 lists the features from the dataset ranked by importance by the Random Forest Classifier. TIME_CONNECTED (the number of days the subscriber has been active on the network) had the highest rate of importance followed by SUBSCRIBER_MAIN_KEY (the subscriber's main offering package), and the TOTAL_MONTHLY_SPEND (the monthly spending rate of the subscriber). A low importance ranking was placed on DATA_USED (the amount of data used for the month) and DATA_CURRENT_BALANCE (the current balance of data available).

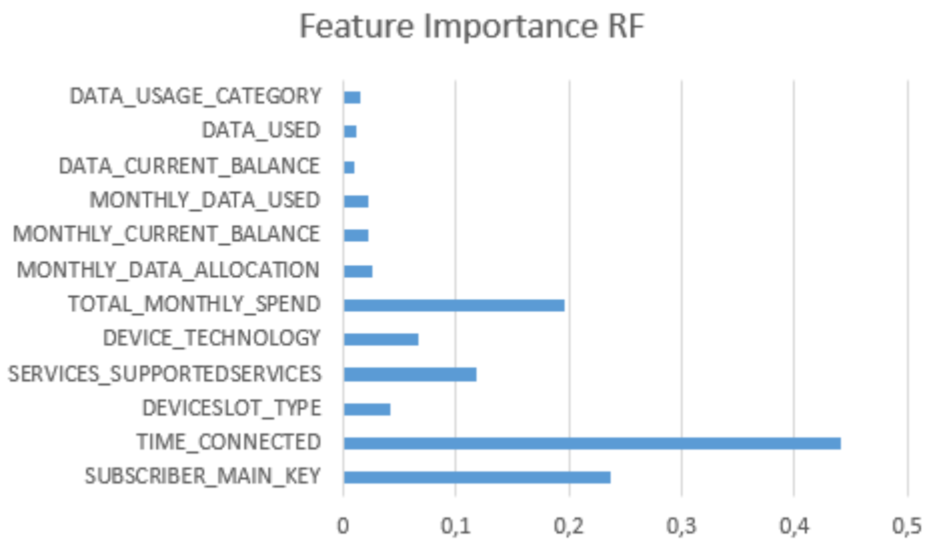

Figure 5.4. Random forest feature importance

5.6.2 Multi-Layer Perceptron Artificial Neural Network

Figure 5.5 lists the features from the dataset ranked by importance by the MLP ANN. As with the Random Forest Classifier TIME_CONNECTED (the number of days the subscriber has been active on the network) had the highest rate of importance. In this instance the MONTHLY_CURRENT_BALANCE (the subscriber's current balance) followed by SUBSCRIBER_MAIN_KEY (the subscriber's main offering package). Unlike the Random

Forest Classifier, the DEVICE_SLOT_TYPE (the type of device being used by the subscriber) and the TOTAL_MONTHLY_SPEND (the monthly spending rate of the subscriber) had an equal importance ranking. As with the Random Forest Classifier a low importance ranking was placed on DATA_USED (the amount of data used for the month) and DATA_CURRENT_BALANCE (the current balance of data available).

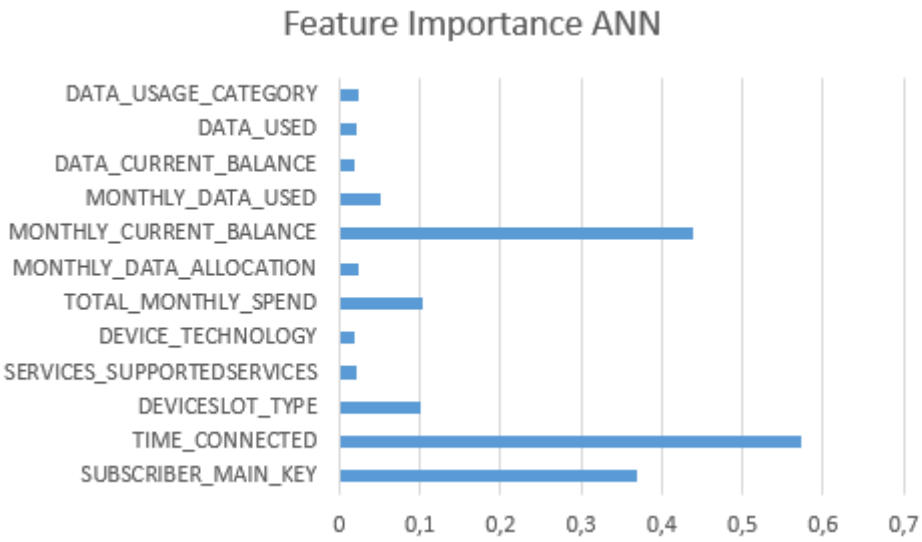

Figure 5.5. ANN feature importance

5.6.3 Support Vector Machines (SVM) Classifier

Figure 5.6 lists the features from the dataset ranked by importance by the SVM Classifier. Unlike both the Random Forest Classifier and the MLP ANN model the SVM Classifier placed a greater importance on the SUBSCRIBER_MAIN_KEY (the subscriber's main offering package) instead of the TIME_CONNECTED (the number of days the subscriber has been active on the network). The SVM Classifier also placed a greater importance on the DEVICE_TECHNOLOGY (which indicates if the subscriber's using 2G, 3G or 4G) than the previous two models. As with the previous two models a low importance ranking were placed on DATA_USED (the amount of data used for the month) and DATA_CURRENT_BALANCE (the current balance of data available). Similar to the previous models DEVICESLOT_TYPE and TIME_CONNECTED are among the most important features.

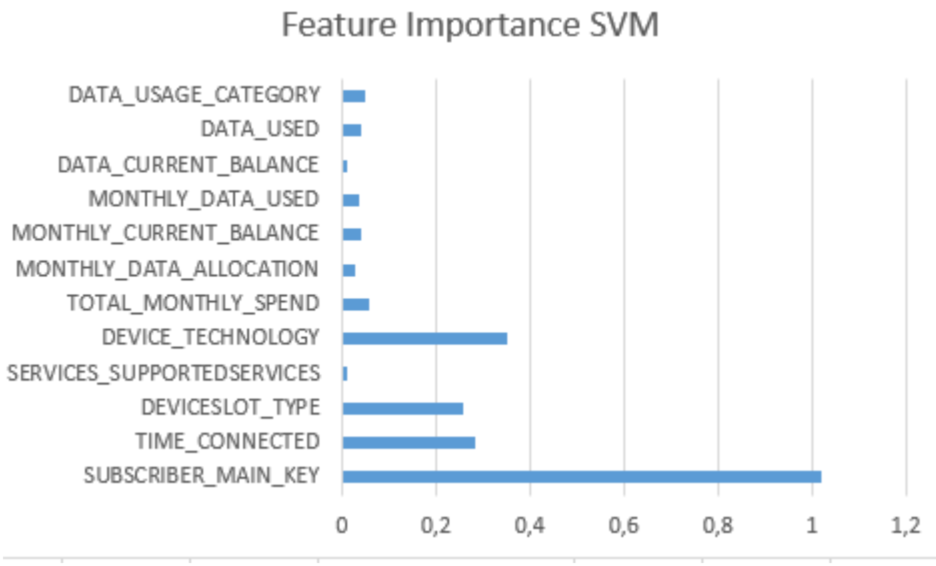

Figure 5.6. SVM feature importance

5.7 Snapshots of System Implementation

The following shows snapshots from the software tool that was developed as a realization of the proposed hybrid ML framework for data-driven direct marketing. The home screen consists of a Social, Video, Data and No Offering buttons. When the button is clicked, the system will display the customers predicted for each respective category.

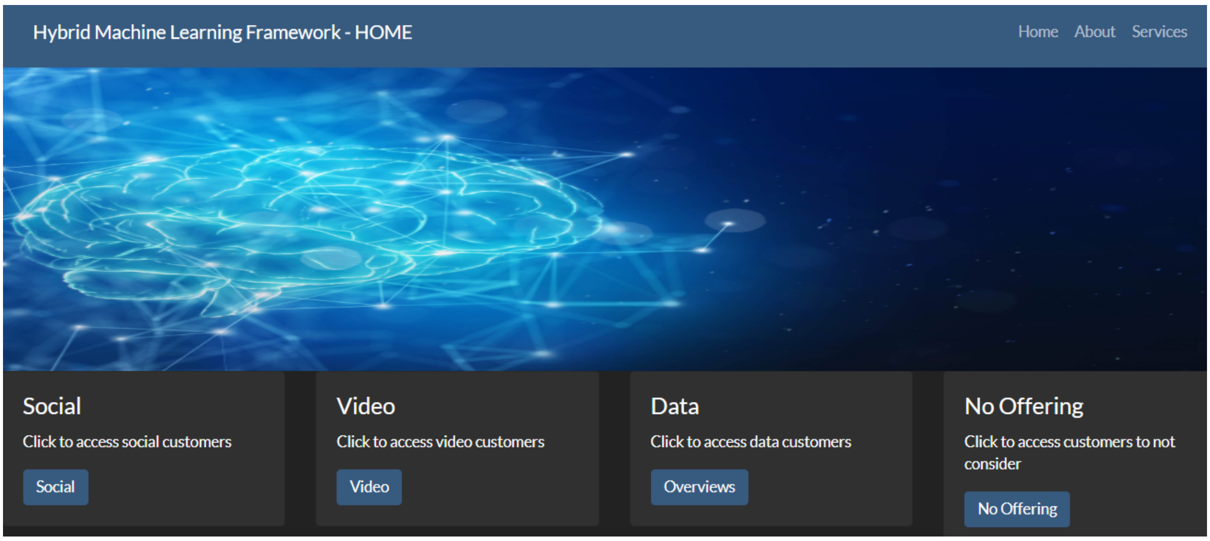

Figure 5.7. Home page

5.7.1 Social Page

The Social page displays all the subscribers to consider for a social category product, in the case below (Figure 5.8) it is for a WhatsApp ticket accompanied with an explanation.

Hybrid Machine Learning Framework - Social Customer Customer About Services

SUBSCRIBER_ID	EXPLANATION	PRODUCT_PREDICTION
CUST110677	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A WhatsApp Ticket is recommended.	Whatsapp Ticket
CUST110793	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A WhatsApp Ticket is recommended.	Whatsapp Ticket
CUST110804	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A WhatsApp Ticket is recommended.	Whatsapp Ticket
CUST112149	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A WhatsApp Ticket is recommended.	Whatsapp Ticket
CUST112150	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A WhatsApp Ticket is recommended.	Whatsapp Ticket
CUST112386	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A WhatsApp Ticket is recommended.	Whatsapp Ticket
CUST112445	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A WhatsApp Ticket is recommended.	Whatsapp Ticket
CUST111662	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A WhatsApp Ticket is recommended.	Whatsapp Ticket
CUST111718	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A WhatsApp Ticket is recommended.	Whatsapp Ticket

Figure 5.8. Social page

5.7.2 Video Page

The Video page displays all the subscribers to consider for a Video category product, in the case below (Figure 5.9) it is for a YouTube ticket accompanied with an explanation.

Hybrid Machine Learning Framework - Video Customer Customer About Services

SUBSCRIBER_ID	EXPLANATION	PRODUCT_PREDICTION
CUST111394	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A YouTube product is recommended.	YouTube Ticket
CUST111395	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A YouTube product is recommended.	YouTube Ticket
CUST111396	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A YouTube product is recommended.	YouTube Ticket
CUST110607	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A YouTube product is recommended.	YouTube Ticket
CUST110608	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A YouTube product is recommended.	YouTube Ticket
CUST110609	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A YouTube product is recommended.	YouTube Ticket
CUST110610	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A YouTube product is recommended.	YouTube Ticket
CUST110611	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A YouTube product is recommended.	YouTube Ticket
CUST110612	As the subscriber is a high spender and medium consumer of allocation with streaming capability. A YouTube product is recommended.	YouTube Ticket

Figure 5.9. Video Page

5.7.3 Data Page

The Data page displays all the subscribers to consider for a data category product, in the case below (Figure 5.10) it is for a 1GB Data Bundle accompanied with an explanation.

Hybrid Machine Learning Framework - Data Customer

Customer About Services

SUBSCRIBER_ID	EXPLANATION	PRODUCT_PREDICTION
CUST110869	The subscriber is a medium spender, high data user, with a medium time duration as a customer, a high spender and with some consumption of the monthly data allocation. Therefore a 1GB Data Bundle is recommended.	1GB Data Bundle
CUST110607	The subscriber is a medium spender, high data user, with a medium time duration as a customer, a high spender and with some consumption of the monthly data allocation. Therefore a 1GB Data Bundle is recommended.	1GB Data Bundle
CUST110608	The subscriber is a medium spender, high data user, with a medium time duration as a customer, a high spender and with some consumption of the monthly data allocation. Therefore a 1GB Data Bundle is recommended.	1GB Data Bundle
CUST110869	The subscriber is a medium spender, high data user, with a medium time duration as a customer, a high spender and with some consumption of the monthly data allocation. Therefore a 1GB Data Bundle is recommended.	1GB Data Bundle
CUST110607	The subscriber is a medium spender, high data user, with a medium time duration as a customer, a high spender and with some consumption of the monthly data allocation. Therefore a 1GB Data Bundle is recommended.	1GB Data Bundle
CUST110607	The subscriber is a medium spender, high data user, with a medium time duration as a customer, a high spender and with some consumption of the monthly data allocation. Therefore a 1GB Data Bundle is recommended.	1GB Data Bundle
CUST105630	The subscriber is a medium spender, high data user, with a medium time duration as a customer, a high spender and with some consumption of the monthly data allocation. Therefore a 1GB Data Bundle is recommended.	1GB Data Bundle
CUST105630	The subscriber is a medium spender, high data user, with a medium time duration as a customer, a high spender and with some consumption of the monthly data allocation. Therefore a 1GB Data Bundle is recommended.	1GB Data Bundle
CUST105677	The subscriber is a medium spender, high data user, with a medium time duration as a customer, a high spender and with some consumption of the monthly data allocation. Therefore a 1GB Data Bundle is recommended.	1GB Data Bundle

Figure 5.10. Data page

5.7.4 No Product Customer Page

The No offering page displays all the subscribers to not consider for a product, in the below case it displays the subscribers accompanied with an explanation.

Hybrid Machine Learning Framework - No Product Customer

Customer About Services

SUBSCRIBER_ID	EXPLANATION	PRODUCT_PREDICTION
CUST105917	As the subscriber is a low spender and a new user of the network. No product will be offered at this point in time.	NO AVAILABLE OFFERING
CUST105918	As the subscriber is a low spender and a new user of the network. No product will be offered at this point in time.	NO AVAILABLE OFFERING
CUST105919	As the subscriber is a low spender and a new user of the network. No product will be offered at this point in time.	NO AVAILABLE OFFERING
CUST105920	As the subscriber is a low spender and a new user of the network. No product will be offered at this point in time.	NO AVAILABLE OFFERING
CUST105921	As the subscriber is a low spender and a new user of the network. No product will be offered at this point in time.	NO AVAILABLE OFFERING

Figure 5.11. No product customer page

5.8 Evaluation of the Framework

After the implementation of the hybridized framework, an evaluation was conducted which focused on two perspectives. Firstly, the framework performance was assessed by evaluating the framework's accuracy of generated predictions and the quality of explanations that accompanied the predictions. The hybridized framework allows one of SVM, ANN and RF to be chosen per time as the ML algorithmic model to obtain a prediction. Secondly, the ease of use of the hybridized framework was evaluated. The two-phased evaluation approach was essential to determine the effectiveness of the hybridized framework within an organization.

5.8.1 Evaluation of Performance

The Random Forest Classifier model used the TOTAL_MONTHLY_SPEND, TIME_CONNECTED and SUBSCRIBER_MAIN_KEY as the most important features to execute a prediction. As per the below confusion matrix the model performance was quite good. The model predicted 92.93% of instances correctly with a 98.19% sensitivity and a 31.94% specificity as shown in Table 5.7.

Table 5.7. Random forest classifier confusion matrix

Confusion Matrix		**Positive**	**Negative**		
	Positive	534898	31956	94.36257	
Random Forest	**Negative**	9856	15003	60.35239	
				92.93374	**Accuracy**
		98.19074298	31.94914713		
		Sensitivity	**Specificity**		

The ANN model on the other hand used the MONTHLY_CURRENT_BALANCE, TIME_CONNECTED AND SUBSCRIBER_MAIN_KEY as the most important features to execute a prediction. As per the below confusion matrix the model was also quite good. The model predicted 97.09% of instance correctly with a 97.32% sensitivity and a 28.4% specificity as shown in Table 5.8.

Table 5.8. ANN Confusion Matrix

Confusion Matrix		Positive	Negative		
	Positive	573962	1432	99.75113	
ANN	**Negative**	15751	568	3.480605	
				97.09606	**Accuracy**
		97.32903972	28.4		
		Sensitivity	**Specificity**		

The SVM model used DEVICE_TECHNOLOGY, DEVICESLOT_TYPE, TIME_CONNECTED and SUBSCRIBER_MAIN_KEY as the most important features to execute a prediction. As per the confusion matrix of the SVM (see Table 5.9) the model performance was good. The model predicted 92% of instances correctly with a 98.81% sensitivity and a 2.3% specificity as shown in Table 5.9.

Table 5.9. SVM Confusion Matrix

Confusion Matrix		Positive	Negative		
	Positive	543376	40810	93.01421	
SVM	**Negative**	6524	1003	13.32536	
				92.00051	**Accuracy**
		98.81360247	2.3987755		
		Sensitivity	**Specificity**		

Table 5.10 displays a comparative view of the Random Forest, Support Vector Machine and Artificial Neural Network machine learning models. The table displays the performance data which was compared on the basis of accuracy, sensitivity and specificity.

Accuracy was calculated by taking the number of correct predictions divided by the number of total predictions. It was noted that the Artificial Neural Network model yielded the highest accuracy at 97.09%, followed by the Random Forest model at 92.93% and the Support Vector Machine model at 92.00%.

Sensitivity, also known as the true sensitivity rate was calculated by taking the number of correct positive predictions divided by the number of total positive predictions. It was noted that the Support Vector Machine model yielded the highest sensitivity at 98.81%, followed by the Random Forest model with 98.19% and the Artificial Neural Network model at 97.32%.

Specificity, also known as the true negative rate was calculated by taking the number of correct negative predictions divided by the number of total negative predictions. It was noted that the Random Forest model yielded the highest specificity at 31.94%, followed by the Artificial Neural Network model at 28.4% and the Support Vector Machine model at 2.3%.

Table 5.10 Comparative View of the Algorithmic Model Performance

ML model	Accuracy	Sensitivity	Specificity
Random Forest	92.93374	98.19074298	31.94914713
ANN	97.09606	97.32903972	28.4
SVM	92.00051	98.81360247	2.3987755

5.8.2 Evaluation of Explanation

ML models have been found to be effective in predicting a specific outcome but lacks an explanation, resulting in a lack of trust in the ML model (Zafar and Khan, 2019). Where an explanation can be accompanied by a prediction it is imperative that the explanation needs to be of quality and effective. Gunning (2017) proposed an explanation framework that consists of the ML model which produces a prediction accompanied by an explanation that will contribute to the user's decision making (see Figure 5.12). The explanation framework of Gunning (2017) was used to evaluate the quality of explanation generated by the hybrid ML framework for direct marketing that is proposed in this study.

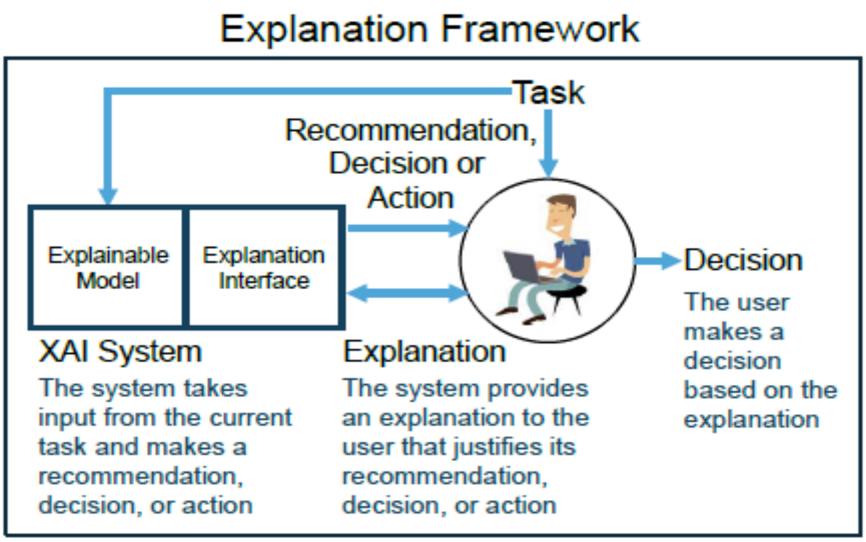

Figure 5.12. Explanation Framework (Gunning, 2017)

5.9 Measurement of the Quality of Explanation

The field of Explainable Artificial Intelligence (XAI) cited five layers in measuring the effectiveness of an explanation. These five layers are User Satisfaction, Mental Model, Task Performance, Trust Assessment and Correctability as shown in Figure 5.13 (Gunning, 2017).

User Satisfaction
• Clarification of the explanation • Use of the explanation
Mental Model
• Understanding individual decisions • Understanding the overall model • Strength/weakness of assessment
Task Performance
• Does the explanation improve the user's decision task performance? • Artificial decision tasks introduced to diagnose the user's understanding
Trust Assessment
• Appropriate future use and trust
Correctability
• Identifying errors • Correcting errors • Continuous training

Figure 5.13 Layers to measure an explanation

In evaluating the quality and effectiveness of the explanation that accompanied the prediction produced by the hybridized system, the Goal Question Metric (GQM) approach was used. The GQM used for evaluation was based on the explanation framework. A goal for each layer is described in Table 5.11.

Table 5.11 Goals of the evaluation

User satisfaction (Clarity of explanations)	**GOAL 1**: Determine if explanation was clear and understandable
Mental model (Understanding decisions)	**GOAL 2**: Determine how decision making is understood by the user
Correctness (Correcting errors and training)	**GOAL 3**: Verify the correctness of predictions and explanations
Trust assessment (Future use and trust)	**GOAL 4**: Determine the hybrid ML framework's contribution to trust
Task performance (Explanation improving decision making)	**GOAL 5**: Establish contribution to decision making

5.10 Response Results

Each goal contained a set of questions which the user answered using a five-point rating scale which is (not at all) 1 - 2 - 3 - 4 - 5 (very much). A total number of eight participants consisting of call centre agents who interact directly with customers and managerial team members participated in the survey.

The survey participants included 5 customer care agents, and this was viewed as very important input as these agents interact directly with customers. Customer care agents interact with customers telephonically and has first-hand insights as to what purchases customers are likely to make.

The survey also included 1 supervisor and 2 managers. The role of these participants was to enable the gathering of perception of the overall improvement in decision making and trustworthiness of the proposed hybrid ML system. Each participant was presented with the same 5 goals and the questions linked to each goal. The median of the scores of the questions were calculated to compute a median goal rating for each goal by an individual. Thereafter, the median of the ratings by all the participants' goal rating was computed to obtain an overall goal rating. This is shown in Table 5.12.

Table 5.12 Response to questions

	GOAL1	Q1	Q2	Q3	Q4	Q5	GOAL2	Q1	Q2	Q3
Participant 1	4	4	3	4	5	4	3	3	3	5
Participant 2	4	5	5	4	3	3	3	2	3	4
Participant 3	5	5	5	4	5	5	4	4	4	5
Participant 4	4	4	3	4	5	5	3	3	3	5
Participant 5	4	5	4	4	4	4	3	3	3	4
Participant 6	5	5	5	4	5	4	4	4	4	5
Participant 7	4	4	4	4	5	4	4	5	5	3
Participant 8	4	5	4	4	4	4	4	4	4	4

	GOAL3	Q1	Q2	Q3	GOAL4	Q1	Q2	GOAL5	Q1	Q2	Q3
Participant 1	5	4	5	5	5	5	5	4	5	4	4
Participant 2	4	4	2	4	3.5	3	4	3	3	3	4
Participant 3	5	4	5	5	5	5	5	4	4	4	4
Participant 4	3	5	2	3	4.5	5	4	4	4	4	5
Participant 5	2	3	2	2	3.5	3	4	2	2	2	2
Participant 6	4	4	4	5	5	5	5	5	5	5	5
Participant 7	5	5	5	4	4	4	4	4	4	4	4
Participant 8	4	5	4	4	4	4	4	4	4	3	5

5.11 Results and Discussion

The median for each goal and a percentage for the median was computed based on all the responses captured from the survey as shown below in Table 5.13.

Table 5.13 Median scores of individual goals

GOAL NAME	MEDIAN	PERCENTAGE
User satisfaction (Clarity of explanations)		
GOAL 1: Determine if explanation was clear and understandable.	**4**	**80%**
Mental model (Understanding decisions)		
GOAL 2: Determine how decision making was understood by the user.	**3.5**	**70%**
Correctness (Correcting errors and training)		
GOAL 3: Verify the correctness of predictions and explanations.	**4**	**80%**
Trust assessment (Future use and trust)		
GOAL 4: Determine the framework's contribution to trust in ML solutions.	**4.25**	**85%**
Task performance (Explanation improving decision making)		
GOAL 5: Establish contribution to decision making.	**4**	**80%**

5.12 Discussion

The hybridized data-driven direct product marketing framework evaluation rating was derived from eight participants. The selected participants installed an instance of the system and they evaluated all its functionalities. Based on the feedback received from the participants the overall implementation and use of the framework were successful.

The explanation produced by the framework is an important aspect of the framework as this reduces the need for highly skilled personnel to interpret results. Out of the eight participants the median scoring was 4 out 5 (80%) when they assessed if the explanation was clear and understandable. A median scoring of 3.5 out of 5 (70%) was obtained for how well decision-making of the system was understood. These two scores suggest that a user can interpret predictions quite comfortably and can gather an understanding of how the framework derives at a decision.

The framework received a median score of 4 out of 5 (80%) for correctness and performance which suggests that the framework can be utilized operationally as customers are predicted accurately and performance of functionality is at an optimal level. Participants gave a median score of 4.25 out of 5 (85%) for trustworthiness. This suggests that users and management are likely to have a positive uptake of the framework, but also contribute to the overall implementation of machine learning solutions with an organization. Participants also gave a median score of 4 out of 5 (80%) for task performance. This suggest that users and management believe that the explanation produced contributes to the decision making in a positive manner.

As the framework was implemented with actual telecommunication data with evidence showing it to be successful, the framework can contribute to any organization within the telecommunication industry as a tool used to upsell products, but also contribute significantly to the understanding of customer behaviour by the generated explanations accompanied with the prediction output.

The design of the framework enables it to be adaptable as any type of supervised machine learning (ML) algorithmic model can be added to the framework. The addition of a ML algorithmic model will be by importing the required algorithm library package into the hybrid framework using the python import functionality. In the event that the user wants to use a different set of features to predict different products, the desired features can be selected as loaded into the training data set, which can then be used to train selected ML models.

5.13 Summary

In this chapter, the description of the practical implementation of the hybrid framework was presented. The data gathering, dataset, data wrangling and data labelling processes were discussed. In addition to this, the data processing by the various ML algorithms, case-based and rule-based reasoning were also explained. A detailed description of the computational processes was outlined which also includes the generation of prediction and explanation.

The chapter also presented the evaluation of the performance of the hybrid framework and the evaluation of the explanation. The chapter was concluded with a discussion of the implications of the evaluation results and the adaptability of the framework.

CHAPTER 6
SUMMARY, CONCLUSION AND RECOMMENDATIONS

6.1 Summary

In this study, a hybridized framework for data-driven direct product marketing was presented which consists of machine learning algorithmic models, case-based and rule-based reasoning. The framework used an actual dataset from a telecommunication company as input and executed predictions with explanations for voice, data and video streaming products. The framework presented can be used with additional machine learning models to upsell any other specified products within the telecommunication industry.

6.2 Conclusion

Based on the objectives and research questions defined for the study the following was achieved:

1. In determining the attributes required for an easy to use and explainable ML process framework for data-driven direct product marketing. The study had to establish the requirements to satisfy such a need. In obtaining a clear understanding of what was needed in such a framework, an extensive review of the literature and existing tools were conducted. It was evident that many existing direct product marketing tools exist, but many lacked the ability to process large amounts of data, execute a product prediction by use of machine learning algorithmic models and produce an explanation with the predictive output to give an understanding as to why the prediction was generated. Once a clear understanding was obtained a clear set of requirements was defined by hosting JAD sessions with domain experts and users.

2. In addressing the requirements of a hybridized framework consisting of machine learning algorithmic models and intelligent reasoning was expected to be designed and developed. The design was delivered by the use of architectural modelling and system design. The development was conducted using only open source development technologies. It consisted of an easy-to-use graphical interface that allowed for a user to log in securely, load a dataset, select an ML algorithmic model and produce prediction accompanied with an explanation as output. The machine learning algorithmic models consisted of the Random Forest, Support Vector Machine and Artificial Neural Network models. Intelligent reasoning was achieved by using a combination of case-based and rule-based reasoning. This was to leverage the statistical, predictive capabilities provided by machine learning algorithmic models and

the explanation capabilities provided by CBR and rule-based reasoning. A fully functional artefact was then presented and demonstrated to domain experts and users.

3. To ensure that the requirements were successfully addressed, and the artefact was easy to use. Accompanied by a comprehensive explanation an evaluation was conducted in two parts. The first part of the evaluation focused on the performance of the machine learning algorithmic models. The metrics used were accuracy, sensitivity and specificity of the models. The second part of the evaluation focused on user satisfaction, mental model, task performance, trust assessment and correctability. The Goal Question Metric approach was used to achieve this. The evaluation of the framework has shown that the machine learning models within the framework produced predictions with high levels of accuracy and in a timely manner. Users have indicated that the produced explanations have contributed positively to the users' decision making and trust in machine learning solutions. The produced explanations were clear and understandable to the user, although an understanding of how the predictions were generated was only up to a certain acceptable level, this can be improved.

6.3 Contributions

The study's contribution can be derived from a theoretical, methodological and practical perspective. Theoretically, the study utilizes an existing telecommunication data retrieval process that fetches data from five different layers (resource layer, service layer, customer layer and the billing layer), within a telecommunication organization. The data retrieved from these layers was then passed into a hybridized framework that leverages the strength of intelligent reasoning processes for explanations and the ability to process complex datasets from artificial intelligence models. This provides a new perspective on the creation of a process framework that can facilitate data-driven direct marketing using machine learning in the telecom industry.

Methodologically, the framework provides a method for multiple algorithmic models to be applied to a complex dataset for the generation of product offering predictions. The framework allows for reasonably skilled resources to operate and execute predictions. As the predictions are accompanied by an explanation it allows for a better understanding of customer behaviour and a better understanding as to why a prediction or recommendation was generated. The hybrid architecture that provides alternative means of obtaining accurate predictions was accompanied by explanations by using a combination of ML, CBR, and rule-based reasoning which is an innovative approach to direct-product marketing in the telecom industry.

Practically, the framework can be applied by any telecommunication organization with customer data. All that is required is to install the framework and immediately start executing predictions for the product offering. As the framework is based on open source technology and does not require highly skilled personnel, setup costs are very low. These product offerings in turn would contribute to sales and add business value.

Generally, the study provides a new perspective on the application of explainable machine learning for data-driven direct marketing within the telecommunication industry through the utilization of a hybridized ML architecture. It presents a way to attain improvement on the existing direct marketing tools that mostly lack explanation capability to justify the predictions that they generate.

6.4 Recommendations and Future work

The framework is designed for any machine learning model to be added with any customer dataset which is derived from the resource, customer, service, and billing layers at any telecommunication company. However, it is important to pay special attention to the data quality of the data set to ensure the training of the models is effective.

An improvement in the area of how a new algorithmic model can be implemented in a more automated manner and by the use of a graphical interfacing wizard. The initial setup of a new algorithmic model takes a bit of time and has the potential for improvement. A larger and more granular data set can be applied in the training process to improve the prediction quality. The generation of the explanations needs to be tested with data from an out of the ordinary usage period such as a holiday period or an emergency lockdown period to evaluate the type of explanation accompanied by the recommendation.

Future work should include a wider variety of cases within the knowledge base of the case-based reasoning module and to verify the performance when a larger variety of products are introduced. In addition to this, more security features need to be introduced if the framework is introduced to a wider user group. Also, more experiments with explainable deep learning models and various types of datasets from the telecom industry needs to be conducted in order to assess the robustness and reliability of the process framework.

REFERENCES

Aamodt, A. and Plaza, E., 1994. Case-based reasoning: Foundational issues, methodological variations, and system approaches. AI communications, 7(1), pp.39-59.

Aamodt, A., 1995, April. Case-based reasoning-an introduction. In a new force in advanced systems development, Unicom Seminars (pp. 9-23).

Abu Dalffa, M., Abu-Nasser, B.S. and Abu-Naser, S.S., 2019. Tic-Tac-Toe Learning Using Artificial Neural Networks.

Abu-Nasser, B.S. and Abu Naser, S.S., 2018. Rule-Based System for Watermelon Diseases and Treatment. *International Journal of Academic Information Systems Research* (IJAISR), 2(7), pp.1-7.

Achtaich, A., Roudies, O., Souissi, N., Salinesi, C. and Mazo, R., 2019, September. Evaluation of the state-constraint transition modelling language: a goal question metric approach. In Proceedings of the 23rd International Systems and Software Product Line Conference-Volume B (pp. 106-113).

Adams, D., Oh, D.H., Kim, D.W., Lee, C.H. and Oh, M., 2020. Prediction of SOx–NOx emission from a coal-fired CFB power plant with machine learning: Plant data learned by deep neural network and least square support vector machine. *Journal of Cleaner Production*, 270, p.122310.

Ahmad, A.S., Hassan, M.Y., Abdullah, M.P., Rahman, H.A., Hussin, F., Abdullah, H. and Saidur, R., 2014. A review on applications of ANN and SVM for building electrical energy consumption forecasting. Renewable and Sustainable Energy Reviews, 33, pp.102-109.

Alanen, A.: Efficient direct marketing: Case: Valtapinnoite Oy (2016).

Alpaydin, E., 2009. Introduction to machine learning. MIT press.
Cruz, J.A. and Wishart, D.S., 2006. Applications of machine learning in cancer prediction and prognosis. Cancer informatics, 2, p.117693510600200030.

Al-Rifaie, M.M. and Alhakbani, H.A., 2016, July. Handling class imbalance in direct marketing dataset using a hybrid data and algorithmic level solutions. *In 2016 SAI Computing Conference* (SAI) (pp. 446-451). IEEE.

Al-Sharif, L., Al Sukkar, G., Hakouz, A. and Al-Shamayleh, N.A., 2017. Rule-based calculation and simulation design of elevator traffic systems for high-rise office buildings. Building Services Engineering Research and Technology, 38(5), pp.536-562.

Amoako, G.K., 2019. Relationship marketing, orientation, brand equity and firm value: The mediating role of customer value—an emerging market perspective. *Journal of Relationship Marketing*, 18(4), pp.280-308.

Angilica, D., Ianni, G. and Pacenza, F., 2019, November. Tight integration of rule-based tools in game development. *In International Conference of the Italian Association for Artificial Intelligence* (pp. 3-17). Springer, Cham.

Aoyagi, K., Wang, H., Sudo, H. and Chiba, A., 2019. Simple method to construct process maps for additive manufacturing using a support vector machine. Additive Manufacturing, 27, pp.353-362.

Appiahene, P., Missah, Y.M. and Najim, U., 2020. Predicting Bank Operational Efficiency Using Machine Learning Algorithm: Comparative Study of Decision Tree, Random Forest, and Neural Networks. Advances in Fuzzy Systems, 2020.

Arditi, D. and Tokdemir, O.B., 1999. Comparison of case-based reasoning and artificial neural networks. *Journal of Computing in Civil Engineering*, 13(3), pp.162-169.

Asteris, P.G. and Mokos, V.G., 2019. Concrete compressive strength using artificial neural networks. Neural Computing and Applications, pp.1-20.

Aung, M.M., Han, T.T. and Ko, S.M., 2019. Customer churn prediction using association rule mining. *International Journal of Trend in Scientific Research and Development*, *3*(5), pp.1886-1890.

Auria, L. and Moro, R.A., 2008. Support vector machines (SVM) as a technique for solvency analysis.

Avery, J. and Israeli, A., 2018. Hubble Contact Lenses: Data Driven Direct-to-Consumer Marketing.

Ayyaz, S. and Hussain, N., 2019. Direct Marketing of Fresh Mangoes: A Case Study of Mango Smallholders in Pakistan.

Babu, M.M., Dey, B.L., Rahman, M., Roy, S.K., Alwi, S.F.S. and Kamal, M.M., 2020. Value co-creation through social innovation: A study of sustainable strategic alliance in telecommunication and financial services sectors in Bangladesh. Industrial Marketing Management, 89, pp.13-27.

Bach, K., Mathisen, B. M., & Jaiswal, A. Demonstrating the myCBR Rest API. Available at:https://iccbr2019.com/wpcontent/uploads/2019/09/01_paper_Demonstrating_the_myCBR_REST_API.pdf

Bacon-Shone, J.H., 2013. *Introduction to quantitative research methods*. Graduate School, The University of Hong Kong.

Barnett, M.L., Bernal, N.A. and Sanchez, B.E.L., 2020. Direct-to-Consumer Marketing for Parent-Child Interaction Therapy: Impact of Language and Messenger. *Journal of Child and Family Studies,* 29(1), pp.71-81.

Baroni, M., 2020. Linguistic generalization and compositionality in modern artificial neural networks. Philosophical Transactions of the Royal Society B, 375(1791), p.20190307.

Bashar, A., 2019. Survey on evolving deep learning neural network architectures. Journal of Artificial Intelligence, 1(02), pp.73-82.

Basili, V., Briand, L., Bianculli, D., Nejati, S., Pastore, F. and Sabetzadeh, M., 2018. Software engineering research and industry: a symbiotic relationship to Foster impact. IEEE Software, 35(5), pp.44-49.

Basili, V.R., 1994. Goal question metric paradigm. Encyclopaedia of software engineering, pp.528-532.

Batzias, F. and Kopsidas, O., 2019, December. Knowledge acquisition for reconstruction and identification of dimensionless groups by means of case-based reasoning. *In AIP Conference Proceedings* (Vol. 2186, No. 1, p. 170033). AIP Publishing LLC.

Bayoude, K., Ouassit, Y., Ardchir, S. and Azouazi, M.: How machine learning potentials are transforming the practice of digital marketing: State of the art. Periodicals of Engineering and Natural Sciences, 6(2), 373-379, (2018).

Beam, A.L. and Kohane, I.S., 2018. Big data and machine learning in health care. Jama, 319(13), pp.1317-1318.

Beheshtian-Ardakani, A., Fathian, M. and Gholamian, M.: A novel model for product bundling and direct marketing in e-commerce based on market segmentation. Decision Science Letters, 7(1), 39-54 (2018).

Belgiu, M. and Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS *Journal of Photogrammetry and Remote Sensing*, 114, pp.24-31.

Belkin, M., Hsu, D., Ma, S. and Mandal, S., 2019. Reconciling modern machine-learning practice and the classical bias–variance trade-off. Proceedings of the National Academy of Sciences, 116(32), pp.15849-15854.

Bergmann, R., Althoff, K.D., Minor, M., Reichle, M. and Bach, K., 2009. Case-Based Reasoning. *KI*, *23*(1), pp.5-11.

Berka, P., 2020. Sentiment analysis using rule-based and case-based reasoning. *Journal of Intelligent Information Systems*, pp.1-16.

Berry, M.W., Mohamed, A. and Yap, B.W. eds., 2019. Supervised and Unsupervised Learning for Data Science. Springer Nature.

Bhatt, U., Xiang, A., Sharma, S., Weller, A., Taly, A., Jia, Y., Ghosh, J., Puri, R., Moura, J.M. and Eckersley, P., 2020, January. Explainable machine learning in deployment. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 648-657).

Bi, X.A., Wang, Y., Shu, Q., Sun, Q. and Xu, Q., 2018. Classification of autism spectrum disorder using random support vector machine cluster. Frontiers in genetics, 9, p.18.

Białek, Ł., Dunin-Kęplicz, B. and Szałas, A., 2017, June. Rule-based reasoning with belief structures. *In International Symposium on Methodologies for Intelligent Systems* (pp. 229-239). Springer, Cham.
Biswas, S.K., Sinha, N., Purakayastha, B. and Marbaniang, L., 2014. Hybrid expert system using case-based reasoning and neural network for classification. Biologically Inspired Cognitive Architectures, 9, pp.57-70.

Boehm, B., Rombach, H.D. and Zelkowitz, M.V. eds., 2005. Foundations of empirical software engineering: the legacy of Victor R. Basili. Springer Science & Business Media.

Bonacina, M.P., 2017, August. Automated Reasoning for Explainable Artificial Intelligence. In *The First International ARCADE (Automated Reasoning: Challenges, Applications, Directions, Exemplary Achievements) Workshop (in association with CADE-26), Gothenburg, Sweden*.

Brezak, S., Vlastelica, T. and Kostić, S.C., 2019. Are marketing communications relevant for the preparation of buyer-seller negotiations? teme, pp.511-526.

Brieuc, M.S., Waters, C.D., Drinan, D.P. and Naish, K.A., 2018. A practical introduction to Random Forest for genetic association studies in ecology and evolution. Molecular ecology resources, 18(4), pp.755-766.

Brożek, B., 2017. Two faces of legal reasoning: rule-based and case-based. *Kraków [bdw], http://www. academia. edu/28674043/Two_ Faces_of_Legal_Reasoning_Rule-Based_and_Case-Based*.

Bui, D.K., Nguyen, T.N., Ngo, T.D. and Nguyen-Xuan, H., 2020. An artificial neural network (ANN) expert system enhanced with the electromagnetism-based firefly algorithm (EFA) for predicting the energy consumption in buildings. Energy, 190, p.116370.

Burbidge, R. and Buxton, B., 2001. An introduction to support vector machines for data mining. Keynote papers, young OR12, pp.3-15.

Buttle, F. and Maklan, S.: Customer relationship management: concepts and technologies. Routledge (2019).

Cai, Z. and Liu, J., 2018. Approximating quantum many-body wave functions using artificial neural networks. Physical Review B, 97(3), p.035116.

Camilleri, M.A., 2020. The use of data-driven technologies for customer-centric marketing. *International Journal of Big Data Management*, 1(1), pp.50-63.

Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., Vogt-Maranto, L. and Zdeborová, L., 2019. Machine learning and the physical sciences. Reviews of Modern Physics, 91(4), p.045002.

Carral, D., Dragoste, I., González, L., Jacobs, C., Krötzsch, M. and Urbani, J., 2019, October. VLog: A rule engine for knowledge graphs. In International Semantic Web Conference (pp. 19-35). Springer, Cham.

Caruana, R. and Niculescu-Mizil, A., 2006, June. An empirical comparison of supervised learning algorithms. In *Proceedings of the 23rd international conference on Machine learning* (pp. 161-168). ACM.

Castanedo, F., Valverde, G., Zaratiegui, J. and Vazquez, A.: Using deep learning to predict customer churn in a mobile telecommunication network (2014).

Cervantes, J., Garcia-Lamont, F., Rodriguez-Mazahua, L. and Lopez, A., 2020. A comprehensive survey on support vector machine classification: Applications, challenges and trends. Neurocomputing, 408, pp.189-215.

Che, J., Zhao, S., Li, Y. and Li, K., 2020. Bank Telemarketing Forecasting Model Based on t-SNE-SVM. *Journal of Service Science and Management*, 13(3), pp.435-448.

Chen, C.: Use cases and challenges in telecom big data analytics. APSIPA Transactions on Signal and Information Processing, 5 (2016).

Chen, C.M., 2016. Use cases and challenges in telecom big data analytics. APSIPA Transactions on Signal and Information Processing, 5.

Chen, W., Sun, Z. and Han, J., 2019. Landslide susceptibility modelling using integrated ensemble weights of evidence with logistic regression and random forest models. Applied sciences, 9(1), p.171.

Cheng, L., Chen, X., De Vos, J., Lai, X. and Witlox, F., 2019. Applying a random forest method approach to model travel mode choice behavior. Travel behaviour and society, 14, pp.1-10.

Cheng, Y., Feris, R.S., Pickover, C.A., Vukovic, M. and Wang, J., International Business Machines Corp, 2018. *Finding missing persons by learning features for person attribute classification based on deep learning*. U.S. Patent Application 10/163,042.

Chicco, D. and Jurman, G., 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC genomics, 21(1), p.6.

Cichy, R.M. and Kaiser, D., 2019. Deep neural networks as scientific models. Trends in cognitive sciences, 23(4), pp.305-317.

Coita, D.C., Popa, A.L., Tarcza, T.M., Ţarcă, N.N., Sasu, D.V. and Roşca, R., 2019. The End of Marketing as You Know It. Or Not!. In Strategic Innovative Marketing and Tourism (pp. 405-412). Springer, Cham.

Couronné, R., Probst, P. and Boulesteix, A.: Random forest versus logistic regression: a large-scale benchmark experiment. BMC bioinformatics, 19(1), 270, (2018).

Covington, P., Adams, J. and Sargin, E., 2016, September. Deep neural networks for YouTube recommendations. In Proceedings of the 10th ACM conference on recommender systems (pp. 191-198).

Crossan, F., 2003. Research philosophy: towards an understanding. *Nurse researcher*, *11*(1), pp.46-56.

Cucchiara, R., Piccardi, M. and Mello, P., 2000. Image analysis and rule-based reasoning for a traffic monitoring system. IEEE Transactions on Intelligent Transportation Systems, 1(2), pp.119-130.

Dabowsa, N.I.A., Amaitik, N.M., Maatuk, A.M. and Aljawarneh, S.A., 2017, August. A hybrid intelligent system for skin disease diagnosis. In 2017 International Conference on Engineering and Technology (ICET) (pp. 1-6). IEEE.

Dakouan, C., Benabdelouahed, R. and Anabir, H., 2019. Inbound marketing vs. outbound marketing: independent or complementary strategies. Expert Journal of Marketing, 7(1).

Daramola, O., Stålhane, T., Omoronyia, I., & Sindre, G.: Using ontologies and machine learning for hazard identification and safety analysis. In Managing requirements knowledge (pp. 117-141). Springer, Berlin, Heidelberg (2013).

Daugherty, T., Li, H. and Biocca, F., 2008. Consumer learning and the effects of virtual experience relative to indirect and direct product experience. Psychology & Marketing, 25(7), pp.568-586.

Debortoli, S., Müller, O. and vom Brocke, J., 2014. Comparing business intelligence and big data skills. *Business & Information Systems Engineering*, *6*(5), pp.289-300.

Deffner, A., Karachalis, N., Psatha, E., Metaxas, T. and Sirakoulis, K., 2020. City marketing and planning in two Greek cities: plurality or constraints. European Planning Studies, pp.1-22.

Delgoshaei, P., Austin, M.A. and Veronica, D., 2017, April. Semantic models and rule-based reasoning for fault detection and diagnostics: Applications in heating, ventilating and air conditioning systems. In the Twelfth International Conference on Systems (ICONS 2017) (pp. 48-53).

DeSoto, L.J., Pearson, M.M., Adkinson, K.A., Achanta, V.R. and Peng, F., Experian Information Solutions LLC, 2019. Systems and methods for providing a direct marketing campaign planning environment. U.S. Patent 10,311,466.

Dieterle, S. and Bergmann, R., 2014, September. A hybrid CBR-ANN approach to the appraisal of internet domain names. *In International Conference on Case-Based Reasoning* (pp. 95-109). Springer, Cham.

Differding, C., Hoisl, B. and Lott, C.M., 1996. Technology package for the goal question metric paradigm.

Du, M., Liu, N. and Hu, X.: Techniques for interpretable machine learning. Communications of the ACM, 63(1), 68-77, (2019).

Ducamp, G., Bonnard, P. and Wuillemin, P.H., 2020, May. Uncertain reasoning in rule-based systems using prm. In FLAIRS 33-33rd Florida Artificial Intelligence Research Society Conference (pp. 617-620). AAAI.

Durstewitz, D., Koppe, G. and Meyer-Lindenberg, A., 2019. Deep neural networks in psychiatry. Molecular psychiatry, 24(11), pp.1583-1598.

Eisenstadt, V., Langenhan, C. and Althoff, K.D., 2019, September. Flea-cbr–a flexible alternative to the classic 4r cycle of case-based reasoning. In International Conference on Case-Based Reasoning (pp. 49-63). Springer, Cham.

Elsalamony, H.A., 2014. Bank direct marketing analysis of data mining techniques. International Journal of Computer Applications, 85(7).

El-Sappagh, S.H. and Elmogy, M., 2015. Case based reasoning: case representation methodologies. *International Journal of Advanced Computer Science and Applications*, *6*(11), pp.192-208.

Erel, I., Stern, L.H., Tan, C. and Weisbach, M.S., 2018. *Selecting Directors Using Machine Learning* (No. w24435). National Bureau of Economic Research.

Faia, R., Pinto, T., Abrishambaf, O., Fernandes, F., Vale, Z. and Corchado, J.M., 2017. Case based reasoning with expert system and swarm intelligence to determine energy reduction in buildings energy management. Energy and Buildings, 155, pp.269-281.

Fei, L. and Feng, Y., 2020. A novel retrieval strategy for case-based reasoning based on attitudinal Choquet integral. Engineering Applications of Artificial Intelligence, 94, p.103791.

Feng, X., Weng, C., He, X., Han, X., Lu, L., Ren, D. and Ouyang, M., 2019. Online state-of-health estimation for Li-ion battery using partial charging segment based on support vector machine. IEEE Transactions on Vehicular Technology, 68(9), pp.8583-8592.

Firman, F., 2019. Pengaruh Direct Marketing Terhadap Keputusan Pembelian Kamar di Imelda Hotel-Waterpark-Convention Padang (Doctoral dissertation, Universitas Negeri Padang).

Fischer, M., Richards, S. and Bolt, B.G., 2020. Direct Marketing of Beef, Pork, or Lamb: Is This an Opportunity for You?.

Flici, A., 2011. A conceptual framework for the direct marketing process using business intelligence (Doctoral dissertation, Brunel University Brunel Business School PhD Theses).

Flici, A.: A conceptual framework for the direct marketing process using business intelligence (Doctoral dissertation, Brunel University Brunel Business School PhD Theses), (2011).

Flores, A., Tito, H. and Silva, C., 2019, November. CBRi: a case-based reasoning-inspired approach for univariate time series imputation. *In 2019 IEEE Latin American Conference on Computational Intelligence* (LA-CCI) (pp. 1-6). IEEE.

Fu, W., Wang, K., Zhang, C. and Tan, J., 2019. A hybrid approach for measuring the vibrational trend of hydroelectric unit with enhanced multi-scale chaotic series analysis and optimized least squares support vector machine. Transactions of the Institute of Measurement and Control, 41(15), pp.4436-4449.

Gallina, B., 2020. Quantitative evaluation of tailoring within SPICE-compliant security-informed safety-oriented process lines. *Journal of Software: Evolution and Process*, 32(3), p.e2212.

Gao, X., Shan, C., Hu, C., Niu, Z. and Liu, Z., 2019. An adaptive ensemble machine learning model for intrusion detection. IEEE Access, 7, pp.82512-82521.

Gardan, D., Epuran, G. and Gardan, I.P., 2015. The Use of Direct Marketing Techniques in Sport Marketing. Annals of Spiru Haret University. Economic Series, 15(4), pp.49-59.

Gemechu, G., 2020. Performance Comparison of Classifiers for Prospective Buyers Identification in ethio telecom Mobile Cross-Selling Market (Doctoral dissertation, Addis Ababa University).

Ghahramani, M., Qiao, Y., Zhou, M., Hagan, A.O. and Sweeney, J., 2020. AI-based modelling and data-driven evaluation for smart manufacturing processes. *IEEE/CAA Journal of Automatica Sinica,* 7(4), pp.1026-1037.

Ghasemaghaei, M. and Calic, G., 2019. Does big data enhance firm innovation competency? The mediating role of data-driven insights. *Journal of Business Research*, 104, pp.69-84.

Gigih, P., Andriani, C., Istiadi, I. and Abd Wahab, M.H., 2020. Case Service System at the Child Welfare Institution using the Case-Based Reasoning Method. *International Journal of Electrical, Energy and Power System Engineering*, 3(2), pp.40-45.

Gillwald, A., 2005. Good intentions, poor outcomes: Telecommunications reform in South Africa. Telecommunications Policy, 29(7), pp.469-491.

Goebel, R., Chander, A., Holzinger, K., Lecue, F., Akata, Z., Stumpf, S., Kieseberg, P. and Holzinger, A.: Explainable AI: the new 42?. *In International Cross-Domain Conference for Machine Learning and Knowledge Extraction* (pp. 295-303). Springer, Cham (2018).

Goel, A.K. and Diaz-Agudo, B., 2017, February. What's hot in case-based reasoning. *In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* (pp. 5067-5069).

Golding, A.R. and Rosenbloom, P.S., 1991, July. Improving Rule-Based Systems Through Case-Based Reasoning. In AAAI (Vol. 1, pp. 22-27).

Gordon, J., Perrey, J. and Spillecke, D., 2013. Big data, analytics and the future of marketing and sales. McKinsey: Digital Advantage.

Gorgevik, D., Cakmakov, D. and Radevski, V., 2001, June. Handwritten digit recognition by combining support vector machines using rule-based reasoning. *In Proceedings of the 23rd International Conference on Information Technology Interfaces*, 2001. ITI 2001. (pp. 139-144). IEEE.

Grandhi, B., Patwa, N. and Saleem, K., 2020. Data-driven marketing for growth and profitability. *EuroMed Journal of Business*.

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F. and Pedreschi, D.: A survey of methods for explaining black box models. ACM computing surveys (CSUR), 51(5), 1-42 (2018).

Gunning, D., 2017. Explainable artificial intelligence (xai). *Defense Advanced Research Projects Agency (DARPA), nd Web*.

Hadzic, S., 2019. A Global South Perspective on Alternative Spectrum Policy.

Hawlk, K., 2018. Outbound Versus Inbound Marketing: Which Strategy Is Best? *Journal of Financial Planning*, 31(6), pp.30-31.

Hegdal, S.S. and Kofod-Petersen, A., 2019. A CBR-ANN hybrid for dynamic environments. CEUR Workshop Proceedings.

Himanen, L., Jäger, M.O., Morooka, E.V., Canova, F.F., Ranawat, Y.S., Gao, D.Z., Rinke, P. and Foster, A.S., 2020. DScribe: Library of descriptors for machine learning in materials science. Computer Physics Communications, 247, p.106949.

Hinson, E., VanZyl, H., Nimako, S.G., Chinje, N. and Asiamah, E., 2016. Extending the four-stage brand loyalty framework in African Telecoms. *African Journal of Business and Economic Research*, 11(2-3), pp.53-82.

Hoffman, R.R., Mueller, S.T., Klein, G. and Litman, J., 2018. Metrics for Explainable AI: Challenges and Prospects. arXiv preprint arXiv:1812.04608.

Hollon, M.F., 2004. Direct-to-consumer marketing of prescription drugs. *CNS drugs*, *18*(2), pp.69-77.

Holzinger, A., Plass, M., Kickmeier-Rust, M., Holzinger, K., Crişan, G.C., Pintea, C.M. and Palade, V., 2019. Interactive machine learning: experimental evidence for the human in the algorithmic loop. Applied Intelligence, 49(7), pp.2401-2414.

Hossain, M.A., Akter, S. and Yanamandram, V., 2020. Revisiting customer analytics capability for data-driven retailing. *Journal of Retailing and Consumer Services*, 56, p.102187.

Hussain, A. and Ferneley, E., 2008, November. Usability metric for mobile application: a goal question metric (GQM) approach. *In Proceedings of the 10th International Conference on Information Integration and Web-Based Applications & Services* (pp. 567-570).

Huynh, A.K., Hamilton, A.B., Farmer, M.M., Bean-Mayberry, B., Stirman, S.W., Moin, T. and Finley, E.P., 2018. A pragmatic approach to guide implementation evaluation research: strategy mapping for complex interventions. Frontiers in public health, 6, p.134.

Hyseni, L., and Dika, Z.: An integrated framework of conceptual modelling for performance improvement of the information systems. *In 2017 Seventh International Conference on Innovative Computing Technology* (INTECH), pp. 174-180. IEEE, (2017).

Isa, M., 2019. Why is South Africa still waiting for cheaper data? Finweek, 2019(15 August), pp.44-45.

Iwuchukwu, E.I., Amoakoh, E.O. and Chipunza, C., 2019. A theoretical analysis on the impact of relationship marketing practices on customer retention: A case of transformed small grocery retail shops in Mangaung Municipality area in South Africa.

Iyengar, R., Ansari, A. and Gupta, S., 2007. A model of consumer learning for service quality and usage. *Journal of Marketing Research*, *44*(4), pp.529-544.

Jakkula, V., 2006. Tutorial on support vector machine (svm). School of EECS, Washington State University, 37.

James, G., Witten, D., Hastie, T. and Tibshirani, R., 2013. An introduction to statistical learning (Vol. 112, p. 18). New York: Springer.

Jetinai, K., 2018, March. Rule-based reasoning for resource recommendation in personalized e-learning. In 2018 International Conference on Information and Computer Technologies (ICICT) (pp. 150-154). IEEE.

Jiang, W., 2020. Applications of deep learning in stock market prediction: recent progress. arXiv preprint arXiv:2003.01859.

Johannesson, P. and Perjons, E., 2014. An introduction to design science. Springer.

Johannesson, P. and Perjons, E., 2014. A method framework for design science research. In *An Introduction to Design Science* (pp. 75-89). Springer, Cham.

John, D.A. and John, R.R., 2010, March. A framework for medical diagnosis using hybrid reasoning. *In Proceedings of the International Multi-Conference of Engineers and Computer Scientists (*Vol. 1).

Jozdani, S.E., Johnson, B.A. and Chen, D., 2019. Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification. Remote Sensing, 11(14), p.1713.

Karim, M., & Rahman, R.: Decision Tree and Naïve Bayes Algorithm for Classification and Generation of Actionable Knowledge for Direct Marketing. *Journal of Software Engineering and Applications*, 6, 196-206, (2013).

Karimi, F., Sultana, S., Babakan, A.S. and Suthaharan, S., 2019. An enhanced support vector machine model for urban expansion prediction. Computers, Environment and Urban Systems, 75, pp.61-75.

Kau, F.M. and Kogeda, O.P., 2019, October. Impact of Subscription Fraud in Mobile Telecommunication Companies. In *2019 Open Innovations (OI)* (pp. 42-47). IEEE.

Keane, M.T. and Kenny, E.M., 2019. The twin-system approach as one generic solution for XAI: An overview of ANN-CBR twins for explaining deep learning. arXiv preprint arXiv:1905.08069

Keane, M.T. and Kenny, E.M., 2019. How case-based reasoning explained neural networks: an XAI survey of post-hoc explanation-by-example in ANN-CBR twins. arXiv preprint arXiv:1905.07186.

Ken Peffers, Tuure Tuunanen, Marcus A. Rothenberger, Samir Chatterjee, 2007, A Design Science Research Methodology for Information Systems Research, Available at (https://www.jstor.org/stable/40398896?seq=1#page_scan_tab_contents).

Ketler, K., 1993. Case-based reasoning: an introduction. Expert Systems with Applications, 6(1), pp.3-8.

Khan, H., 2020. Is marketing agility important for emerging market firms in advanced markets? International Business Review, 29(5), p.101733.

Khandelwal, K. and Sharma, D.P., 2013. Hybrid Reasoning Model for Strengthening the problem-solving capability of Expert Systems. Int. J. Adv. Comput. Sci. Appl. (IJACSA), 4.

Khosravani, M.R. and Nasiri, S., 2020. Injection moulding manufacturing process: Review of case-based reasoning applications. *Journal of Intelligent Manufacturing*, 31(4), pp.847-864.

Kim, G., Chae, B.K. and Olson, D.L., 2013. A support vector machine (SVM) approach to imbalanced datasets of customer responses: comparison with other customer response models. Service Business, 7(1), pp.167-182.

Kivunja, C. and Kuyini, A.B., 2017. Understanding and Applying Research Paradigms in Educational Contexts. *International Journal of Higher Education*, *6*(5), pp.26-41.

Kogan, A., Mayhew, B.W. and Vasarhelyi, M.A., 2019. Audit data analytics research—An application of design science methodology. Accounting Horizons, 33(3), pp.69-73.

Koh, H.C. and Tan, G., 2011. Data mining applications in healthcare. *Journal of Healthcare Information Management*, *19*(2), p.65.

Kolahdouz Rahimi, S., 2013. A comparative study of model transformation approaches through a systematic procedural framework and goal question metrics paradigm (Doctoral dissertation, King's College London (University of London)).

Kourtis, G., Kavakli, E. and Sakellariou, R., 2019. A rule-based approach founded on description logics for Industry 4.0 smart factories. IEEE Transactions on Industrial Informatics, 15(9), pp.4888-4899.

Krems, R.V., 2019. Bayesian machine learning for quantum molecular dynamics. Physical Chemistry Chemical Physics, 21(25), pp.13392-13410.

Kumar, A., Shankar, R. and Aljohani, N.R., 2020. A big data driven framework for demand-driven forecasting with effects of marketing-mix variables. Industrial Marketing Management, 90, pp.493-507.

Krupitzer, C., Drechsel, G., Mateja, D., Pollkläsener, A., Schrage, F., Sturm, T., Tomasovic, A. and Becker, C., 2018, March. Using spreadsheet-defined rules for reasoning in self-adaptive systems. *In 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)* (pp. 289-294). IEEE.

Kuo, P.H. and Huang, C.J., 2018. A high precision artificial neural networks model for short-term energy load forecasting. Energies, 11(1), p.213.

Ładyżyński, P., Żbikowski, K. and Gawrysiak, P., 2019. Direct marketing campaigns in retail banking with the use of deep learning and random forests. Expert Systems with Applications, 134, pp.28-35.

Lahmiri, S. and Shmuel, A., 2019. Detection of Parkinson's disease based on voice patterns ranking and optimized support vector machine. Biomedical Signal Processing and Control, 49, pp.427-433.

Lamy, J.B., Sekar, B., Guezennec, G., Bouaud, J. and Séroussi, B., 2019. Explainable artificial intelligence for breast cancer: A visual case-based reasoning approach. Artificial intelligence in medicine, 94, pp.42-53.

Lawi, A., Velayaty, A.A. and Zainuddin, Z.: August. On identifying potential direct marketing consumers using adaptive boosted support vector machine. *In 2017 4th International Conference on Computer Applications and Information Processing Technology* (CAIPT) (pp. 1-4). IEEE, (2017).

Lee, C.H., Chen, C.H., Li, F. and Shie, A.J., 2020. Customized and knowledge-centric service design model integrating case-based reasoning and TRIZ. Expert Systems with Applications, 143, p.113062.

Lessmann, S., Stahlbock, R. and Crone, S.F., 2006, July. Genetic algorithms for support vector machine model selection. *In the 2006 IEEE International Joint Conference on Neural Network Proceedings* (pp. 3063-3069). IEEE.

Lezama, O.B.P., 2020. Data Processing for Direct Marketing Through Big Data. Computational Vision and Bio-Inspired Computing: ICCVBIC 2019, 1108, p.187.

Li, C., Pavlu, V., Aslam, J., Wang, B. and Qin, K., 2019, September. Learning to Calibrate and Rerank Multi-label Predictions. *In Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 220-236). Springer, Cham.

Li, M., Huang, Y. and Sinha, A., 2020. Data-Driven Promotion Planning for Paid Mobile Applications. Information Systems Research, 31(3), pp.1007-1029.

Lian-Ying, Z., Amoh, D.M., Boateng, L.K. and Okine, A.A.: Combined appetency and upselling prediction scheme in telecommunication sector using support vector machines. *International Journal of Modern Education and Computer Science*, 11(6),1, (2019).

Liaw, A. and Wiener, M., 2002. Classification and regression by random forest. R news, 2(3), pp.18-22.

Lin, X., 2019. A Road Network Traffic State Identification Method Based on Macroscopic Fundamental Diagram and Spectral Clustering and Support Vector Machine. Mathematical Problems in Engineering, 2019.

Ling, H., Qian, C., Kang, W., Liang, C. and Chen, H., 2019. Combination of Support Vector Machine and K-Fold cross validation to predict compressive strength of concrete in marine environment. Construction and Building Materials, 206, pp.355-363.

Low, S.A. and Vogel, S.J., 2011. Direct and intermediated marketing of local foods in the United States.

Lubeniqi, G., 2020. Advantages, Disadvantages and the Performance of Foreign Direct Investment in the Republic of Kosovo 2008-2019. *Prizren Social Science Journal*, 4(1), pp.14-19.

Mabkhot, M.M., Al-Samhan, A.M. and Hidri, L., 2019. An ontology-enabled case-based reasoning decision support system for manufacturing process selection. Advances in Materials Science and Engineering, 2019.

Maind, S.B. and Wankar, P., 2014. Research paper on basic of artificial neural network. *International Journal on Recent and Innovation Trends in Computing and Communication*, 2(1), pp.96-100.

McCoy, C.L., 2019. A Systematic Review Characterizing Farm Direct Marketing Challenges, Strategies, and Opportunities (Doctoral dissertation).

Micheaux, A. and Bosio, B., 2019. Customer journey mapping as a new way to teach data-driven marketing as a service. *Journal of Marketing Education*, 41(2), pp.127-140.

Miguéis, V.L., Camanho, A.S. and Borges, J., 2017. Predicting direct marketing response in banking: comparison of class imbalance methods. Service Business, 11(4), pp.831-849.

Mijwel, M.M., 2018. Artificial neural networks advantages and disadvantages. Retrieved from LinkedIn: https://www. linkedin. com/pulse/artificial-neuralnet works-advantages-disadvantages-maad-m-mijwel.

Mironova, E.Y., 2018. OVERVIEW OF STARLINK–WIRELESS ACCESS GLOBAL NETWORK. Международный научно-исследовательский журнал, (12 (78) Часть 1), pp.117-120.

Mkhize, S. and Ellis, D., 2020. Creativity in marketing communication to overcome barriers to organic produce purchases: The case of a developing nation. *Journal of Cleaner Production*, 242, p.118415.

Molchanov, D., Ashukha, A. and Vetrov, D., 2017. Variational dropout sparsifies deep neural networks. arXiv preprint arXiv:1701.05369.

Montavon, G., Samek, W. and Müller, K.R., 2018. Methods for interpreting and understanding deep neural networks. Digital Signal Processing, 73, pp.1-15.

Moore, M.M., Slonimsky, E., Long, A.D., Sze, R.W. and Iyer, R.S., 2019. Machine learning concepts, concerns and opportunities for a paediatric radiologist. Paediatric radiology, 49(4), pp.509-516.

Morgan, S. and Govender, K., 2017. Exploring customer loyalty in the South African mobile telecommunications sector. Cogent Business & Management, 4(1), p.1273816.

Movilla-Pateiro, L., Mahou-Lago, X.M., Doval, M.I. and Simal-Gandara, J., 2020. Toward a sustainable metric and indicators for the goal of sustainability in agricultural and food production. Critical Reviews in Food Science and Nutrition, pp.1-22.

Mpwanya, M.F. and Letsoalo, M.E., 2019. Relationship between service quality, customer satisfaction and behavioural intentions in South Africa’s mobile telecommunication industry. *African Journal of Business and Economic Research*, 14(2), pp.67-89.

Muhammedrisaevna, T.M.S., Bayazovna, G.N. and Kakhramonovna, D.A., 2020. Goal and objectives of integrated marketing communications. Economics, (2 (45)).

Muketha, G., Ondago, C.O. and Ondimu, K.O., 2008. Complexity metrics for measuring the understandability and maintainability of business process models using goal-question-metric (GQM).

Musa, A.G., Daramola, O., Owoloko, E.A. and Olugbara, O.O., 2013. A neural-CBR system for real property valuation. Journal of Emerging Trends in Computing and Information Sciences, 4(8), pp.611-622.

Nachev, A.: Application of data mining techniques for direct marketing. Computational Models for Business and Engineering Domains, 86-95, (2015).

Naderpour, H., Rafiean, A.H. and Fakharian, P., 2018. Compressive strength prediction of environmentally friendly concrete using artificial neural networks. *Journal of Building Engineering*, 16, pp.213-219.

Nagappan, N., Murphy, B. and Basili, V., 2008, May. The influence of organizational structure on software quality. *In 2008 ACM/IEEE 30th International Conference on Software Engineering* (pp. 521-530). IEEE.

Nakhjiri, N., Salamó, M. and Sànchez-Marrè, M., 2020. Reputation-Based Maintenance in Case-Based Reasoning. Knowledge-Based Systems, 193, p.105283.

Nguyen, V.V., Pham, B.T., Vu, B.T., Prakash, I., Jha, S., Shahabi, H., Shirzadi, A., Ba, D.N., Kumar, R., Chatterjee, J.M. and Tien Bui, D., 2019. Hybrid machine learning approaches for landslide susceptibility modelling. Forests, 10(2), p.157.

Ngwenya, M., 2017, December. Analysing service quality using customer expectations and perceptions in the South African telecommunication industry. In 2017 *IEEE International Conference on Industrial Engineering and Engineering Management* (IEEM) (pp. 1094-1097). IEEE.

Nilsson, N.J., 1998. Introduction to Machine Learning. An Early Draft of a Proposed Textbook (1998). Software available at http://robotics. stanford. edu/people/nilsson/mlbook. html.

Niu, W.J., Feng, Z.K., Feng, B.F., Min, Y.W., Cheng, C.T. and Zhou, J.Z., 2019. Comparison of multiple linear regression, artificial neural network, extreme learning machine, and support vector machine in deriving operation rule of hydropower reservoir. Water, 11(1), p.88.

Olsson, T., Gillblad, D., Funk, P. and Xiong, N., 2014. Case-based reasoning for explaining probabilistic machine learning. *International Journal of Computer Science and Information Technology*, *6*(2), pp.87-101.

Ouedraogo, I., Defourny, P. and Vanclooster, M.: Application of random forest regression and comparison of its performance to multiple linear regression in modelling groundwater nitrate concentration at the African continent scale. *Hydrogeology Journal*, 27(3), 1081-1098, (2019).

Paelo, A. and Robb, G., 2020. Comparative approaches to key issues in the economic regulation of telecommunications markets in South Africa, Tanzania, Zambia, and Zimbabwe. *WIDER Working Paper Series*, (wp-2020-84).

Park, Y.S., Konge, L. and Artino Jr, A.R., 2020. The positivism paradigm of research. Academic Medicine, 95(5), pp.690-694.

Patrutiu-Baltes, L., 2016. Inbound Marketing-the most important digital marketing strategy. Bulletin of the Transylvania University of Brasov. Economic Sciences. Series V, 9(2), p.61.

Pau, L.F., 2011. Telecommunications in South Africa: Who is Actually Regulating? Available at SSRN 1917335.

Pereira, F.C. and Borysov, S.S.: Machine learning fundamentals. In Mobility Patterns, Big Data and Transport Analytics, pp. 9-29. Elsevier, (2019).

Petersen, R. and Daramola, O., 2020, July. Towards Explainable Direct Marketing in the Telecom Industry through Hybrid Machine Learning. In *International Conference on Computational Science and Its Applications* (pp. 471-486). Springer, Cham.

Pham, T.B., Singh, S.K. and Ly, H.B., 2020. Using Artificial Neural Network (ANN) for prediction of soil coefficient of consolidation. Vietnam Journal of Earth Sciences, 42(4), pp.311-319.

Pintelas, E., Liaskos, M., Livieris, I.E., Kotsiantis, S. and Pintelas, P., 2020. Explainable Machine Learning Framework for Image Classification Problems: Case Study on Glioma Cancer Prediction. Journal of Imaging, 6(6), p.37.

Pittz, T.G., Steiner, S.D. and Pennington, J.R., 2019. An Ethical Marketing Approach to Wicked Problems: Macromarketing for the Common Good. *Journal of Business Ethics*, pp.1-10.

Pouladi, N., Møller, A.B., Tabatabai, S. and Greve, M.H., 2019. Mapping soil organic matter contents at field level with Cubist, Random Forest and kriging. Geoderma, 342, pp.85-92.

Prašljivić, A. and Ramić, A., 2019, May. Targeted multichannel marketing campaign in Telecommunications. In 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (pp. 387-391). IEEE.

Prentzas, J. and Hatzilygeroudis, I., 2007. Categorizing approaches combining rule-based and case-based reasoning. Expert Systems, 24(2), pp.97-122.

Putit, L. and Abdullah, M.F., 2019. Exploring relationship marketing tactics & customer loyalty: issues in the Malaysian mobile telecommunication sector. *Journal of International Business, Economics and Entrepreneurship* (JIBE), 4(1), pp.21-26.

Qu, K., Guo, F., Liu, X., Lin, Y. and Zou, Q., 2019. Application of machine learning in microbiology. Frontiers in Microbiology, 10, p.827.

Quinn, S., Bond, R. and Nugent, C., 2017. Ontological modelling and rule-based reasoning for the provision of personalized patient education. Expert Systems, 34(2), p.e12134.

Rahman, A., Slamet, C., Darmalaksana, W., Gerhana, Y.A. and Ramdhani, M.A., 2018. Expert System for Deciding a Solution of Mechanical Failure in a Car using Case-based Reasoning. In IOP Conf. Ser. Mater. Sci. Eng (Vol. 288, No. 1, p. 012011).

Rajput, H. and Singh, L.K., 2011. Improvement of software quality attributes in object-oriented analysis and design phase using goal-question-metric paradigm. *Journal of Software Engineering and Applications*, 4(6), p.345.

Rakib, A. and Uddin, I., 2019. An efficient rule-based distributed reasoning framework for resource-bounded systems. Mobile Networks and Applications, 24(1), pp.82-99.

Relich, M. and Pawlewski, P., 2018. A case-based reasoning approach to cost estimation of new product development. Neurocomputing, 272, pp.40-45.

Rem, B.S., Käming, N., Tarnowski, M., Asteria, L., Fläschner, N., Becker, C., Sengstock, K. and Weitenberg, C., 2019. Identifying quantum phase transitions using artificial neural networks on experimental data. Nature Physics, 15(9), pp.917-920.

Renton, M. and Simmonds, H., 2019. Effectuation and Morphogenesis in the New Zealand Fairtrade Marketing System. *Journal of Macromarketing*, 39(4), pp.385-399.

Reutterer, T., Mild, A., Natter, M. and Taudes, A., 2006. A dynamic segmentation approach for targeting and customizing direct marketing campaigns. *Journal of Interactive Marketing*, *20*(3-4), pp.43-57.

Rezvani, S., Wang, X. and Pourpanah, F., 2019. Intuitionistic fuzzy twin support vector machines. IEEE Transactions on Fuzzy Systems, 27(11), pp.2140-2151.

Richhariya, B. and Tanveer, M., 2020. A reduced universum twin support vector machine for class imbalance learning. Pattern Recognition, 102, p.107150.

Richter, M.M. and Weber, R.O., 2016. Case-based reasoning (p. 27). Springer-Verlag Berlin An.

Roguet, A., Eren, A.M., Newton, R.J. and McLellan, S.L., 2018. Faecal source identification using random forest. Microbiome, 6(1), p.185.

Ruangthong, P. and Jaiyen, S.: Bank direct marketing analysis of asymmetric information based on machine learning. *In 2015 12th International Joint Conference on Computer Science and Software Engineering* (JCSSE), pp. 93-96. IEEE (2015).

Ruder, S., 2017. An overview of multi-task learning in deep neural networks. arXiv preprint arXiv:1706.05098.

Sadq, Z.M., Othman, B. and Khorsheed, R.K., 2019. The impact of tourism marketing in enhancing competitive capabilities. *African Journal of Hospitality, Tourism and Leisure*, 8(5), pp.1-11.

Samek, W., Wiegand, T. and Müller, K.R., 2017. Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. arXiv preprint arXiv:1708.08296.

Santucci, L.C., McHugh, R.K. and Barlow, D.H., 2012. Direct-to-consumer marketing of evidence-based psychological interventions: Introduction.

Sathya, R. and Abraham, A., 2013. Comparison of supervised and unsupervised learning algorithms for pattern classification. *International Journal of Advanced Research in Artificial Intelligence*, 2(2), pp.34-38.

Saul, C. and Wuttke, H.D., 2013. An adaptation model for personalized e-assessments. *International Journal of Emerging Technologies in Learning* (iJET), 8(2013).

Schwabe, K., Teizer, J. and König, M., 2019. Applying rule-based model-checking to construction site layout planning tasks. Automation in construction, 97, pp.205-219.

Seeger, J., Bröring, A., Pahl, M.O. and Sakic, E., 2019, June. Rule-Based Translation of Application-Level QoS Constraints into SDN Configurations for the IoT. In 2019 European Conference on Networks and Communications (EuCNC) (pp. 432-437). IEEE.

Sekhar, C.R. and Madhu, E., 2016. Mode choice analysis using random forest decision trees. Transportation Research Procedia, 17, pp.644-652.

Seydoux, N., Drira, K., Hernandez, N. and Monteil, T., 2019. EDR: A Generic Approach for the Dynamic Distribution of Rule-Based Reasoning in a Cloud-Fog continuum. *Semantic Web Journal.*

Shen, L., Yan, H., Fan, H., Wu, Y. and Zhang, Y., 2017. An integrated system of text mining technique and case-based reasoning (TM-CBR) for supporting green building design. Building and Environment, 124, pp.388-401.

Sheng, Y., Zhang, J., Wang, C., Yin, F.F., Wu, Q.J. and Ge, Y., 2019. Incorporating case-based reasoning for radiation therapy knowledge modelling: a pelvic case study. Technology in cancer research & treatment, 18, p.1533033819874788.

Sheng-sheng, C.A.I. and Fan-liang, B.U., 2019. An Assistant Decision-making Algorithm for Public Security Emergencies# br# Based on Case-based Reasoning and Rule-based Reasoning. Computer and Modernization, (09), p.7.

Shi, H., Liu, S., Chen, J., Li, X., Ma, Q. and Yu, B., 2019. Predicting drug-target interactions using Lasso with random forest based on evolutionary information and chemical structure. Genomics, 111(6), pp.1839-1852.

Shukla, A.K., 2020. Identification of cancerous gene groups from microarray data by employing adaptive genetic and support vector machine technique. Computational Intelligence, 36(1), pp.102-131.

Shwartz-Ziv, R. and Tishby, N., 2017. Opening the black box of deep neural networks via information. arXiv preprint arXiv:1703.00810.

Siau, K. and Wang, W., 2018. Building trust in artificial intelligence, machine learning, and robotics. *Cutter Business Technology Journal*, *31*(2), pp.47-53.

Silva, G.C., Carvalho, E.E. and Caminhas, W.M., 2020. An artificial immune systems approach to case-based reasoning applied to fault detection and diagnosis. Expert Systems with Applications, 140, p.112906.

Simeone, O., 2018. A very brief introduction to machine learning with applications to communication systems. IEEE Transactions on Cognitive Communications and Networking, 4(4), pp.648-664.

Siregar, A.M., Faisal, S., Handayani, H.H. and Jalaludin, A., 2020. Classification Data for Direct Marketing using Deep Learning. *Scientific Journal of PPI-UKM*, 7(2), pp.33-37.

Slik, J. and Bhulai, S., 2019. Data-driven Direct Marketing via Approximate Dynamic Programming. DATA ANALYTICS 2019, p.72.

Smiti, A. and Elouedi, Z., 2020. Dynamic maintenance case base using knowledge discovery techniques for case-based reasoning systems. Theoretical Computer Science, 817, pp.24-32.

Sokol, K. and Flach, P.A., 2019, January. Counterfactual explanations of machine learning predictions: opportunities and challenges for AI safety. In SafeAI@ AAAI.

Somdyala, B., Rananga, S., Mfupe, L., Masonta, M. and Mekuria, F., 2017, May. Spectrum regulation for future internet networks in developing economies. *In 2017 IST-Africa Week Conference* (IST-Africa) (pp. 1-12). IEEE.

Song, C. and Alkhalifah, T., 2020, December. Identifying microseismic events in time-reversed source images using support vector machine. *In 82nd EAGE Annual Conference & Exhibition* (Vol. 2020, No. 1, pp. 1-5). European Association of Geoscientists & Engineers.

Song, M., Gofuku, A. and Lind, M., 2020. Model-based and rule-based synthesis of operating procedures for planning severe accident management strategies. Progress in Nuclear Energy, 123, p.103318.

Stafoggia, M., Bellander, T., Bucci, S., Davoli, M., De Hoogh, K., De'Donato, F., Gariazzo, C., Lyapustin, A., Michelozzi, P., Renzi, M. and Scortichini, M., 2019. Estimation of daily PM10 and PM2. 5 concentrations in Italy, 2013–2015, using a spatiotemporal land-use random-forest model. Environment international, 124, pp.170-179.

Statnikov, A., Wang, L. and Aliferis, C.F., 2008. A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. BMC bioinformatics, 9(1), p.319.

Stone, M., Woodcock, N., Ekinci, Y., Aravopoulou, E. and Parnell, B., 2019. SCHEMA: Information on marketing and customer engagement performance–reality versus dreams. The Bottom Line, 32(1), pp.98-116.

Su, J., Vargas, D.V. and Sakurai, K., 2019. One pixel attack for fooling deep neural networks. IEEE Transactions on Evolutionary Computation, 23(5), pp.828-841.

Sun, Y., Huang, X., Kroening, D., Sharp, J., Hill, M. and Ashmore, R., 2018. Testing deep neural networks. arXiv preprint arXiv:1803.04792.

Tajul Urus, S., 2013. Living with enterprise resource planning (ERP): an investigation of end user problems and coping mechanisms.

Tan, X., Su, S., Huang, Z., Guo, X., Zuo, Z., Sun, X. and Li, L., 2019. Wireless sensor networks intrusion detection based on SMOTE and the random forest algorithm. Sensors, 19(1), p.203.

Tang, F., Kawamoto, Y., Kato, N. and Liu, J., 2019. Future intelligent and secure vehicular network toward 6G: Machine-learning approaches. Proceedings of the IEEE, 108(2), pp.292-307.

Tang, J., Yuan, F., Shen, X., Wang, Z., Rao, M., He, Y., Sun, Y., Li, X., Zhang, W., Li, Y. and Gao, B., 2019. Bridging biological and artificial neural networks with emerging neuromorphic devices: Fundamentals, progress, and challenges. Advanced Materials, 31(49), p.1902761.

Tang, L., Tian, Y. and Pardalos, P.M., 2019. A novel perspective on multiclass classification: regular simplex support vector machine. Information Sciences, 480, pp.324-338.

Tang, L. and Volponi, A.J., 2019. Intelligent reasoning for gas turbine fault isolation and ambiguity resolution. Journal of Engineering for Gas Turbines and Power, 141(4).

Thangaramya, K., Kulothungan, K., Indira Gandhi, S., Selvi, M., Santhosh Kumar, S.V.N. and Arputharaj, K., 2020. Intelligent fuzzy rule-based approach with outlier detection for secured routing in WSN. Soft Computing, pp.1-15.

Tharwat, A., Hassanien, A.E. and Elnaghi, B.E., 2017. A BA-based algorithm for parameter optimization of support vector machine. Pattern Recognition Letters, 93, pp.13-22.

Thike, P.H., Xu, Z., Cheng, Y., Jin, Y. and Shi, P., 2019. Materials failure analysis utilizing rule-case based hybrid reasoning method. Engineering Failure Analysis, 95, pp.300-311.

Thomas, D., Medoff, B., Anger, J. and Chughtai, B., 2020. Direct-to-consumer advertising for robotic surgery. Journal of robotic surgery, pp.1-4.

Tkachenko, Y., Kochenderfer, M.J. and Kluza, K., 2016, October. Customer simulation for direct marketing experiments. *In 2016 IEEE International Conference on Data Science and Advanced Analytics* (DSAA) (pp. 478-487). IEEE.

Too, E.C., Yujian, L., Njuki, S. and Yingchun, L., 2019. A comparative study of fine-tuning deep learning models for plant disease identification. Computers and Electronics in Agriculture, 161, pp.272-279.

Troisi, O., Maione, G., Grimaldi, M. and Loia, F., 2020. Growth hacking: Insights on data-driven decision-making from three firms. Industrial Marketing Management, 90, pp.538-557.

Tseng, M.L., Wu, K.J., Lim, M.K. and Wong, W.P., 2019. Data-driven sustainable supply chain management performance: a hierarchical structure assessment under uncertainties. *Journal of Cleaner Production*, 227, pp.760-771.

Uhrig, R.E., 1995, November. Introduction to artificial neural networks. *In Proceedings of IECON'95-21st Annual Conference on IEEE Industrial Electronics* (Vol. 1, pp. 33-37). IEEE.

Van Solingen, R., Basili, V., Caldiera, G. and Rombach, H.D., 2002. Goal question metric (gqm) approach. Encyclopaedia of Software Engineering.

Vandevijvere, Stefanie, Karuna Sagar, Bridget Kelly, and Boyd A. Swinburn. "Unhealthy food marketing to New Zealand children and adolescents through the internet." (2017): 32.

Vantara, H., Kalakota, R.,Partner, LiquidHub, 2018, Transform Telecom: A Data-Driven Strategy for Digital Transformation.

Vásquez-Morales, G., Martínez-Monterrubio, S., Moreno-Ger, P., & Recio-García, J.: Explainable Prediction of Chronic Renal Disease in the Colombian Population Using Neural Networks and Case-Based Reasoning. IEEE Access, 7, 152900-152910 (2019).

Vieira, S.M., Kaymak, U. and Sousa, J.M., 2010, July. Cohen's kappa coefficient as a performance measure for feature selection. In *International Conference on Fuzzy Systems* (pp. 1-8). IEEE.

Wadkar, M., Di Troia, F. and Stamp, M., 2020. Detecting malware evolution using support vector machines. Expert Systems with Applications, 143, p.113022.

Walker, R.S. and Brown, I., 2019. Big data analytics adoption: A case study in a large South African telecommunications organisation. *South African Journal of Information Management*, 21(1), pp.1-10.

Wan, S., Li, D., Gao, J. and Li, J., 2019. A knowledge-based machine tool maintenance planning system using case-based reasoning techniques. Robotics and Computer-Integrated Manufacturing, 58, pp.80-96.

Wang, A., Wang, Y. and Chen, Y., 2019. Hyperspectral image classification based on convolutional neural network and random forest. Remote sensing letters, 10(11), pp.1086-1094.

Wang, M. and Chen, H., 2020. Chaotic multi-swarm whale optimizer boosted support vector machine for medical diagnosis. Applied Soft Computing, 88, p.105946.

Wang, S., Cao, J. and Yu, P., 2020. Deep learning for spatio-temporal data mining: A survey. IEEE Transactions on Knowledge and Data Engineering.

Wen, W., Wu, C., Wang, Y., Chen, Y. and Li, H., 2016. Learning structured sparsity in deep neural networks. arXiv preprint arXiv:1608.03665.

Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V.X., Doshi-Velez, F., Jung, K., Heller, K., Kale, D., Saeed, M. and Ossorio, P.N., 2019. Do no harm: a roadmap for responsible machine learning for health care. Nature medicine, 25(9), pp.1337-1340.

Wimmo, F., 2020. Assessment of the Marketing Strategies Used by Public Telecommunication Companies: The Case of Tanzania Telecommunication Company Limited (Doctoral dissertation, Mzumbe University).

Wisaeng, K.: A comparison of different classification techniques for bank direct marketing. *International Journal of Soft Computing and Engineering* (IJSCE), 3(4), 116-119, (2013).

Wojciechowski, A. and Wrembel, R., 2020, June. On Case-Based Reasoning for ETL Process Repairs: Making Cases Fine-Grained. *In International Baltic Conference on Databases and Information Systems* (pp. 235-249). Springer, Cham.

Wongsuphasawat, K., Smilkov, D., Wexler, J., Wilson, J., Mane, D., Fritz, D., Krishnan, D., Viégas, F.B. and Wattenberg, M., 2017. Visualizing dataflow graphs of deep learning models in tensorflow. IEEE transactions on visualization and computer graphics, 24(1), pp.1-12.

Wright, J.C., 2019. Leveraging data insights to direct marketing strategy.
Xhemali, D., Hinde, C.J. and Stone, R.G., 2009. Naïve Bayes vs. decision trees vs. neural networks in the classification of training web pages.

Xu, L., Huang, C., Li, C., Wang, J., Liu, H. and Wang, X., 2020. An improved case-based reasoning method and its application in estimation of surface quality toward intelligent machining. *Journal of Intelligent Manufacturing*, pp.1-15.

Xu, L., Huang, C., Li, C., Wang, J., Liu, H. and Wang, X., 2020. A novel intelligent reasoning system to estimate energy consumption and optimize cutting parameters toward sustainable machining. *Journal of Cleaner Production*, p.121160.

Xuan, S., Liu, G., Li, Z., Zheng, L., Wang, S. and Jiang, C., 2018, March. Random forest for credit card fraud detection. *In 2018 IEEE 15th International Conference on Networking, Sensing and Control* (ICNSC) (pp. 1-6). IEEE.

Yahya, F., Walters, R.J. and Wills, G.B., 2015, November. Using goal-question-metric (GQM) approach to assess security in cloud storage. *In International Workshop on Enterprise Security* (pp. 223-240). Springer, Cham.

Yiyou, C., 2019. Innovation Direction of the Operation Mode of Insurance Marketing Channel.

Yu, C., Zhang, Z., Lin, C. and Wu, Y.J.: Can data-driven precision marketing promote user ad clicks? Evidence from advertising in WeChat moments. Industrial Marketing Management. doi: 10.1016/j.indmarman.2019.05.001 (2019).

Zafar, M.R. and Khan, N.M., 2019. DLIME: A deterministic local interpretable model-agnostic explanations approach for computer-aided diagnosis systems. arXiv preprint arXiv:1906.10263.

Zamani Joharestani, M., Cao, C., Ni, X., Bashir, B. and Talebiesfandarani, S., 2019. PM2. 5 Prediction based on random forest, XGBoost, and deep learning using multisource remote sensing data. Atmosphere, 10(7), p.373.

Zhang, J., Petersen, S.D., Radivojevic, T., Ramirez, A., Pérez-Manríquez, A., Abeliuk, E., Sánchez, B.J., Costello, Z., Chen, Y., Fero, M.J. and Martin, H.G., 2020. Combining mechanistic and machine learning models for predictive engineering and optimization of tryptophan metabolism. Nature communications, 11(1), pp.1-13.

Zhang, Q., Yu, H., Barbiero, M., Wang, B. and Gu, M., 2019. Artificial neural networks enabled by nanophotonics. Light: Science & Applications, 8(1), pp.1-14.

Zhang, W., Paudel, B., Wang, L., Chen, J., Zhu, H., Zhang, W., Bernstein, A. and Chen, H., 2019, May. Iteratively learning embeddings and rules for knowledge graph reasoning. *In The World Wide Web Conference* (pp. 2366-2377).

Zhang, Z., 2016. A gentle introduction to artificial neural networks. Annals of translational medicine, 4(19).

Zhao, X., Tan, Y., Shen, L., Zhang, G. and Wang, J., 2019. Case-based reasoning approach for supporting building green retrofit decisions. Building and Environment, 160, p.106210.

Zhu, J., Liapis, A., Risi, S., Bidarra, R. and Youngblood, G.M., 2018, August. Explainable AI for designers: A human-centered perspective on mixed-initiative co-creation. *In 2018 IEEE Conference on Computational Intelligence and Games* (CIG) (pp. 1-8). IEEE.

Zupan, J., 1994. Introduction to artificial neural network (ANN) methods: what they are and how to use them. *Acta Chimica Slovenica*, *41*, pp.327-327.

Zurada, J.M., 1992. Introduction to artificial neural systems (Vol. 8). St. Paul: West.

APPENDICES

APPENDIX 1

Publication from the Dissertation

1. Petersen, R., & **Daramola, O**. (2020). Towards Explainable Direct Marketing in the Telecom Industry Through Hybrid Machine Learning. Lecture Notes in Computer Science, 12254 LNCS, pp. 471-486, Springer. https://link.springer.com/chapter/10.1007/978-3-030-58817-5_35

Papers in Preparation

- Data-driven Direct Marketing Using an Explainable Machine Learning Approach (Planned Article in a DHET Accredited Journal)

- Evaluation of an explainable Machine Learning based Direct Marketing Tool for the Telecom Industry (Planned Article in a DHET Accredited Journal)

APPENDIX 2

SAMPLE OF CUSTOMER INPUT DATA

MANUFACTURER	MARKETINGNAME	DEVICESLOT_TYPE	SERVICES_SUPPORTEDSERVICES	WDSNWCHARACT_WIRELESSTECH	TOTAL_MONTHLY_SPEND
(null)	(null)	(null)	0	(null)	329.56
Vodafone	Vodafone Mobile WiFi R216	Wireless Router	0	4G	262.27
Apple	iPhone SE	Mobile Handset	1	4G	399.12
Samsung	Galaxy S6 Edge	Mobile Handset	1	4G	181.73
(null)	(null)	(null)	0	(null)	382.6
(null)	(null)	(null)	0	(null)	781.74
Samsung	Galaxy A7 2018	Mobile Handset	1	4G	378.26
Vodafone	R209-Z	Wireless Router	0	3G	129.57
Huawei	P10	Mobile Handset	1	4G	332.17
Vodafone	R218h Mobile WiFi	Mobile Hotspot	0	4G	286.09
(null)	(null)	(null)	0	(null)	738.25
Samsung	Galaxy Tab A 9.7	Tablet	1	4G	129.57
Qingdao Hisense Communications	Infinity E8	Mobile Handset	1	4G	343.48
Huawei	Y5 2019	Mobile Handset	1	4G	48.69
Samsung	Galaxy J5	Mobile Handset	1	4G	105.21
Samsung	Galaxy S9	Mobile Handset	1	4G	612.56
Samsung	Galaxy S9	Mobile Handset	1	4G	1226.78
Samsung	Galaxy A30s	Mobile Handset	1	4G	71.38
Vodafone	Smart Tab N8	Tablet	1	4G	155.65
Vodafone	R218h Mobile WiFi	Mobile Hotspot	0	4G	129.57
Samsung	Galaxy A30s	Mobile Handset	1	4G	370.42

APPENDIX 3

SAMPLE OF CUSTOMER DATA POST LABELLING

SPENDING_CATEGORY	DATA_USAGE_CATEGORY	PERIOD_OF_CONNECTION	SPEND_RATE_PER_DAY	DATA_USAGE_RATE_PER_DAY
HIGH	(null)	HIGH	9	(null)
HIGH	(null)	HIGH	32	(null)
HIGH	MEDIUM	HIGH	9	-22.7191466666666666666666666666666667
HIGH	(null)	HIGH	6	(null)
HIGH	(null)	LOW	7	(null)
HIGH	(null)	MEDIUM	13	(null)
LOW	(null)	MEDIUM	0	(null)
HIGH	(null)	MEDIUM	7	(null)
HIGH	(null)	MEDIUM	13	(null)
HIGH	(null)	MEDIUM	8	(null)
HIGH	(null)	HIGH	12	(null)
MEDIUM	(null)	HIGH	2	(null)
LOW	(null)	MEDIUM	0	(null)
HIGH	(null)	LOW	6	(null)
HIGH	(null)	MEDIUM	8	(null)
HIGH	(null)	HIGH	12	(null)
HIGH	(null)	LOW	9	(null)
HIGH	(null)	HIGH	6	(null)
HIGH	(null)	MEDIUM	36	(null)
HIGH	HIGH	HIGH	47	(null)

APPENDIX 4

SAMPLE OF CUSTOMER DATA POST LABELLING AND PREDICTION

DEVICE_CAPABILITY	STREAMING_CAPABILITY	STREAMING_USAGE_PER_DAY	PRODUCT
HIGH	STREAMING	(null)	YouTube Ticket
HIGH	STREAMING	(null)	1GB Data Bundle
(null)	NON-STREAMING	(null)	Whatsapp Ticket
MEDIUM	NON-STREAMING	(null)	Whatsapp Ticket
HIGH	STREAMING	(null)	Whatsapp Ticket
HIGH	STREAMING	(null)	1GB Data Bundle
HIGH	STREAMING	(null)	1GB Data Bundle
HIGH	NON-STREAMING	(null)	Whatsapp Ticket
(null)	NON-STREAMING	(null)	Whatsapp Ticket
HIGH	NON-STREAMING	(null)	Whatsapp Ticket
HIGH	STREAMING	(null)	YouTube Ticket
HIGH	STREAMING	(null)	1GB Data Bundle
HIGH	STREAMING	(null)	YouTube Ticket
HIGH	STREAMING	(null)	Whatsapp Ticket
HIGH	STREAMING	(null)	1GB Data Bundle
HIGH	STREAMING	(null)	YouTube Ticket
HIGH	NON-STREAMING	(null)	Whatsapp Ticket
HIGH	STREAMING	(null)	YouTube Ticket
HIGH	STREAMING	(null)	Whatsapp Ticket
(null)	NON-STREAMING	(null)	Whatsapp Ticket

APPENDIX 5

SAMPLE OF GQM QUESTIONS PRESENTED TO USERS

Name & Surname: Mr X		
Title: Customer Care Agent		
User Satisfaction (Clarity of Explanations)		
Goal	Purpose	Determine if explanation was clear and understandable.
	Issue	Establish quality of explanation.
	Object (process)	Output of explanation.
	Viewpoint	Users understanding of the explanation.
Question 1		How detailed are the explanations?
Metric (very low detail) 1 - 2 - 3 - 4 - 5 (very high detail)		5
Question 2		How clear are the explanations?
Metric (very low clarity) 1 - 2 - 3 - 4 - 5 (very high clarity)		5
Question 3		How much does the explanation contribute to your decision making?
Metric (very low contribution) 1 - 2 - 3 - 4 - 5 (very high contribution)		4
Question 4		How easy was it to generate the prediction and explanation?
Metric (very difficult) 1 - 2 - 3 - 4 - 5 (very easy)		3
Question 5		From your perspective, how reliable do you find the explanations and predictions?
Metric (not reliable at all) 1 - 2 - 3 - 4 - 5 (very reliable)		3
Mental Model (Understanding Decisions)		
Goal	Purpose	Determine how decision making is understood by the user.

	Issue	Understanding decisions produced.
	Object (process)	The processing of data to produce predictions.
	Viewpoint	User understanding of how a decision is produced.
Question 1		How easy is it to establish how a prediction was produced?
Metric (very difficult) 1 - 2 - 3 - 4 - 5 (very easy)		2
Question 2		How clear is it to you to establish how a decision was reached?
Metric (not clear at all) 1 - 2 - 3 - 4 - 5 (very clear)		3
Question 3		How much does knowing how a prediction was produced contribute to your decision making?
Metric (very low contribution) 1 - 2 - 3 - 4 - 5 (very high contribution)		5
Correctness (Correcting errors and training)		
Goal	Purpose	Verify the correctness of predictions and explanations.
	Issue	Establish accuracy of the predictions
	Object (process)	Executing predictions
	Viewpoint	Quality assurance
Question 1		How accurate are the predictions?
Metric (not accurate) 1 - 2 - 3 - 4 - 5 (very accurate)		4
Question 2		From your perspective, ow timeously was the explanation produced?
Metric (very slow) 1 - 2 - 3 - 4 - 5 (very speedily)		2
Question 3		How easy was it to familiarise yourself with the system?
Metric (very difficult) 1 - 2 - 3 - 4 - 5 (very easy)		4

Trust Assessment (Future use and trust)		
Goal	Purpose	Determine the frameworks contribution to trust.
	Issue	Establish if trust has increased.
	Object (process)	Framework as a whole.
	Viewpoint	Users perception of trust.
Question 1		How much do you trust the results produced?
Metric (low amount of trust) 1 - 2 - 3 - 4 - 5 (high amount of trust)		3
Question 2		How confident do you feel that a customer would purchase the predicted product should you call the customer?
Metric (not confident at all) 1 - 2 - 3 - 4 - 5 (very much confident)		4
Task Performance (Explanation improving decision making)		
Goal	Purpose	Establish contribution to decision making.
	Issue	Influence on decision making
	Object (process)	Framework as a whole
	Viewpoint	Senior/Managerial users
Question 1		How much does the explanation contribute to your decision making?
Metric (not at all) 1 - 2 - 3 - 4 - 5 (very much)		3
Question 2		How much has your understanding of customer behaviour improved due to the explanation?
Metric (not at all) 1 - 2 - 3 - 4 - 5 (very much)		3
Question 3		How much has your trust in machine learning solutions increased due to the explanations?
Metric (not at all) 1 - 2 - 3 - 4 - 5 (very much)		4