**Topic: Intelligent Marketing**

Student’s Name

Course Name

Date

**Table of Contents**

[Abstract 4](#_Toc102791237)

[1.0 Introduction 5](#_Toc102791238)

[Research Background 5](#_Toc102791239)

[Project Rationale & Significance 5](#_Toc102791240)

[Problem statement 5](#_Toc102791241)

[Research Questions 6](#_Toc102791242)

[Research Aim & Objectives 6](#_Toc102791243)

[Objectives 6](#_Toc102791244)

[Problem Solutions & Use cases 6](#_Toc102791245)

[2.0 Literature Review 9](#_Toc102791246)

[3.0 Research Methodology 14](#_Toc102791247)

[4.0 Data Analysis/Implementation 16](#_Toc102791248)

[About the dataset 16](#_Toc102791249)

[Exploratory data analysis 18](#_Toc102791250)

[Data cleaning and transformation 19](#_Toc102791251)

[Modeling to predict campaign response 24](#_Toc102791252)

[Supervised Machine Learning Models 25](#_Toc102791253)

[Unsupervised Machine learning models 29](#_Toc102791254)

[Reinforcement Learning 30](#_Toc102791255)

[ROC Curve 31](#_Toc102791256)

[Model Training 32](#_Toc102791257)

[Machine Learning Training 33](#_Toc102791258)

[Benefits of Machine Learning Model Training 33](#_Toc102791259)

[How to train a model 33](#_Toc102791260)

[Dataset Splitting 34](#_Toc102791261)

[Selection of machine learning algorithms to test 35](#_Toc102791262)

[Tuning the Hyperparameters 35](#_Toc102791263)

[Fitting and tunning machine learning models 35](#_Toc102791264)

[Choosing the Best Model 36](#_Toc102791265)

[Methodical approach to model training 36](#_Toc102791266)

[5.0 Key findings and discussions 38](#_Toc102791267)

[6.0 Conclusions and Recommendations 40](#_Toc102791268)

[6.1 Conclusions 40](#_Toc102791269)

[6.2 Recommendations 42](#_Toc102791270)

[References 43](#_Toc102791271)

**Figures**

[Figure 1: Steps in the data analysis 15](#_Toc102791272)

[Figure 2: The dataset information 19](#_Toc102791273)

[Figure 3: Data cleaning and feature transformation 20](#_Toc102791274)

[Figure 4: Outliers detection 23](#_Toc102791275)

[Figure 5: Outlier removal 23](#_Toc102791276)

[Figure 6: Machine Learning models 24](#_Toc102791277)

[Figure 7: Learning models for machine learning (Source link) 25](#_Toc102791278)

[Figure 8: Linear regression model 26](#_Toc102791279)

[Figure 9: A typical decision tree 26](#_Toc102791280)

[Figure 10: Neural network 27](#_Toc102791281)

[Figure 11: Support Vector Machine (SVM) 28](#_Toc102791282)

[Figure 12: Clustering patterns 29](#_Toc102791283)

[Figure 13: ROC Curve 31](#_Toc102791284)

**Tables**

[Table 1: Feature Description 18](#_Toc102791285)

[Table 2: Results ranking 30](#_Toc102791286)

**Equations**

[Equation 1: Bayes theorem 29](#_Toc102791287)

# Abstract

AI marketing tools provide businesses with valuable insight from collected datasets by analyzing the most effective advertisement placements or recommendations to increase customer engagement. Businesses personalize their products, services, and advertisements that specifically suit their customers based on their information. As customers' loyalty to a business increases based on their experiences, intelligent marketing allows organizations to build unique customer profiles that will further tailor personalize messages to customers at a prime stage of the consumer's lifecycle. Intelligent marketing has surpassed traditional marketing methods like direct emailing, media advertising, and physical retail marketing, which are no longer as effective as they were. Most importantly, Artificial Intelligence draws customers' patterns and identifies multiple drivers that will influence their decision to engage with a business brand. With the help of machine learning algorithms, Artificial Intelligence allows businesses to obtain and analyze data from different sources. There are different types of data that marketers can focus on in intelligent marketing, including financial, customer, and operational data. Businesses can improve their performance by analyzing their operations data, such as shipping logistics providing a better customer experience. The analysis of collected information from customers will help businesses understand their consumers more and get more visibility on potential targeted customers. Lastly, financial data allows marketers to measure business performance and operate more effectively by dropping marketing strategies that do not generate revenue. Above all, available relevant data determine marketer efforts in an organization. Collected data that is accurately analyzed guides businesses in their decision-making process, such as in marketing campaigns. Besides intelligence marketing being used to improve different marketing approaches, it allows businesses to navigate a technical world that keeps adopting automation. Big Data analytics and Business Intelligence can use this data for analysis and garnering insights that can be helpful in advertising and marketing. By doing this, organizations will improve their effectiveness to understand and reach consumers at different levels in the customer journey. The following research questions will help narrow down the research scope; What is Artificial intelligence and how does it impact marketing? What AI tools are used in Intelligent marketing? How can businesses employ Artificial intelligence in marketing? The objective was to investigate how the concepts and methods of artificial intelligence can be applied in the advertising and marketing process for business success. The purpose of this analysis was to study the effectiveness of previous marketing campaigns and propose better data-driven marketing solutions that will increase the response rates.

# 1.0 Introduction

## Research Background

Intelligent marketing entails incorporating different technologies to help optimize marketing strategies and improve customer experience through personalization (Rekha, Abdulla, & Asharaf, 2016). The rise of eCommerce and big data set the foundation for intelligent marketing. To be precise, as soon as businesses started adapting to eCommerce, marketers were forced to change their marketing approach, especially since they could reach a wider audience through the internet. The eCommerce concept allows businesses to operate on several digital platforms such as websites, social media networks, mobile applications, and so on to market their brand consistently. However, these platforms provide real-time communication and have lots of data. As Lies (2019) has mentioned, big data in intelligent marketing has proven to be indispensable as marketers can gather, analyze and use the massive amount of digital data to improve business operations. Additionally, enterprises with lots of traffic coming in on their online platforms require intelligent marketing tools to provide quick responses to customers and perform more tactical tools with less human intervention. Therefore, the intelligent marketing approach is popular among many businesses today as enterprises are rapidly adopting intelligent technology to make their business operations more efficient while maximizing profit and minimizing costs.

## Project Rationale & Significance

The study intends to discuss how Artificial intelligence has disrupted marketing and how businesses benefit from adopting this marketing approach. It will also highlight some of the advantages intelligent marketing has over traditional methods of marketing. The world is very advanced, and automation of activities is the norm today. Artificial Intelligence technology in marketing is used to automate marketing processes through data analysis where marketers can observe their target audience and economic trends that may impact their business. AI automated software is therefore vital in intelligent marketing as it streamlines and automates marketing tasks and workflows. Examples of AI automated marketing solutions include image recognition, chatbots, content recommendation engines, and dynamic pricing on eCommerce sites. AI technology primarily uses statistical models to analyze data and predicts future actions based on past behavior.

## Problem statement

Understanding the consumer journey is not a simple task. Using appropriate technology tools, customers use various forms to express their attitudes, needs, or values such as through comments, searches, blogs, likes, or even tweets. The supply of this consumer-curated data is seemingly endless and continues to grow. Big Data analytics and Business Intelligence can use this data for analysis and garnering insights that can be helpful in advertising and marketing. By doing this, organizations will improve their effectiveness to understand and reach consumers at different levels in the customer journey.

## Research Questions

The following research questions will help narrow down the research scope.

1. What is Artificial intelligence and how does it impact marketing?
2. What AI tools are used in Intelligent marketing?
3. How can businesses employ Artificial intelligence in marketing?

## Research Aim & Objectives

To investigate how the concepts and methods of artificial intelligence can be applied in the advertising and marketing process for business success.

## Objectives

* Understand what is artificial intelligence and intelligent marketing
* Analyze current marketing and advertising techniques in organizations
* Investigate how marketing can optimally utilize the AI technologies for maximizing customer satisfaction, market share, and profitability
* Understand how Artificial Intelligence affects advertising along the consumer journey

## Problem Solutions & Use cases

AI marketing tools provide businesses with valuable insight from collected datasets by analyzing the most effective advertisement placements or recommendations to increase customer engagement. Businesses personalize their products, services, and advertisements that specifically suit their customers based on their information. As customers' loyalty to a business increases based on their experiences, intelligent marketing allows organizations to build unique customer profiles that will further tailor personalize messages to customers at a prime stage of the consumer's lifecycle. Intelligent marketing has surpassed traditional marketing methods like direct emailing, media advertising, and physical retail marketing, which are no longer as effective as they were. Most importantly, Artificial Intelligence draws customers' patterns and identifies multiple drivers that will influence their decision to engage with a business brand (Dimitrieska, Stankovska, & Efremova, 2018)

With the help of machine learning algorithms, Artificial Intelligence allows businesses to obtain and analyze data from different sources (Siau & Yang, 2017). There are different types of data that marketers can focus on in intelligent marketing, including financial, customer, and operational data. Businesses can improve their performance by analyzing their operations data, such as shipping logistics providing a better customer experience. The analysis of collected information from customers will help businesses understand their consumers more and get more visibility on potential targeted customers. Lastly, financial data allows marketers to measure business performance and operate more effectively by dropping marketing strategies that do not generate revenue. Above all, available relevant data determine marketer efforts in an organization. Collected data that is accurately analyzed guides businesses in their decision-making process, such as in marketing campaigns. Besides intelligence marketing being used to improve different marketing approaches, it allows businesses to navigate a technical world that keeps adopting automation.

There are several use cases that highlight AI benefits in marketing. The benefits can be quantifiable, i.e., increased sales and revenue, or non-quantifiable, for example, customer satisfaction and risk reduction. A good example is the tailored recommendations on Netflix once a user has watched a few movies. Amazon has readily available data such as customers' browsing history, items they have purchased, and product review to provide their customers with the best suggestion. Machine-learning algorithms complement AI marketing efforts by analyzing customers' data to offer a hyper-personalized customer experience. AI is being used to create customer service chatbots that increase customer engagement. Customer engagement reflects customers' interests in a brand, and thus marketers need to get such insight. Another benefit of AI in marketing is depicted in social media platforms such as WhatsApp and Facebook, which have chatbots integrated into them. Kaczorowska-Spychalska (2019) states that chatbots create a more convenient way for consumers to directly contact their suppliers or service providers. It is a better and cheaper option compared to assigning customer service agents to handle different business social media accounts. Companies use virtual assistants such as Siri, Alexa, Google Assistant, and Cortana to collect information from users and later use the feedback to make advancements in their products and provide a better user experience.

Utilizing a methodical as well as repeatable model training procedure can be said to be of paramount importance for any business entity or company strategizing on building successful machine learning or artificial intelligence models at scale. The central point of having all these are strategizing on the available resources, tools as well as libraries, and proper documentation in a single enterprise platform that will amplify team collaboration rather than hindering the process. In conclusion, the above section has had relative discussions on various machine learning models as well as algorithms. However, there may be rising questions relative to any beginner that may be regarding which type of model should they choose? The answer to this question is; that it depends on the business requirement or project requirements. However, it also depends on the associated features, the available dataset volume, the number of features, and complexity level among other things. In addition, in real-case practice, it is always recommended that one needs to start with the simplest machine learning model that can be utilized for a given problem and then gradually foster the complexity level as well as carry out testing on the accuracy with the aid of parameter tuning as well as cross-validation.

# 2.0 Literature Review

**Evolution of Intelligent Marketing**

The intelligent marketing strategic approach started with the availability of the worldwide web in 1991 (Khanzode & Sarode, 2016). Millions of internet users kept increasing as they wanted to be part of the technologically advanced society. Many businesses preferred investing in Customer Relationship Management (CRMs) tools to create and diversify their current interaction with current and potential customers. Marketers primarily focused on search engines to get more online information about their targeted consumers. Unfortunately, businesses were starting to get overwhelmed with heaps of customer data and had a challenge in analyzing and using the collected data for their business operations. The 90s decade saw a rapid change in the way businesses and customers used the internet to make their buying and selling activities easier. Society’s overdependence on mobile phones drove the online marketing strategy (Yadav, Joshi, & Rahman, 2015). Individuals’ usage of mobile phones involves different activities, which include communication, checking mails, photo snapping and video recording, payment of bills, and so on. Businesses took advantage of the situation and included ads in mobile applications to target a broader audience. Today, Intelligent marketing is at its peak, and businesses make sure they are updated with the current marketing practices and the advancements that are to come. The latest advancement in marketing includes artificial intelligence marketing, which adopts insight and data-driven marketing, i.e., content marketing and customer service through social media platforms. The vastness and increased pace of intelligent marketing can be overwhelming for businesses when there is less research work and evidence that proves the benefits of employing Artificial Intelligence in digital marketing. The literature review aims to give insight to the reader on marketing intelligence evolution.

There is no specific founding father for digital marketing as different technology innovations improved the marketing process. However, Philip Kotler stands out among many innovators responsible for the improved online marketing methods. Kotler is a well-known American professor who invested his time and resources in establishing better marketing strategies. He argued that marketing played a huge role in maintaining a successful business and also pointed out that the demand for goods and services was also influenced by advertising and promotions through direct mail and any other distributing channels (Kotler, Kartajaya, & Setiawan, 2019). The progression of modern marketing can be categorized into different eras. The sale orientation era describes business challenges in selling their mass-made products. Marketers saw the need to enforce branding in their marketing strategy as supplies had surpassed demand, and companies were now competing for customers. The heavily saturated market led to the marketing orientation era, where companies prioritized hiring marketing professionals to increase their sales. In this era, marketers had more say in the company's services, the product produced, distribution channels, and pricing decisions. Marketers gathered understanding consumers' needs was essential in marketing. Therefore, it founded the relationship marketing era where companies focus on building long-term relationships with their customers. As Hunt (2018) mentions, generic hard sales and marketing campaigns were outdated as companies did not gain customers' trust. For customers to be loyal to a particular business, they had to be comfortable and be assured that their suppliers prioritized their interests. The social marketing era became popular as businesses could easily interact with their customers in real-time. In this era, marketing automation was adopted because it was challenging to maintain a one-to-one relationship with clients. Social media applications such as Twitter, Instagram, LinkedIn, and Facebook provide marketers with necessary data that will help increase operational efficiencies. Marketing automation became a critical success factor for businesses because it increased their revenues and minimized costs by reducing manual processes.

**Automation & Its Impact on Intelligent Marketing**

Marketing automation entails using software tools to execute, manage, and automate marketing tasks and processes. Its main purpose is to replace the tiresome manual and repetitive marketing processes with software programs developed to improve marketing. Intelligent marketing coupled with traditional media marketing channels such as radio stations, billboards, and television ads enable businesses to attain a broader customer reach. As marketing strategies keep evolving, the primary purpose remains to increase customer engagement and improve customer experience (Gerrikagoitia, Castander, Rebón, & Alzua-Sorzabal, 2015). Artificial Intelligence equips marketers with more sophisticated tools to personalize the customer experience. These advancements in marketing allow businesses to meet customers’ expectations in a much simpler, faster, and more efficient way than the traditional marketing approaches. AI incorporates Machine Learning in marketing which facilitates deep learning-based recommendations that make marketers' work easier. The deployment of cloud infrastructure has also made AI more affordable and scalable. Most importantly, businesses are adopting AI because it uses algorithms that provide useful insights that help in decision-making while at the same time giving companies some level of control. The continuous change in consumer behavior keeps confusing marketers as they cannot settle on the right marketing strategy. Fortunately, these challenges can be solved by marketing intelligence which allows marketers to monitor the performances of different mediums of marketing that will give them precise analytics for a better understanding of customers’ behavior. A good example is the use of cookies on websites. The technique tracks common browsing habits and product usage patterns of website users to tailor personalized promotions. The continuous increasing influence that technology has on consumer behavior makes intelligent marketing more appealing. Paschen, Kietzmann, and Kietzmann (2019) discussion confirms that data collected from customers in marketing undoubtedly provides the foundation of business plans. Additionally, data analysis technology keeps evolving, and thus companies marketing strategies need to be revised to adopt new digital resources effectively.

**Reasons to Adopt Intelligent Marketing**

Throughout the history of marketing, the focus on marketing has been to set up businesses in the direction that allows them to conduct more in-depth research and use the findings to stay ahead of the market, reduce risk in investments, and make strategic decision-making. Intelligent marketing implements this by managing digital consumer data and providing a comprehensive outlook on consumer behavior that deduces conclusions for marketing decisions (Kumar, Rajan, Venkatesan, & Lecinski, 2019). Intelligent marketing can be classified into four categories. First, it is the competitive intelligence where marketers observe consumers' and competitors' behavior and establish strategic actions based on their findings. Businesses' product and service decisions are in response to market trends. Second, strategic marketing entails using internal and external data that will positively disrupt business operations. Third, there is predictive intelligence which focuses on understanding the market by monitoring subjects that are relevant to the organization. Marketers use the findings to anticipate future threats and opportunities. Lastly, data intelligence objectives are to organize, analyze and integrate it with the company’s software system.

Artificial Intelligence has multiple subfields that use different techniques to achieve marketing automation. Marketing is rich in data, and Machine Learning (ML) has exceptional algorithms that can handle complex datasets and provide accurate insights into marketing campaigns. Shah, Engineer, Bhagat, Chauhan, and Shah (2020) argue that ML can transform existing data on a product or service into a detailed list of insights that describe customer behavior and expectations. ML algorithms are more reliable in mining unstructured data for sentiment analysis. However, Machine Learning is more suitable for industries that work with large and complex datasets, including many human demographic variables. On the other hand, deep learning helps create content, develop chatbots, real-time bidding on ad networks, speech recognition, and natural language processing. Marketers can use Natural Language Processing to analyze social post reviews and user-generated content that is related to their company. AI helps marketers have operational workflows, easily create marketing content, perform predictive analysis, and optimize marketing activities (De Bruyn, Viswanathan, Beh, Brock, & von Wangenheim, 2020).

**What to Consider before Adopting Intelligent Marketing**

Intelligent marketing is vital in obtaining marketers' required information to construct marketing strategies. When businesses know the market's competitive state, they can confidently enter a new or existing market with confidence. They can also use the insights attained from intelligent marketing to make informed decisions that bring consistency to their business. As Dimitrieska et al. (2018) have discussed, technologies keep improving every day, and so are the marketing intelligence processes. Therefore, organizations need to understand their dependency on digital technologies and how they are to improve their business operation before adopting marketing intelligence. Although different organizations have different marketing targets, marketers can use the following common steps to get started with Artificial intelligence marketing. The marketing team can identify suitable metrics and KPIs that will be used to measure the effectiveness of an AI-based marketing strategy. As data is crucial in planning future events, marketers should create a data privacy strategy for all AI operations. For instance, AI tools can be programmed to anonymize personal data and then store it in default ways that implement the access controls, encryption, and other privacy protections required to regulate frameworks. It is also important that the organization has data sources, for example, CRM systems and website logs. Having the necessary skills is key in adopting intelligent marketing, especially when deciding on suitable AI tools (Iazzi, Trio, Pandurino, & Caione, 2015). Artificial intelligence computing systems can work toward solving problems without needing written codes. The system is initially programmed to learn from human interactions through a predetermined set of rules. Before automation and online presence, businesses had limited interactive marketing where customers found it challenging to share their experiences with a certain brand. Today, advanced AI algorithms have shown great results in enabling quality consumer experience and easing up the customer journey. In addition, brands can evaluate their AI marketing-based tools according to their distinct business goals and objectives.

**Role of AI in Intelligence Marketing**

AI at its core perform marketing tasks that require human intelligence more efficiently and thus saves a lot of time and costs in doing business. It emulates the capacity of human power and surpasses its ability by remaining accurate. Jain and Aggarwal (2020) state that the decision made based on the insights provided are more reliable because intelligent marketing solely relies on data and defined patterns. AI allows marketers to recognize and categorize customer segments by their behavioral patterns and then optimize performance by monitoring how well different content performs against individual segments. Furthermore, intelligent marketing using AI helps curb human errors, especially in maintaining data security. The majority of the businesses hesitate to use customers' data because they are worried their employees may misuse customers' data. AI provides companies with better data security options and prevents information from being harmed by cyber-attacks (Safdar, Banja, & Meltzer, 2020). Another successful use case for AI marketing is image recognition. Using computer vision technology, programmed AI systems can understand visual information. Marketers can scan millions of images online and get to know how their products and their competitors are performing in the market. Popular eCommerce websites such as IKEA and Home Depot allow clients to try their products using Augmented Reality before confirming the purchase (Sung, 2021). AI enables this by creating realistic-looking composite images, generally in real-time, as the user is looking through the camera on their phone. Being ranked at the top of search engines is one of the most critical steps in marketing. Intelligent marketing uses Machine Learning algorithms to understand the intent behind keyword usage and the content of internet users’ searches. However, AI cannot completely replace the human role in marketing, i.e., creating creative content such as video and image ads to attract customers. Customers may need a human-to-human connection to try a company’s product or service. Therefore, marketers should know when to apply intelligent marketing and human skills. For instance, AI chatbots used to recommend and enhance customer service are limited to their programmed functionalities. The study aims to expound on the impacts of intelligent marketing and the technologies used to automate marketing activities.

# 3.0 Research Methodology

Research methodology is the structured process of how to conduct research. It majorly entails the research design, how to gather data, and how to analyze the data. The major aim of the report was to investigate how artificial intelligence and machine learning can be applied in marketing and advertising for business success.

**About the data**

The data used for this analysis contains marketing information and is sourced from the Kaggle [link](https://www.kaggle.com/seananguyen/marketing-campaign-analysis-python/data). It contains information about customers, their demographic information, and details of how they reacted to various marketing campaigns.

**Methodology**

This research applies both quantitative and qualitative research methodologies. The quantitative approach analyses the information about numbers such as the number of items that the customers bought. The qualitative approach will entail applying analysis tools and techniques to explain why some phenomena were observed. For instance, to understand why people with kids or high income like to purchase items of a given category. The motive behind using the quantitative approach is to produce knowledge that can be generalized about the causes of various marketing phenomena. On the other hand, the qualitative approach will aim at producing contextual real-world knowledge about the behaviors of various customers of different demographic categories and how they react to marketing campaigns targeted to them.

**Steps in the data analysis**



Figure 1: Steps in the data analysis

The image above depicts the steps that the data analysis process will entail. The first step is the definition of the need to perform the analysis. In this case, it is to unravel how machine learning can be applied in marketing and advertising for business success after understanding customer purchase patterns from recorded data. The second step is sourcing the data from the relevant sources. The data is then cleaned to make it ready for analysis using proper tools. The analysis produces results that the analyst can interpret to answer the research questions and aims.

# 4.0 Data Analysis/Implementation

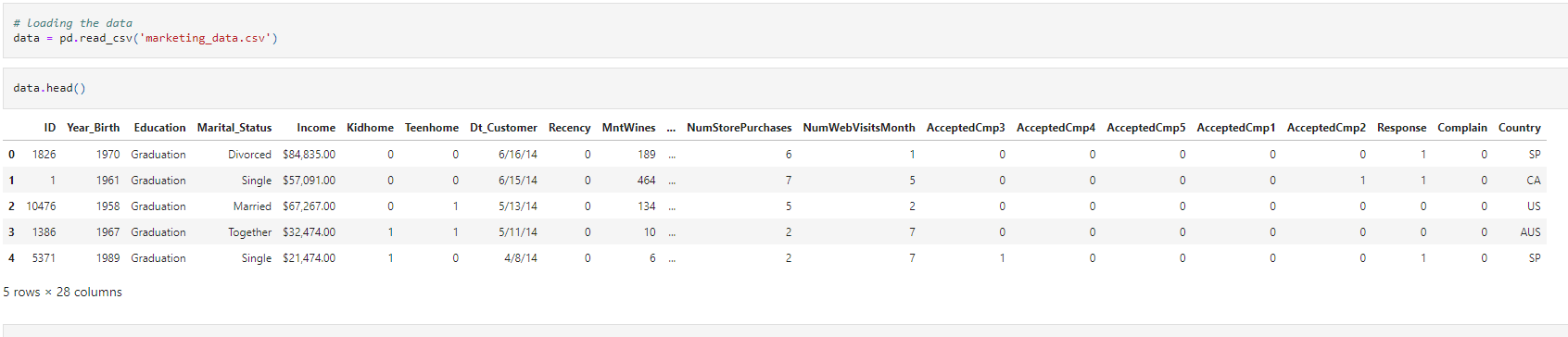
This chapter illustrates the data analysis process, which entails cleaning, processing, and transforming raw data before training on a model. In the process, actionable insights are generated, which are helpful to businesses in making informed decisions. Data analysis will help in reducing inherent risks that affect the decision-making process and it provides useful insights presented using graphs, charts, images, and tables. The reason behind analyzing the dataset used in this study is to enable firms to develop better advertisement strategies and optimize their marketing campaigns by leveraging machine learning tools and techniques. Also, the analysis will help an organization target their customers, reduce operational costs, improve revenues, and come up with better problem-solving methods. The major activities in this stage are cleaning the data, analyzing it using relevant tools, interpreting the results, visualizing the insights, and building a model to predict the acceptability of marketing campaigns.

## About the dataset

The dataset used in this study was sourced Kaggle [link](https://www.kaggle.com/code/seananguyen/marketing-campaign-analysis-python/data). It contains information about customers for an e-commerce website that recorded information about customers such as their income, age, number of products they purchase, and how they respond to marketing campaigns.

**Loading the dataset**

Pandas library has been used to lead the dataset from a CSV file to a data frame that is easy to perform statistical operations on.



The data has 2240 records or rows, and 28 columns or features. Applying machine learning techniques to the dataset can lead to the creation of a response model that can offer a significant boost in the efficiency of marketing campaigns. Analyzing such data will help an organization identify the products that customers buy most and how they respond to marketing campaigns. This can help an organization to focus on products that sell most, reduce expenses on marketing products to unresponsive populations, and better the channels that customers use to purchase their items. The objective of this study is to create a model that can predict whether customers will respond to offers or marketing campaigns for products before they are launched. Information derived will help the management to make appropriate planning and dispense marketing to the right group. This reduces marketing expenses and maximizes the returns from the effective marketing campaigns launched by the company. The table below describes the different features of the dataset that will be explored for the study.

|  |  |
| --- | --- |
| **Feature** | **Description** |
| ID | The customer's unique identifier |
| Year\_Birth | The customer's year of birth |
| Education | The customer's level of education level |
| Marital\_Status | The customer's marital status |
| Income | The annual household income of the customer |
| Kidhome | The number of small kids in the customer's household |
| Teen home | The number of teenagers in the customer's household |
| Dt\_Customer | The date when the customer enrolled in the company |
| Recency | The number of days since the customer's last purchase |
| MntWines | The amount the customer spent on wines in the last 2 years |
| MntFruits | The amount the customer spent on fruits in last 2 years |
| MntMeatProducts | The amount the customer spent on meat products in the last 2 years |
| MntFishProducts | The amount the customer spent on fish products in the last 2 years |
| MntSweetProduct | The amount the customer spent on sweet products in the last 2 years |
| MntGoldProds | The amount the customer spent on gold products in the last 2 years |
| NumDealsPurchases | The number of purchases made with discount |
| NumWebPurchases | The number of purchases made through the company's website |
| NumCatalogPurchases | The number of purchases made using catalog |
| NumStorePurchases | The number of purchases made directly in stores |
| NumWebVisitsMonth | The number of visits to the company's website last month |
| AcceptedCmp1 | 1 if the customer accepted the offer in the 1st campaign, 0 otherwise |
| AcceptedCmp2 | 1 if the customer accepted the offer in the 2nd campaign, 0 otherwise |
| AcceptedCmp3 | 1 if the customer accepted the offer in the 3rd campaign, 0 otherwise |
| AcceptedCmp4 | 1 if the customer accepted the offer in the 4th campaign, 0 otherwise |
| AcceptedCmp5 | 1 if the customer accepted the offer in the 5th campaign, 0 otherwise |
| Response(target) | 1 if the customer accepted the offer in the last campaign, 0 otherwise |
| Complain | 1 if the customer complained in the last 2 years |
| Country | The country where the customer is located |

Table : Feature Description

## Exploratory data analysis

Exploratory Data Analysis (EDA) entails analyzing the dataset to summarise its major characteristics of the same. It is majorly conducted via visualization methods. It has to be conducted before the modeling part. The technique allows the analyst to identify errors, and understand patterns or interesting relations in the dataset. It is in this stage that outliers are addressed as they might affect the final results, especially when creating the predictive model. EDA has been performed in this study to ensure that the results are reproducible and applicable in any business that has similar goals and objectives. It is in this stage that stakeholders get the answers to the questions they have raised. When the process is completed, the final features selected are then used to build the final model using sophisticated techniques and tools such as machine learning. The activities performed in this stage were data cleaning and variable transformation, checking outliers, checking nulls, feature engineering, and visualizing trends in the dataset.

### Data cleaning and transformation

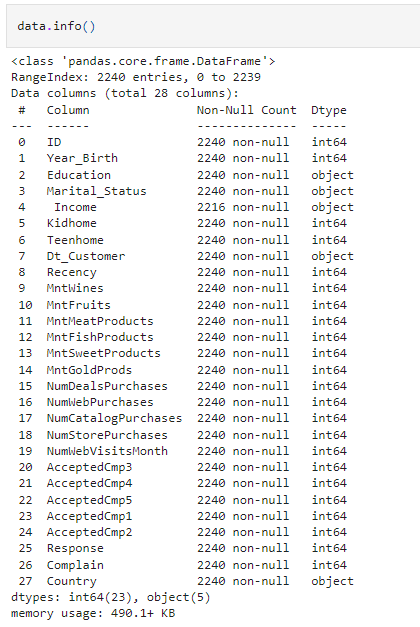


Figure 2: The dataset information

The figure above shows the information about the dataset. The Dtype column specifies the different data types of the features of the dataset. The income column contains records in string or object format. This implies that it has to be changed to float or double format. The code snippet below shows how the characters in the income column were cleaned and the income feature was transformed to a float data type.

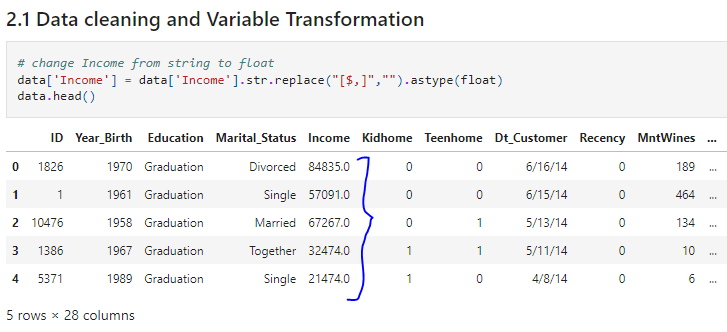
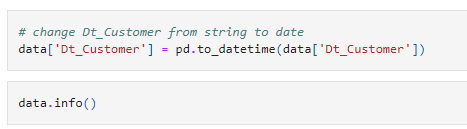


Figure 3: Data cleaning and feature transformation

The Dt\_Customer column that shows the date when the customer enrolled in the company was also changed from string to a date format.



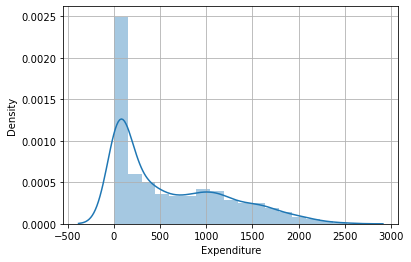
The marital status and education columns were transformed as shown below.



A new column, expenditure, was added to contain all the amount that each customer has spend in various products.

*# new expenditure column*

data['Expenditure'] **=** data['MntWines'] **+** data['MntSweetProducts'] **+** data['MntMeatProducts'] **+** data['MntGoldProds'] **+** data['MntFruits'] **+** data['MntFishProducts']



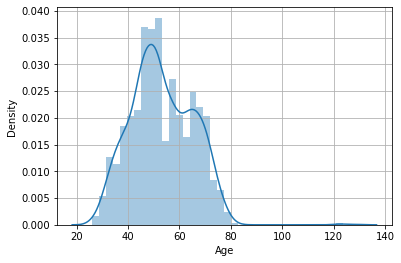
Visualizing the new expenditure column revealed that most expenditure ranges between 0 and $ 2500.

A new column, age, was also added to reveal the ages of the customers. It was derived by subtracting the customer’s year of birth from the current year.

*# new age column*

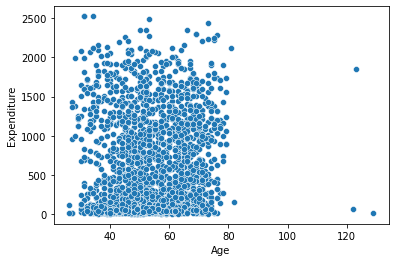
data['Age'] **=** datetime**.**datetime**.**now()**.**year **-** data['Year\_Birth']

data**.**head()



Visualizing the new “Age” column reveals that the majority of the customers are aged between 20 and 80 years.

Below is a scatter graph showing the relationship between expenditure and age.



**Handling missing values**

Often, real-world data will have missing values. They are caused by data corruption or failure to record data. These missing values have to be addressed as they can distort findings. Machine learning models do not support missing values. The common ways of dealing with missing values are deleting or dropping the rows that contain the missing values and replacing the missing values with the right values such as median, mean, or mode. The percentage of the missing values for this case study was about 1.07 %. Being a small percentage, the rows with the missing values were just dropped.

**Dealing with outliers**

Outliers are values or observations that lie at an abnormal distance compared to the majority of values in the data sample. When used in the analysis, they bring out different results compared to when they are eradicated. Outliers, in this case, were detected via appropriate visualizations using the matplotlib python library. Histograms and boxplots were used to detect outliers as shown in the figures below.

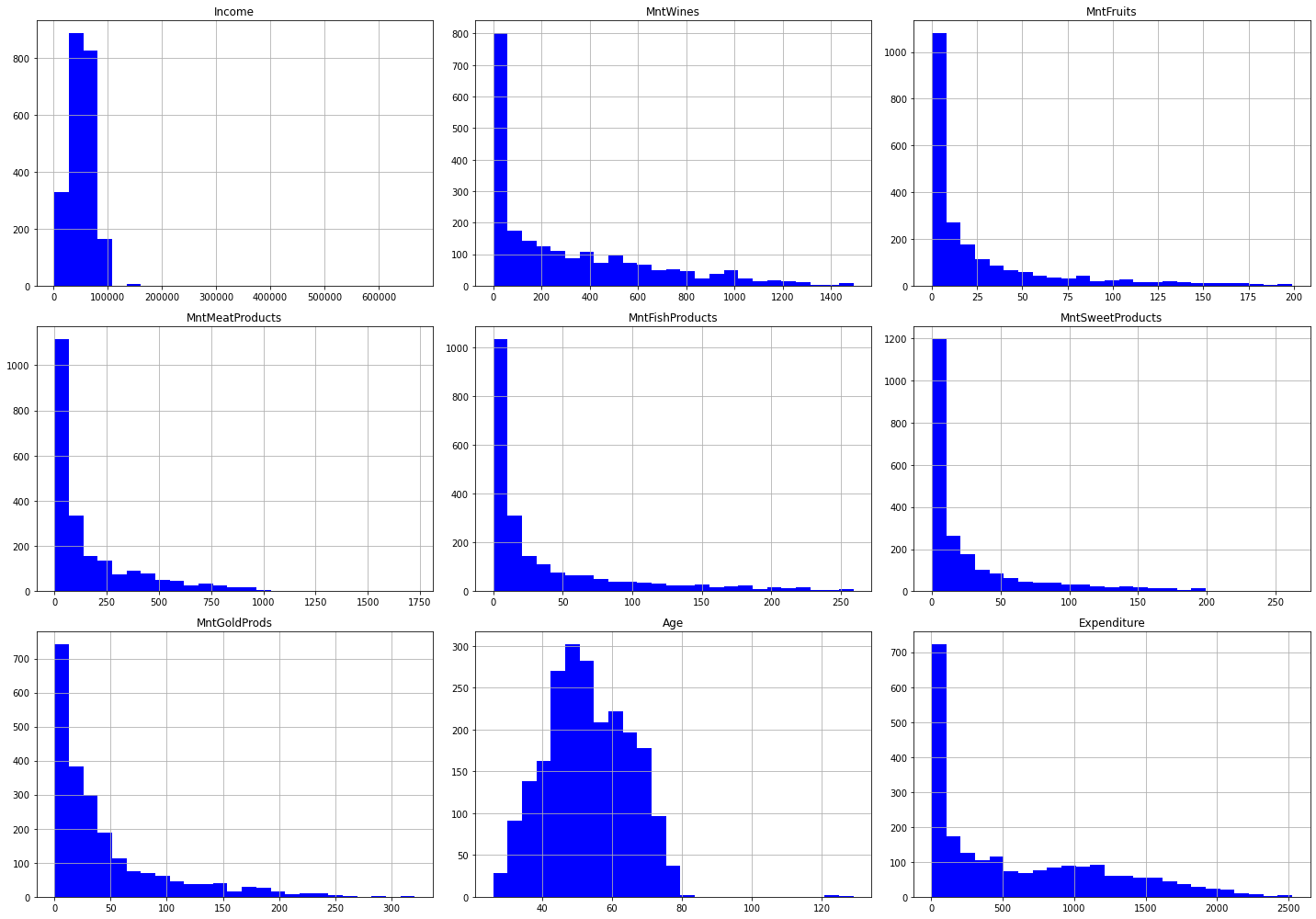


Figure : Outliers detection

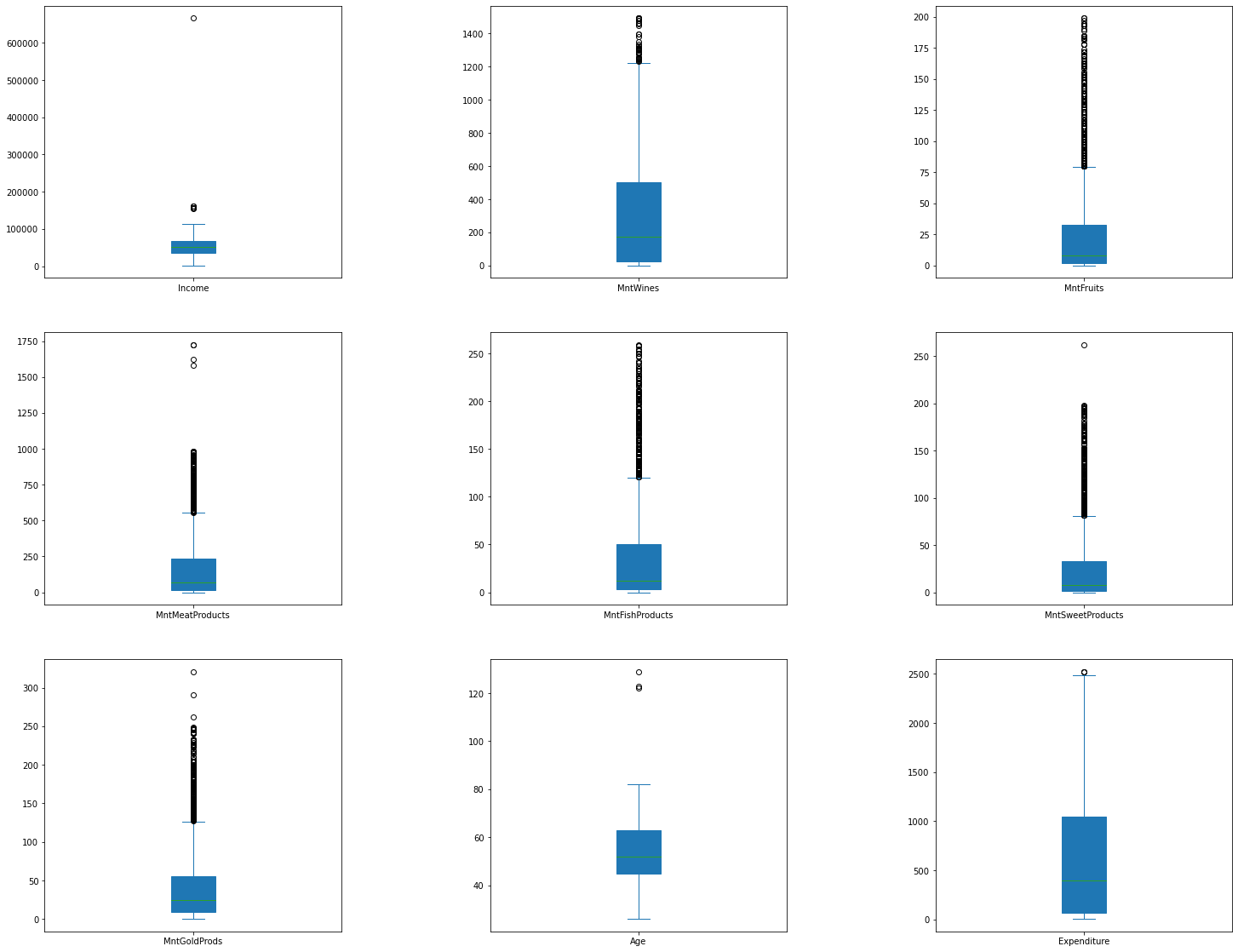


Figure : Outlier removal

The income and age columns have few outliers. To deal with the outliers, the rows where age is above 120 were dropped, and the rows where income was 600,000 were dropped too.

## Modeling to predict campaign response

In this case study’s dataset, the target variable is “Response”, whose value is 1 if the customer accepted the offer in the last campaign, or 0 otherwise. This implies that the target variable is in a binary form, calling for a classification algorithm to be used to build a machine learning model that will predict the response value based on the data fed. A machine learning model is a mathematical representation of the output from a trained dataset. A machine learning model is represented as a mathematical function. It takes in input data and makes predictions on that data, and outputs a response. Machine learning entails studying different algorithms to build a model which can improve automatically through learning and experience. The learning algorithm is trained on past data to discover patterns, and then creates a machine learning model that captures the same insights when fed with similar data. There are various machine learning models as illustrated in the figure below.

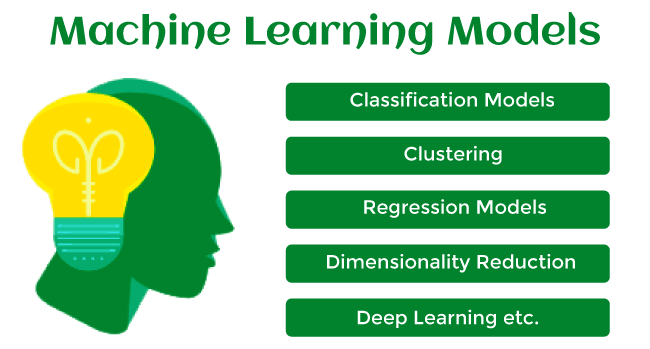


Figure 6: Machine Learning models

Machine learning falls into three major categories – supervised, unsupervised, and reinforcement learning. In supervised learning, the machine learning algorithm uses labeled input data for training and testing, and the output is known. For instance, in this case, study, the dataset is in labeled columns, and the output is known – response (0 or 1). In many cases, the machine learning algorithm in supervised learning does not make 100% accuracy in prediction and hence has an error. More data for training is needed to improve the model’s performance. On the other hand, in unsupervised learning, the machine learning algorithms used do not learn on labeled data, but rather learn from unlabelled datasets to discover underlying insights. The model learns hidden patterns in unlabelled data without any supervision. For reinforcement learning, the machine learning algorithm learns from a given set of states leading to a goal state. It takes feedback for each action or state as it interacts with the environment. There is a reward for each action – positive for good action, and negative for a bad action. To improve the performance of the model, the algorithm or agent aims at maximizing the positive rewards. These three categories of machine learning are illustrated in the figure below.

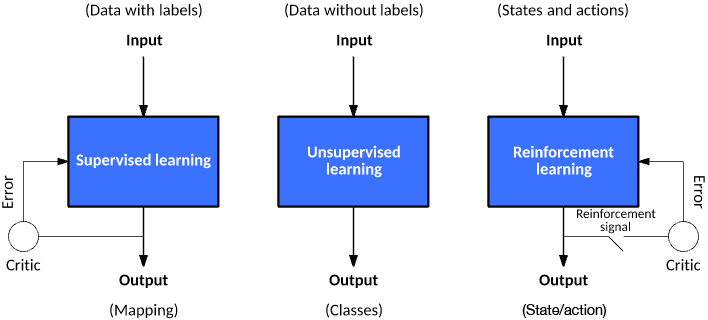


Figure 7: Learning models for machine learning (Source [link](https://developer.ibm.com/articles/cc-models-machine-learning/))

Supervised machine learning is further divided into two categories i.e. classification and regression. Unsupervised learning on the other hand is divided into three major categories – clustering, association rule, and dimensionality reduction.

### Supervised Machine Learning Models

1. **Regression**

In regression machine learning problems, the output is a continuous variable. For instance, predicting stock exchange rates, weather elements such as rainfall or temperature, and predicting profits for a company, among others. Some of the commonly used regression models include:

1. ***Linear Regression***

This is the simplest regression model that uses one or more input variables to predict one continuous output variable. It is represented as a linear equation that combines a set of input variables labeled as **x** and the predicted output labeled as **y** outlined as **y** = b**x**+ c.

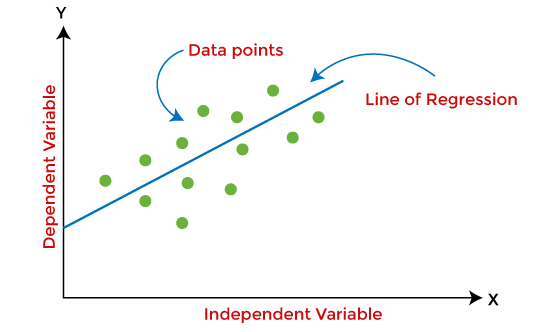


Figure 8: Linear regression model

The linear regression model finds the best fit line that linearly separates the data points. In multiple linear regression, the model tries to find the plane of the best fit, while in polynomial regression, the model tries to find the best fit curve.

1. ***Decision Tree***

Decision trees are used for both regression and classification problems. A decision tree uses a tree-like structure to make decisions on how to separate the data points based on their consequences and outcomes. A test on an attribute is represented by an internal node, and the outcome is represented by a branch. The deeper the tree or the more nodes a tree has, the more accurate the modeling results will be. Decision trees are intuitive and easy to implement, but they are prone to overfitting and less accuracy. They are used in decision analysis, operations research, and strategic planning.

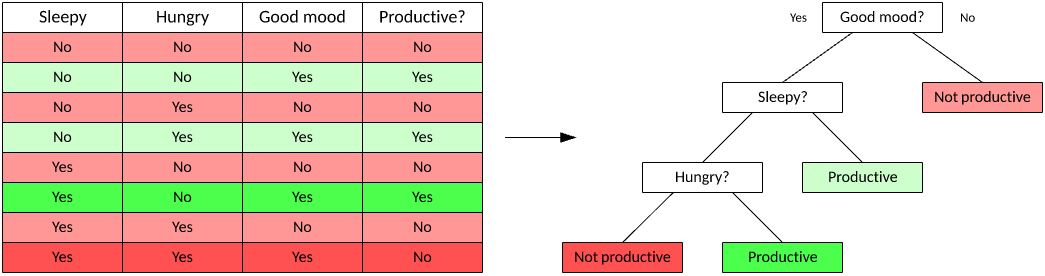


Figure 9: A typical decision tree

1. **Random Forest**

This is an ensemble learning method that consists of multiple decision trees that have been combined to improve accuracy. Each decision tree in the random forest produces an outcome, and the tree with the best outcome is picked (for a classification task), or the average of the outcomes from all the trees is determined and used (for a regression task). The Random Forest model is used for both regression and classification problems.

1. **Neural Networks**

Neural networks, commonly referred to as Artificial Neural Networks (ANNs) are a subset of machine learning. They are made of artificial neurons which imitate the brain structure of human beings. Each of these artificial neurons connects with others in the neural network. The resultant is a dense and complicated cognitive structure with an input layer, hidden layers, and an output layer as illustrated in the image below.

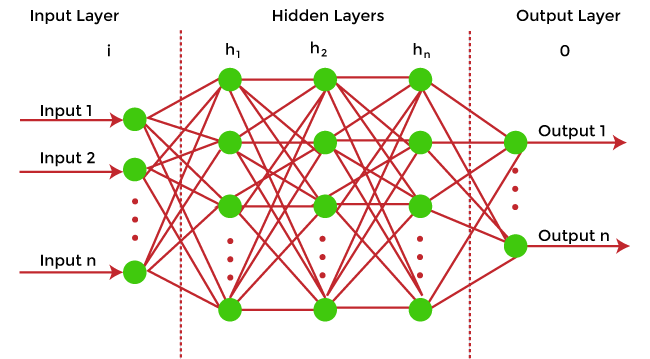


Figure 10: Neural network

In a Neural Network, the connected neurons transmit data from one layer to another. The networks rely on training data to train, learn and improve their accuracy. When a neural network is trained perfectly, it becomes a powerful AI tool for clustering data. The Google Search algorithm is one of the best-known neural networks.

1. **Classification**

The second type of supervised learning is classification. Classification generates conclusions from observed values in a categorical form. For instance, classification can classify whether an email is spam or not, whether a customer will purchase an item or not, etc. this report is about classification as it tries to predict whether a customer will accept a marketing campaign or not. Classification falls into 2 categories – binary and multi-class classification. Since the problem in this report has only two possible outcomes – 0 or 1 for the marketing campaign response, then it falls under the binary classification category where the problem has only two possible classes. In multi-class classification, the problem has more than two possible classes, for instance positive, neutral, and negative. Below is an explanation of some of the most popular classification algorithms.

1. **Logistic Regression**

Logistic Regression is similar to linear regression, but it is used to predict categorical variables in machine learning. However, it provides probabilistic values between 0 and 1, rather than giving the exact values.

1. **Support Vector Machine (SVM)**

The support vector machine (SVM) works by finding the best decision boundaries for an n-dimensional space, where n is the number of features. It segregates data points into classes and the line separating them is known as the hyperplane. The algorithm will select the extreme vectors or data points usually referred to as the support vectors when finding the hyperplane. They are the closest points to the hyperplane from both classes. The SVM algorithm tries to maximize the margin between the support vectors and the hyperplane for better results.

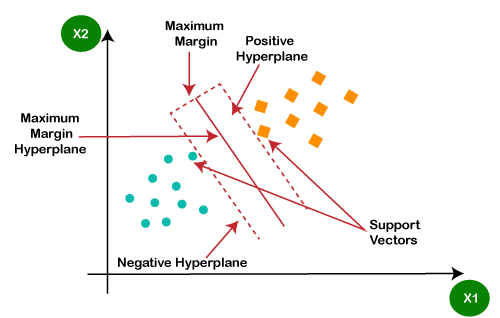
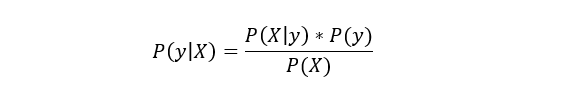


Figure 11: Support Vector Machine (SVM)

SVM algorithm works best where there is a clear separation margin and is very effective in high dimensional spaces. It is memory efficient as it uses a subset of the training points (support vectors) in the decision function.

1. **Naïve Bayes**

Naïve Bayes is based on the Bayes theorem and is a common classification algorithm in machine learning. It follows the naïve or independent assumption between the dataset features.



Equation 1: Bayes theorem

The algorithm assumes that the values of the features in the dataset are independent of each other. The algorithm needs less training compared to other models and is easy and fast to make predictions. It also performs well for categorical data compared to numerical values. Under the scikit learn python library, the algorithm is implemented in three forms – Gaussian, Multinomial, and Bernoulli. Gaussian is used for classification and it assumes that the featured fit in a normal distribution. Multinomial is mainly used for discrete counts. Bernoulli is used if the features are vectorized into binary (0 and 1).

### Unsupervised Machine learning models

These methods work contrary to the supervised machine learning methods. The model learns from a training dataset that is not labeled by itself without any supervision. The model then predicts the outcome based on the insights it has learned from the unlabelled data. Majorly, unsupervised learning models are used for clustering, dimensionality reduction, and association rule learning. Clustering entails clustering data into groups based on their inherent similarities or differences. Data points or objects with similar characteristics are grouped. They have different characteristics from the objects in the other groups. These algorithms are deployed in fields like market segmentation, statistical data analysis, and image segmentation, among others. K-means clustering algorithm is the most popular algorithm in this category.

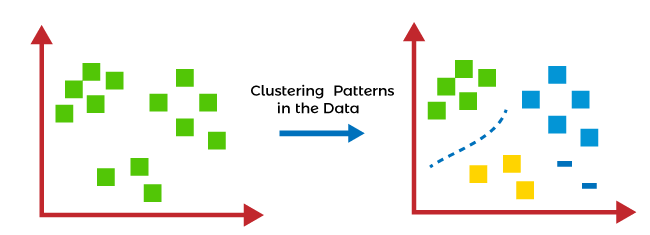


Figure 12: Clustering patterns

Dimensionality Reduction is a technique used to reduce dimensionality in a dataset. Dimensionality is the number of variables or features in a dataset. Even though many features are essential to have a better-performing model, they may as well cause performance shortcomings such as overfitting. This prompts dimensionality techniques to be used to reduce the number of these features. This can be performed using techniques like Principal Component Analysis (PCA) and Singular Value Decomposition (SVD).

Association Rule learning is a technique that discovers interesting relations among dataset features or variables. This is mainly used in finding item dependencies and mapping the variables accordingly so that maximum gain is obtained from the same. For instance, association rules can be used to find items that are commonly purchased with others e.g. sugar and bread or milk. The algorithms under this category are majorly used in market basket analysis, continuous production, and web usage mining, among other applications. Apriori and FP-growth algorithms are the commonly used algorithms under this type of unsupervised learning technique.

## Reinforcement Learning

This type of machine learning is a feedback-based learning model where the algorithm learns from actions that translate to different sets of states leading to a goal state. The model takes feedback signals after every state change based on action completed while interacting with the environment. When the feedback is positive, then the action is good and is hence considered a reward. When the action is not good, then it is negative feedback. The goal of the agent is to maximize the positive rewards and hence improve the model’s performance. The model in this case behaves like a human being. Human beings learn stuff from experiences with interaction with the environment and work on the feedback they get to improve on their goals. The commonly used algorithms in this category are State Action Reward State Action (SARSA), Q-learning, and Deep Q Network (DQN).

After analyzing the produced results, the researcher represented them in a table as shown below;

|  |  |  |
| --- | --- | --- |
| **Rank** | **Model** | **Accuracy (%)** |
|  | Gradient Boosting | 88.25 |
|  | XGBoost | 87.52 |
|  | Decision Tree | 86.44 |
|  | Logistic Regression | 86.8 |
|  | Neural Network | 86.08 |

Table : Results ranking

From the above results, the Gradient Boosting model resulted to be the better model that had a high accuracy level of 88.25%.

### ROC Curve

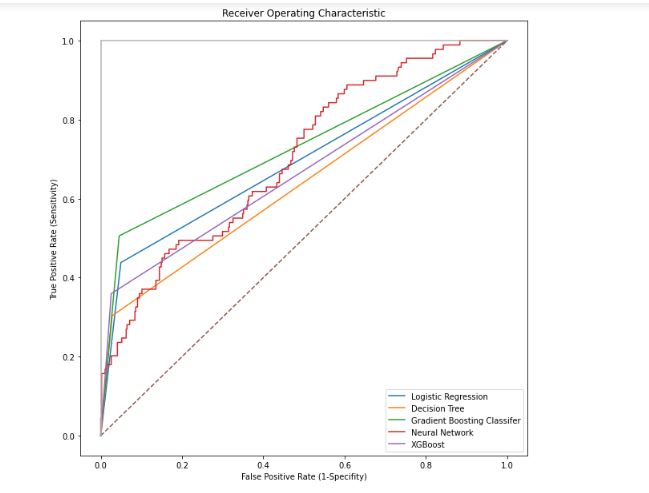


Figure : ROC Curve

The purpose of this analysis was to study the effectiveness of previous marketing campaigns and propose better data-driven marketing solutions that will increase the response rates.

* The exploratory analysis has shown the age group that the company serves, their income ranges, and the amount that the customers have spent on purchases.
* The analysis has shown that all products are positively correlated with income, except for gold products. Wine and meat have the strongest correlation with income. People with high income buy wine and meat products.
* The highest correlation with income is the Catalog purchases (0.7). This implies people with high-income shop via the company's catalog.
* Wine products produced the highest revenue followed by meat products.
* Sweet products and fruits produced the lowest revenues.
* The number of website visits per month is negatively correlated (-0.65) with income. This implies people with high income don’t shop via the website.
* The results indicate that campaign 4 was the most successful, followed by 3, 5, 1, and Campaign 2 was the worst-performing. The company can further investigate what attributes of this campaign made it successful, understand where and why people are accepting it, etc, and apply it to future campaigns.
* The analysis shows that customers between the ages of 25-50, who have an income of more than 60K dollars are the most recent
* Most customers who accepted the campaign were from Spain, followed by South Africa.
* The households without dependents spend more than the ones with one or more dependents.
* To increase the number of in-store purchases, the company should come up with ways of attracting more people to the website. For instance, through better UI/UX and/or more outreach online.
* Also, since income was one of the significant features, the company can further analyze what products each income class bought and improve its strategy of selling those specific products to that target demographic.
* In-store purchases were the top-performing channel, so the company can find other ways to improve the 2 others (website purchases and catalog purchases)

Modeling has been performed to show how the company can use the same data to predict campaign response (whether people will accept campaign offers or not. Different classification algorithms have been compared and the Gradient Boost classifier was the best performing model. The model can be used to predict whether people will accept marketing campaign offers or not.

## Model Training

This was a classification task and the models used to predict whether a campaign will be successful or not are. These models were trained on labeled data before predictions could be done. The dataset was split into training and testing sets.

*# Split the data into training and test sets*

*X = mktdata\_trf.drop(['Response'], axis=1)*

*y = mktdata\_trf['Response']*

*X\_train, X\_test, y\_train, y\_test = train\_test\_split(X, y, test\_size=0.25, random\_state=42)*

*How to choose the best model?*

## Machine Learning Training

Machine learning model training is a procedure that which a machine learning algorithm is usually stipulated with the training part of the dataset and learns from the supplied dataset values. In this context, machine learning models are usually undergoing training to benefit the business entities in different ways through quick processing of huge volumes of datasets, pattern identification, anomaly identification, or data correlations testing that would be quite difficult for human identification. It can be considered as the heart of the data science development lifecycle since it is through this lifecycle that the assigned data mining team toils intending to fit the best weights as well as biases to the machine learning algorithms, therefore, minimizing the loss function over forecasting range. The data mining team may utilize various types of loss functions but it all depends on the predefined project objectives and goals, the type of dataset as well as the type of machine learning algorithm. In this context, when the team utilizes supervised learning procedures, the model training process generates a relative mathematical representation of the dataset features and the dataset's target label. On the other hand, an unsupervised type of machine learning generates a relative mathematical representation among the dataset’s features themselves.

## Benefits of Machine Learning Model Training

This exercise can be referred to as the primary stage in data mining while using machine learning techniques; it results in a fully functioning model that later can be validated, tested as well as deployed. The machine learning models’ performance during the training process is always used to determine how accurate it works when it is deployed for the end-user group. Some factors are always considered during the machine learning model training; they include the quality level of the relative training dataset as well as the choice of the algorithm. In all cases, the training dataset is always divided into two sets and this includes training, validation as well as testing sets. However, the algorithm selection procedure is usually determined by the product’s end-users. In addition, more factors require considerations like the algorithm-model complexity level, model’s performance, results in interpretability, necessary computer resource needs and requirements as well as the model’s speed. The balancing process of these different requirements can turn the algorithm selection procedure to be a complicated procedure.

## How to train a model

Machine learning model training is a procedure that which a machine learning algorithm is usually stipulated with the training part of the dataset and learns from the supplied dataset values. In this context, machine learning models are usually undergoing training to benefit the business entities in different ways through quick processing of huge volumes of datasets, pattern identification, anomaly identification, or data correlations testing that would be quite difficult for human identification. Training any given machine learning model needs a systematic and repeatable procedure that relatively minimizes the use of the supplied training dataset as well as the time used by the data mining team or individual. Before the training phase starts, one needs to determine the project’s problem statement, the relative access process to the dataset as well as the cleaning procedure of the dataset to be presented to the chosen machine learning model. Also, one is required to determine the algorithm to be used as well as the parameters that will be considered during the training phase. Having considered all the discussed factors, splitting the dataset into two that is training dataset and the testing dataset follows and later prepare the chosen models for the training procedure.

### Dataset Splitting

The initially supplied training dataset can be termed to be a limited resource that requires to be allocated carefully. Some of the training datasets can be utilized in training the chosen machine learning models while the other part of it is utilized for testing the models’ accuracy. Note that no one can utilize the same dataset for each step. One cannot, therefore, test any machine learning model unless the model is supplied with new dataset values that have not been encountered earlier before. Splitting the supplied dataset into two or more sets can be considered to allow the data mining team to carry out training procedures as well as validation processes using a single and verified data source. In this context, the team or individual involved in the data mining project can outline the model’s overfit level, which simply refers to the mode that can perform well with the supplied training dataset but performs poorly with the testing set. There is a commonly used method of splitting the supplied training dataset which is referred to as cross-validation. For example, if the team chooses 10-fold cross-validation, the dataset will undergo a splitting process into ten different sets thus allowing one to carry out training as well as testing of the dataset ten times respectively. In this context, one can carry out the following;

1. Split the dataset into 10 equal parts alias folds,
2. Choose one part to be the hold-out fold,
3. Carry out the training process on the other nine parts,
4. Carry out the testing process on the model’s hold-out fold.
5. Repeat the procedure at least 10 times, with each time carrying out the selection of various folds to be the hold-out fold. The average performance across the ten hold-out folds then becomes the model’s performance estimate and it is referred to as the cross-validated score.

### Selection of machine learning algorithms to test

In the machine learning field of study, there are different algorithms that one can choose and this means that there is no relative sure method of determining which model can be the best. In several cases, an individual is likely to set to try dozens. The selection process of the candidate algorithms can depend on various factors such as;

1. The training dataset size,
2. The required output accuracy level and interpretability,
3. The relative training speed as well as time which is always inversely proportional to the accuracy percentage,
4. The training dataset’s linearity,
5. Dataset’s number of features.

### Tuning the Hyperparameters

Hyperparameters in a given dataset can be defined as the high-level features that are usually set by the data mining team or individual before the machine learning model is assembled and trained. It is always recommended for one to note that various attributes can be learned from the supplied training dataset but they cannot learn their inclusive hyperparameters. For example, if one is utilizing a regression algorithm, the relative machine learning model can be used to determine the regression coefficients on itself by analyzing the supplied dataset. In addition, it cannot be able to dictate the penalty strength that should be utilized in regularizing any available overabundance of dataset variables. Another example could be, that a machine learning model utilizing the random forest algorithm can be used to determine where decision trees will undergo splitting but the number of the trees to be utilized requires to be tuned beforehand.

### Fitting and tunning machine learning models

Having the prepared dataset and the determined machine learning model’s hyperparameters; is the favorite moment to train the chosen models. This procedure is usually essential in looping through the various algorithms utilizing each set of dataset’s hyperparameter values that the team or individual has decided to carry out explorations on. To practice all these, the individual or the data mining team requires to carry out the following sub-processes;

1. Splitting the dataset,
2. Selecting the appropriate algorithm,
3. Tuning the dataset’s hyperparameter values,
4. Training the dataset model,
5. Selecting another machine learning algorithm and repeating steps (iii) and (iv) respectively.

Next, carry out the selection of another set of dataset’s hyperparameter values that one needs to try out for the same algorithm, carry out the cross-validation process again and calculate the new scores. Once the team has tried out each hyperparameter value, it can repeat similar steps for any other additional machine learning algorithms. The team or individual carrying out the data mining processes is required to think of the above trials as track and field heats. Each machine learning algorithm demonstrates its capabilities with the various hyperparameter values. Now one can carry out the selection of the best version from each machine learning algorithm as well as send them to the final project competition.

### Choosing the Best Model

In this context, it is fully fit for one to test the best versions of each machine learning algorithm with aim of determining which algorithm stipulates the best model overall.

1. Make forecasts on the supplied testing dataset,
2. Define the ground truth for the target variable during the training phase of that model,
3. Define the performance metrics from the dataset’s forecasts as well as the ground truth target variable,
4. Run each model finalist with the testing dataset.

Once the dataset testing phase is over, the team or individual can carry out comparison exercises to the models' performances to define which models resulted to be better than others. The overall better model is required to have had good results in the training and testing parts respectively. It needs to have performed well in other performance testing metrics like the speed as well as empirical loss and also should adequately provide a solution to the problem statement’s question of the project.

## Methodical approach to model training

Utilizing a methodical as well as repeatable model training procedure can be said to be of paramount importance for any business entity or company strategizing on building successful machine learning or artificial intelligence models at scale. The central point of having all these are strategizing on the available resources, tools as well as libraries, and proper documentation in a single enterprise platform that will amplify team collaboration rather than hindering the process. In conclusion, the above section has had relative discussions on various machine learning models as well as algorithms. However, there may be rising questions relative to any beginner that may be regarding which type of model should they choose? The answer to this question is; that it depends on the business requirement or project requirements. However, it also depends on the associated features, the available dataset volume, the number of features, and complexity level among other things. In addition, in real-case practice, it is always recommended that one needs to start with the simplest machine learning model that can be utilized for a given problem and then gradually foster the complexity level as well as carry out testing on the accuracy with the aid of parameter tuning as well as cross-validation.

# 5.0 Key findings and discussions

Understanding the consumer journey is not a simple task. Using appropriate technology tools, customers use various forms to express their attitudes, needs, or values such as through comments, searches, blogs, likes, or even tweets. The supply of this consumer-curated data is seemingly endless and continues to grow. Big Data analytics and Business Intelligence can use this data for analysis and garnering insights that can be helpful in advertising and marketing. By doing this, organizations will improve their effectiveness to understand and reach consumers at different levels in the customer journey. The study intended to discuss how Artificial intelligence has disrupted marketing and how businesses benefit from adopting this marketing approach. It outlined some of the advantages intelligent marketing has over traditional methods of marketing. The world is very advanced, and automation of activities is the norm today. Artificial Intelligence technology in marketing is used to automate marketing processes through data analysis where marketers can observe their target audience and economic trends that may impact their business. AI automated software is therefore vital in intelligent marketing as it streamlines and automates marketing tasks and workflows. Examples of AI automated marketing solutions include image recognition, chatbots, content recommendation engines, and dynamic pricing on eCommerce sites. AI technology primarily uses statistical models to analyze data and predicts future actions based on past behavior.

The purpose of this analysis was to study the effectiveness of previous marketing campaigns and propose better data-driven marketing solutions that will increase the response rates.

* The exploratory analysis has shown the age group that the company serves, their income ranges, and the amount that the customers have spent on purchases.
* The analysis has shown that all products are positively correlated with income, except for gold products. Wine and meat have the strongest correlation with income. People with high income buy wine and meat products.
* The highest correlation with income is the Catalog purchases (0.7). This implies people with high-income shop via the company's catalog.
* Wine products produced the highest revenue followed by meat products.
* Sweet products and fruits produced the lowest revenues.
* The number of website visits per month is negatively correlated (-0.65) with income. This implies people with high income don’t shop via the website.
* The results indicate that campaign 4 was the most successful, followed by 3, 5, 1, and Campaign 2 was the worst-performing. The company can further investigate what attributes of this campaign made it successful, understand where and why people are accepting it, etc, and apply it to future campaigns.
* The analysis shows that customers between the ages of 25-50, who have an income of more than 60K dollars are the most recent customers
* Most customers who accepted the campaign were from Spain, followed by South Africa.
* The households without dependents spend more than the ones with one or more dependents.
* To increase the number of in-store purchases, the company should come up with ways of attracting more people to the website. For instance, through better UI/UX and/or more outreach online.
* Also, since income was one of the significant features, the company can further analyze what products each income class bought and improve its strategy of selling those specific products to that target demographic.
* In-store purchases were the top-performing channel, so the company can find other ways to improve the 2 others (website purchases and catalog purchases)

Modeling has been performed to show how the company can use the same data to predict campaign response (whether people will accept the campaign offer or not. Different classification algorithms have been compared and the Gradient Boost classifier was the best performing model. The model can be used to predict whether people will accept marketing campaign offers or not.

# 6.0 Conclusions and Recommendations

## 6.1 Conclusions

Intelligent marketing entails incorporating different technologies to help optimize marketing strategies and improve customer experience through personalization. The rise of eCommerce and big data set the foundation for intelligent marketing. To be precise, as soon as businesses started adapting to eCommerce, marketers were forced to change their marketing approach, especially since they could reach a wider audience through the internet. The eCommerce concept allows businesses to operate on several digital platforms such as websites, social media networks, mobile applications, and so on to market their brand consistently. However, these platforms provide real-time communication and have lots of data. Big data in intelligent marketing has proven to be indispensable as marketers can gather, analyze and use the massive amount of digital data to improve business operations. Additionally, enterprises with lots of traffic coming in on their online platforms require intelligent marketing tools to provide quick responses to customers and perform more tactical tools with less human intervention. Therefore, the intelligent marketing approach is popular among many businesses today as enterprises are rapidly adopting intelligent technology to make their business operations more efficient while maximizing profit and minimizing costs. Understanding the consumer journey is not a simple task. Using appropriate technology tools, customers use various forms to express their attitudes, needs, or values such as through comments, searches, blogs, likes, or even tweets. The supply of this consumer-curated data is seemingly endless and continues to grow. Big Data analytics and Business Intelligence can use this data for analysis and garnering insights that can be helpful in advertising and marketing. By doing this, organizations will improve their effectiveness to understand and reach consumers at different levels in the customer journey.

AI marketing tools provide businesses with valuable insight from collected datasets by analyzing the most effective advertisement placements or recommendations to increase customer engagement. Businesses personalize their products, services, and advertisements that specifically suit their customers based on their information. As customers' loyalty to a business increases based on their experiences, intelligent marketing allows organizations to build unique customer profiles that will further tailor personalize messages to customers at a prime stage of the consumer's lifecycle. Intelligent marketing has surpassed traditional marketing methods like direct emailing, media advertising, and physical retail marketing, which are no longer as effective as they were. Most importantly, Artificial Intelligence draws customers' patterns and identifies multiple drivers that will influence their decision to engage with a business brand.

With the help of machine learning algorithms, Artificial Intelligence allows businesses to obtain and analyze data from different sources. There are different types of data that marketers can focus on in intelligent marketing, including financial, customer, and operational data. Businesses can improve their performance by analyzing their operations data, such as shipping logistics providing a better customer experience. The analysis of collected information from customers will help businesses understand their consumers more and get more visibility on potential targeted customers.

Machine learning model training is a procedure that which a machine learning algorithm is usually stipulated with the training part of the dataset and learns from the supplied dataset values. In this context, machine learning models are usually undergoing training to benefit the business entities in different ways through quick processing of huge volumes of datasets, pattern identification, anomaly identification, or data correlations testing that would be quite difficult for human identification. It can be considered as the heart of the data science development lifecycle since it is through this lifecycle that the assigned data mining team toils intending to fit the best weights as well as biases to the machine learning algorithms, therefore, minimizing the loss function over forecasting range. The data mining team may utilize various types of loss functions but it all depends on the predefined project objectives and goals, the type of dataset as well as the type of machine learning algorithm. In this context, when the team utilizes supervised learning procedures, the model training process generates a relative mathematical representation of the dataset features and the dataset's target label. On the other hand, an unsupervised type of machine learning generates a relative mathematical representation among the dataset’s features themselves. Exploratory Data Analysis (EDA) entails analyzing the dataset to summarise its major characteristics of the same. It is majorly conducted via visualization methods. It has to be conducted before the modeling part. The technique allows the analyst to identify errors, and understand patterns or interesting relations in the dataset. It is in this stage that outliers are addressed as they might affect the final results, especially when creating the predictive model. EDA has been performed in this study to ensure that the results are reproducible and applicable in any business that has similar goals and objectives. It is in this stage that stakeholders get the answers to the questions they have raised. When the process is completed, the final features selected are then used to build the final model using sophisticated techniques and tools such as machine learning. The activities performed in this stage were data cleaning and variable transformation, checking outliers, checking nulls, feature engineering, and visualizing trends in the dataset.

## 6.2 Recommendations

Utilizing a methodical as well as repeatable model training procedure can be said to be of paramount importance for any business entity or company strategizing on building successful machine learning or artificial intelligence models at scale. The central point of having all these are strategizing on the available resources, tools as well as libraries, and proper documentation in a single enterprise platform that will amplify team collaboration rather than hindering the process. In conclusion, the above section has had relative discussions on various machine learning models as well as algorithms. However, there may be rising questions relative to any beginner that may be regarding which type of model should they choose? The answer to this question is; that it depends on the business requirement or project requirements. However, it also depends on the associated features, the available dataset volume, the number of features, and complexity level among other things. In addition, in real-case practice, it is always recommended that one needs to start with the simplest machine learning model that can be utilized for a given problem and then gradually foster the complexity level as well as carry out testing on the accuracy with the aid of parameter tuning as well as cross-validation.

However, the following steps are always crucial in any data science related project work;

1. Carrying out the collection process of the required dataset: This means that several data repositories have been made and thus can be used for the analysis.
2. Pre-processing the collected data: This process will incorporate the data cleaning procedures tasks like describing the statistical data characteristics, removal of dataset outliers, data duplicates, and all missing dataset values.
3. Performing feature engineering on the dataset: This procedure will be made of numerical features scaling, encoding the categorical attributes, and the transformation of the back dummy data variables to numerical nature.
4. Carrying out data exploration as well as data visualization: This process will be facilitated by applying the relevant dataset visualizing tools reliable in the data science field such as heatmaps, and histograms among others.

# References

De Bruyn, A., Viswanathan, V., Beh, Y. S., Brock, J. K.-U., and von Wangenheim, F. (2020). Artificial intelligence and marketing: Pitfalls and opportunities. *Journal of Interactive Marketing, 51*, 91-105.

Dimitrieska, S., Stankovska, A. and Efremova, T. (2018). Artificial intelligence and marketing. *Entrepreneurship, 6*(2), 298-304.

Gerrikagoitia, J. K., Castander, I., Rebón, F., and Alzua-Sorzabal, A. (2015). New trends of Intelligent E-Marketing based on Web Mining for e-shops. *Procedia-Social and Behavioral Sciences, 175*, 75-83.

Hunt, S. D. (2018). Advancing marketing strategy in the marketing discipline and beyond: from the promise, to neglect, to prominence, to fragment (to promise?). *Journal of Marketing Management, 34*(1-2), 16-51.

Iazzi, A., Trio, O., Pandurino, A., and Caione, A. (2015). The evaluation of market attractiveness through the marketing intelligence approach: a tool for the SMEs. *International Journal of Markets and Business Systems, 1*(2), 92-107.

Jain, P., and Aggarwal, K. (2020). Transforming marketing with artificial intelligence. *International Research Journal of Engineering and Technology, 7*(7), 3964-3976.

Kaczorowska-Spychalska, D. (2019). How chatbots influence marketing. *Management, 23*(1), 251-270.

Khanzode, C. A., and Sarode, R. D. (2016). Evolution of the world wide web: from web 1.0 to 6.0. *International Journal of Digital Library Services, 6*(2), 1-11.

Kotler, P., Kartajaya, H., and Setiawan, I. (2019). Marketing 3.0: From products to customers to the human spirit. In *Marketing wisdom* (pp. 139-156): Springer.

Kumar, V., Rajan, B., Venkatesan, R., and Lecinski, J. (2019). Understanding the role of artificial intelligence in personalized engagement marketing. *California Management Review, 61*(4), 135-155.

Lies, J. (2019). Marketing intelligence and big data: Digital marketing techniques on their way to becoming social engineering techniques in marketing. *International Journal of Interactive Multimedia & Artificial Intelligence, 5*(5).

Paschen, J., Kietzmann, J., and Kietzmann, T. C. (2019). Artificial intelligence (AI) and its implications for market knowledge in B2B marketing. *Journal of Business & Industrial Marketing*.

Rekha, A., Abdulla, M. S., and Asharaf, S. (2016). Artificial intelligence marketing: an application of a novel lightly trained support vector data description. *Journal of Information and Optimization Sciences, 37*(5), 681-691.

Safdar, N. M., Banja, J. D., and Meltzer, C. C. (2020). Ethical considerations in artificial intelligence. *European journal of radiology, 122*, 108768.

Shah, N., Engineer, S., Bhagat, N., Chauhan, H., and Shah, M. (2020). Research trends on the usage of machine learning and artificial intelligence in advertising. *Augmented Human Research, 5*(1), 1-15.

Siau, K., and Yang, Y. (2017). *Impact of artificial intelligence, robotics, and machine learning on sales and marketing.* Paper presented at the Twelve Annual Midwest Association for Information Systems Conference (MWAIS 2017).

Sung, E. C. (2021). The effects of augmented reality mobile app advertising: Viral marketing via shared social experience. *Journal of Business Research, 122*, 75-87.

Yadav, M., Joshi, Y., and Rahman, Z. (2015). Mobile social media: The new hybrid element of digital marketing communications. *Procedia-Social and Behavioral Sciences, 189*, 335-343.