Impact of AI on Marketing Strategy

**STUDENT NAME: MOHAMMED RUBEZ ARFANI**

**STUDENT ID: 20001411**

**SUPERVISOR: MIKE WATKINS**

**DECLARATION**

I certify that the ideas, designs and experimental work, results, analysis and conclusions set out in this dissertation are entirely my own effort, except where otherwise indicated and acknowledged. I did not use human participants in my MSc project. I further certify that the work is original and has not previously submitted for assessment in any other course or institution, except where specifically stated.

**ABSTRACT**

Artificial intelligence (AI) is becoming increasingly popular in a variety of industries, including research, commerce, healthcare, transportation, and academia. AI has also made its way to the marketing sector. This study is concerned with determining the extent to which AI has affected market strategies and the implications of various AI applications in marketing. Additionally, this study offers a useful application of an AI-based model that may facilitating business strategists make informed decisions about the promotional offers. The two objectives of this project are: 1) A review study on the impact of AI on marketing strategy; and 2) Real-world marketing application of AI (AI-based model creation).

A detailed analysis of several research on the effects of AI in the field of marketing is undertaken in order to attain objective 1. The review also addresses the risks and difficulties presented by AI in the marketing sector. Regarding objective 2, this work provides a regression model based on machine learning that estimates the successful promotional offer. To execute the predictive tasks, the proposed model integrates four machine learning regression algorithms, including k-nearest neighbours, stochastic grading descent, linear regression, and artificial neural network. Algorithm performance is measured in terms of r-squared error and negative root mean square error. According to experiment results, logistic regression is the most effective algorithm.

**TABLE OF CONTENTS**

[1 Introduction 7](#_Toc120329127)

[1.1 AI in Marketing: An Overview 8](#_Toc120329128)

[1.1.1 Working of AI with Chatbots 9](#_Toc120329129)

[1.1.2 NLP Engines 9](#_Toc120329130)

[1.1.3 Voice Processing Technology 10](#_Toc120329131)

[1.1.4 Image recognition and processing technology 11](#_Toc120329132)

[1.1.5 Data Mining 11](#_Toc120329133)

[1.1.6 Machine Learning 12](#_Toc120329134)

[1.1.7 Making up right solution 13](#_Toc120329135)

[1.1.8 AI for Customers 13](#_Toc120329136)

[1.1.9 Decision-making 14](#_Toc120329137)

[1.2 Aims & Objectives 15](#_Toc120329138)

[1.3 Contribution 16](#_Toc120329139)

[1.4 Document Organization 16](#_Toc120329140)

[2 Literature Review 18](#_Toc120329141)

[2.1 Review of Marketing Strategy 18](#_Toc120329142)

[2.2 Review of Artificial Intelligence 22](#_Toc120329143)

[2.3 Review of Impact of AI on Market Strategy 26](#_Toc120329144)

[3 Research Methodology 32](#_Toc120329145)

[3.1 Dataset Description & Processing 32](#_Toc120329146)

[3.2 Training, Validation, & Evaluation 36](#_Toc120329147)

[3.2.1 K-Fold Cross-Validation 36](#_Toc120329148)

[3.2.2 R-Squared Error 36](#_Toc120329149)

[3.2.3 Negative Root Mean Square Error 36](#_Toc120329150)

[3.3 Regression Models 37](#_Toc120329151)

[3.3.1 K-Nearest Neighbors 37](#_Toc120329152)

[3.3.2 Stochastic Grading Descent 37](#_Toc120329153)

[3.3.3 Linear Regression 38](#_Toc120329154)

[3.3.4 Artificial Neural Network 38](#_Toc120329155)

[3.4 Proposed Approach 39](#_Toc120329156)

[4 Results & Discussion 43](#_Toc120329157)

[5 Conclusion 46](#_Toc120329158)

[5.1 Future Work Recommendations 47](#_Toc120329159)

[6 Appendix 56](#_Toc120329160)

[A. Data Processing 56](#_Toc120329161)

[B. Defining Regression Models 57](#_Toc120329162)

[C. Performing Predictions 60](#_Toc120329163)

**TABLE OF FIGURES**

[Figure ‎2.1-Impact of AI on the outcomes of businesses (Jain & Aggarwal, 2020). 29](#_Toc120328774)

[Figure ‎2.2-Implication of AI-based marketing in different sectors (Jain & Aggarwal, 2020). 29](#_Toc120328775)

[Figure ‎3.1- Architecture of devised model for prediction of best performing promotional offer 32](#_Toc120328776)

[Figure ‎3.2-Sample count of offer ids in the dataset. 35](#_Toc120328777)

[Figure ‎3.3-Sample of dataset after being subjected to processing steps. 36](#_Toc120328778)

[Figure ‎3.4-Architecture of the proposed approach 40](#_Toc120328779)

[Figure ‎4.1-Negative RMSE scores of regression models with respect to each fold. 44](#_Toc120328780)

[Figure ‎4.2-R-squared score of regression models corresponding to each fold. 45](#_Toc120328781)

**LIST OF TABLES**

[Table ‎3.1-Description of dataset files and their corresponding features. 33](#_Toc120329380)

[Table ‎4.1-Experimental results of regression models in correspodance to each fold 43](#_Toc120329381)

**CHAPTER 1**

This chapter serves to introduce the topic under investigation. To do so, we explore the broader impact of artificial intelligence on several dimensions of marketing. We also list the goals and objectives of this study. Finally, we present this study's comprehensive contributions.

# Introduction

Artificial intelligence has been a hot topic for quite some time now. It is a type of technology that enables computers to perform tasks that are usually done by humans. In its most basic form, it is the ability to have machines complete tasks like responding to emails and creating spreadsheets, but the term can also refer to the increasing level of automation in society in general. The most significant application of AI is in the field of robotics. The study of AI could be used to assist robots in their work. For example, there are many robots that are currently being developed that can assist humans in their daily routine by transporting objects, monitoring, and cleaning facilities, and more. These technologies have the potential to revolutionize the workforce as they decrease the amount of time and labour needed by humans. Essentially, AI is a technology that has the potential to make life easier in many ways. It can also be used as an effective way to increase productivity by decreasing human interaction with machines (Calvert, 2008).

AI has been used in marketing for quite some time now and it is only getting more sophisticated with time. Some of the most popular uses of AI in marketing include predictive analytics, chatbots, and content creation. The long-term impact of AI is becoming a necessity for today’s marketing strategies. It enables marketers to have a better understanding of customer insights, creating more effective brand messaging, and facilitating personalized experiences. Predictive analytics has been around for decades, but the use of AI has allowed marketers to take this practice a step further. A variety of AI tools such as predictive modelling, machine learning, and natural language processing are used to crunch data and produce actionable insights into customer behaviour.

AI marketing strategies can help increase the effectiveness of a marketing campaign by sifting through data, digesting information, and creating targeted ads and landing pages to maximize engagement and conversions (Kabene, et al., 2006). The most common type of AI marketing is machine learning. This means that a computer program is trained on large amounts of data and then allowed to make predictions and decisions based on that training. Machine learning can be used to generate recommendations, classify information, and anticipate trends in the market. It’s also being used to create personalized ads in order to deliver a better targeted message (Conklin, 2002).

Moreover, the use of AI assistants like Google’s Assistant can also be seen as an important part in the future of digital marketing strategy. The term "assistant" is often used to describe a technology device or application which helps users carry out various functions. For example, a virtual assistant who can help with daily tasks would be an assistant; while the one which only provides answers to specific questions might be called a question answering service. AI-based marketing has been said to make people more loyal and increase the likelihood that they will pay more for a brand. Virtual assistants as forms of AI can be found in many products including smartphones, digital personal assistants such as Google Assistant, Amazon Alexa and Cortana, home automation devices like the Amazon Echo or Apple's Home Pod. A study has shown that people exposed to an artificial intelligence-driven virtual assistant experience a significant increase in their perception of the quality of products, even when those products are unrelated to the artificial intelligence technology (Br, 1987).

It can said that the marketing world has been transformed by AI technologies. These technologies have fuelled rapid growth in the digital marketing industry as a whole and can be used across every aspect of digital marketing to reach more customers, gather more data, and increase the efficiency of campaigns.

## AI in Marketing: An Overview

The use cases of AI technology in marketing are diverse. From helping marketers gain insights into their target audiences to enhancing customer experience, the use cases of AI are diverse and exciting. They can be used to:

* Increase the effectiveness of online ads by targeting the most relevant audience: Online ads have proven to be effective, but they often struggle to reach the right audience. These ads are intended for specific people, but the wrong person might see the ad and click on it. AI software can learn from past data and use that information to target the most relevant audiences with highly targeted ads. This can save companies money and increase overall effectiveness of ad campaigns (Sankrusme, 2011).
* Detect fraudulent transactions: AI is able to detect fake transactions or identify unusual patterns in data that might signal a transaction is fraudulent. AI can automate the process of identifying fraud by using machine learning and natural language processing.
* Personalize content based on customer preferences: Customers would not have to go through the hassle of searching for information on their own. With AI writing assistants, content can be personalized on a customer-by-customer basis.

### Working of AI with Chatbots

Nowadays, more people are using the internet to make purchases online or to dispel any doubts, but this has increased the amount of information available, which helps people make decisions more quickly. There is also the possibility of downloading apps for personal assistance.

While it's understandable for a website to want to use AI assistants, there are certain questions and sections of content that people may leave unchecked. This is because the assistants cannot anticipate all situations. The person writing the content usually has to fix any problems that occur with their article or knowledgebase. In the interest of fairness, the researchers have been working on this problem for a while now. Recently, (Mutolo & Jl, 2011) made some changes to the Knowledge base's AI assistant that will reduce or eliminate bias in certain situations. One of the most common issues authors noticed was that the Knowledgebase AI assistant would recommend content for a newbie audience. This is a good thing, but some people do not want to come across as too easy on their site. To ensure this doesn't happen, the AI assistant now has a learning period before recommending content to help it go through its training. It's also worth noting that this learning period is now shorter. When users are first on the site, they can be automatically sign up for service and start receiving recommendations from the Knowledgebase AI assistant by clicking "Get started." However, they won't get any additional recommendations until after their learning.

### NLP Engines

NLP Engine is a software that helps companies automate their sales process. It creates structured content, triggers responses, and organizes the entire sales process in one seamless software. This tool is crucial for the modern world of digital marketing where everything needs to be automated. It is crucial for the modern world of digital marketing where everything needs to be automated.

NLP Engine uses machine learning to help companies automate their sales process. The platform automatically generates personalized content based on the customer's answers to the questions, and it also helps them understand which of those answers are most important for the customer (Inglis, 2005).

NLP engine has a lot of use cases in different industries such as marketing, sales, recruitment, and so on. It is also used by businesses that want to improve their customer experience by providing them with customized content and targeted messages. NLP is a tool that companies can use to improve the customer experience and make their clients feel valued. With NLP, users are able to explore what makes them tick and reach out to them in the right way. NLP can also be used for other purposes such as predicting user behaviour patterns, improving your website’s UX, and also detect emotions by analysing data generated by users. NLP is used in many different fields to make individuals feel valued. In marketing, companies can use NLP to create personalized content based on user.

### Voice Processing Technology

Voice processing technologies have found their way into our lives, and the influence is going to continue to grow. Voice is now a key driver in how we interact with brands and products, how we buy and sell products, how we research products, and of course how brands interact with us. With voice technology the opportunities for creating custom experiences are immense (Tan & Ong, 2002). Voice assistants like Siri and Alexa have become a popular way to interact with brands. In fact, voice assistants are predicted to account for 50% of all customer interactions by 2020. The era of voice assistants has brought about a new way for brands to connect with customers and create emotional experiences. These interactions are already being used in a variety of industries, from retail to travel to banking. However, there is no standard method for how brands can use voice assistants in their marketing.

Voice is also quickly becoming a primary interface for marketing, as it allows marketers to deliver personalized experiences that are more personal than traditional channels. For example, Amazon’s Echo device uses voice recognition technology to answer users’ questions about products and provide them with information about the product’s price, shipping costs and reviews from other customers who have purchased it before them. Voice also has the potential to enhance interactions with products by allowing users to control their devices using voice commands. This could potentially lead to the development of smart home systems that follow users’ preferences and make it easier for them to manage their homes without having to switch from one application or device to another.

### Image recognition and processing technology

Image recognition technology is one of the most exciting and rapidly advancing areas in marketing. This technology has immense potential to influence the way brands communicate with their target audiences. This can be achieved through an algorithm-driven solution that searches for images which are relevant to a given query and automatically generates text on the basis of those images. This technology is evolving at an exponential speed.

The simple idea of image recognition has now taken the form of multi-objective and multi-modal solutions which can be used to solve a variety of marketing problems in a much more efficient way. The following are the market research services: - Image recognition software solutions that can be used to identify images and extract the text on them automatically. - Image recognition solutions for identifying brand logos. As more companies start using AI for marketing purposes, marketers can leverage this technology to reach their target audience better than ever before.

As AI becomes a bigger part of marketing, its implications also become clearer. Many marketers are concerned about the ethical implications of AI and its ability to create biased results. Though these concerns are valid, they shouldn’t prevent companies from using AI as long as it is being used transparently and with care (Br, 1987).

### Data Mining

A marketer's ability to produce data-driven, accurate marketing campaigns is also referred to as data mining. This can help boost a team's marketing efforts significantly and allows for highly customized messaging. Marketing effectiveness is highly dependent on the data collected and analysed by a company. This can help create targeted messaging that reaches the intended audience. Data-driven market research is also referred to as marketing analytics, predictive modelling, or predictive analysis. A company's ability to determine its target market is referred to as target marketing.

Data mining is an essential skill in today's marketing landscape. Without it, campaigns can be highly ineffective, costing marketers a lot of time and money. With this skill, marketers can produce data-driven campaigns that are highly customized and increase their conversion rates significantly. Mining social data is a great way to get to the bottom of a current situation or launch an effective marketing campaign. With analytics, marketers can identify trends and patterns within social media conversations that they might not have been able to otherwise (Sharp, 1991).

### Machine Learning

Machine learning is an umbrella term used to describe a broad range of tasks that are enabled by computers. "Machine learning" today tends to refer only to the subset of these tasks that are most relevant for computer vision and artificial intelligence applications in computer graphics and computer vision. The term "machine learning" is often used interchangeably with "artificial intelligence". But other terms such as deep learning, statistical learning, and pattern recognition are also used in computer vision and artificial intelligence (AI). Machine learning has applications in areas like data mining, computer security, speech recognition and natural language processing.

We are seeing a great increase in the analysis of data and Machine learning systems are here to help. They can predict event patterns and forecast responses, accordingly, allowing marketers to make sense of this massive data. So, marketers can understand root cause and probability of repeating those responding behaviour patterns. Marketers can map their data to predict the needs of consumers and tailor their messages accordingly. It's all about drill down and know what really matters to your customers and build personalized strategies with them in mind. This is not just a one-sided coin, as marketers gain new insights too. They are seeing an improvement in marketing effectiveness, better ROI, and reduced campaign costs. The best part is that there are many platforms out there and they are getting more advanced by the day. For instance, Google's machine learning systems can now predict your credit score before you apply for a loan.

Moreover, a prevalent use of machine learning has been observed. In the last two decades, machine learning has been used in industries like healthcare, e-commerce, and travel. These technologies can provide businesses with insight on how best to serve their customers and prioritize their efforts (Conklin, 2002).

In healthcare, machine learning could help doctors better predict the potential outcomes of patients with a certain disease. This can save time and money that would otherwise be spent on unnecessary tests or medications. In e-commerce, machine learning could predict what products are going to sell more based on past data about similar items. The use of machine learning for travel could help companies find the best routes to take, or prioritize flights based on factors like time, cost, and weather.

### Making up right solution

Marketing strategies for AI are no different from a human. This means that, extremely quickly, the platforms can recognize informative concepts and trends that will result in more engagement and sales. AI is having an enormous impact on marketing, just as it is in other industries. New AI-powered technology has increased efficiencies and costs across the board, without compromising quality or user experience. The following are some ways marketers can use AI to their advantage:

1. Collect data on all your current digital marketing channels – social media channels, websites, advertisements, and more (Sankrusme, 2011).
2. Use AI to analyse the data collected from your marketing channels and find insights that will help you improve your campaign.
3. Create an AI-powered digital lead generation tool that can deliver personalized results based on the data collected from your website, social media channels, and other sources.

AI solutions are developed using machine learning and artificial intelligence to understand human expression more accurately. These solutions can answer questions and respond to natural language queries in a variety of formats such as social media posts, email responses, blog posts and more.

### AI for Customers

AI has changed the way businesses interact with their clients and how customers interact with companies. In order to keep up, many companies are taking advantage of AI solutions. With the help of AI, customers can get quick answers to any questions they may have, and companies can improve their customer service by providing better personalized responses. AI has been used in customer service to learn from customers' responses and provide them with better services. AI can also help companies understand their customers better by understanding their needs and what they are looking for (Calvert, 2008).

The Internet has slowly introduced a revolution in customer service and shopping habits. With citizens leaving reviews, businesses are able to offer suggestions related to one's purchases for quicker decisions. The revolving door of information is also allowing businesses to better cater their consumers and make them feel valued.

### Decision-making

The Four Ps of Marketing are Product, Place, Promotion and Price. A product is anything that a company produces or sells. Examples of products include food, soap, computers, cars and even books. The purpose of marketing a product is to increase the demand for it so that it can be sold at a higher price and make money for the company. Place refers to where the product is being sold to customers.

A decision is a process that typically involves making a choice between two or more alternatives. A decision is an ordered process of assessing various options, evaluating them to determine which alternatives best serve the needs and interests of the decision maker and weighing the consequences, benefits and risks associated with each alternative.

The decision-making process has an impact on marketing because the decisions made in one stage can directly affect the next one. For example, if you decide based on data that is not representative of your target market, it will result in a poor marketing strategy that won't work for you. The marketing decision-making process is a continuous loop that involves gathering data, analysing it, and deciding on what to do next (Carter, et al., 2010).

The decision-making process starts with being able to identify the market that your product is intended for, followed by determining the target audience, understanding their needs, and defining your objectives in order to assess whether you have met those needs. Market research is the process of gathering data about targeted markets for products or services and their needs. It enables the development of better marketing strategies by identifying potential customers and understanding their needs.

## Aims & Objectives

Decisions are not always easy to make because they require careful consideration of what matters the most - profit or customer satisfaction - and how much risk you are willing to take for each outcome (Carter, et al., 2010). Additionally, in order to receive the greatest benefits from promotional offers, marketers must take knowledgeable decisions. This entails making an effort to take into account many factors of promotional offers, such as the amount spent, the profit made, the user encountered, etc.

Artificial intelligence and machine learning are the new buzzwords in marketing. They have the potential to transform how marketers reach their target audience by creating content that doesn't need human input (Pitta, et al., 1999). Not just in the world of science, but also in marketing and customer service. Today, AI is rewriting the rules of advertising and making it easier for brands to attract customers (Calvert, 2008). Investigating how AI affects marketing strategy is the goal of this project.

Furthermore, AI can improve the effectiveness of marketing campaigns by using machine learning data to identify the best-performing strategies and create personalized content for users. Promotional deals encompassing outsourcing are used by businesses to boost their marketing efforts and increase sales. An effective promotional offer can help with the following: (1) increasing revenue; (2) establishing relationships with customers; and (3) enabling consumers to refer other potential customers to the persuasive product. Effective marketing with promotional offers can be accomplished by concentrating largely on prospect data and anticipating possible clients. Processing various facets of promotional offers manually necessitates a significant amount of money, and work. An automated decision-making system is therefore necessary in order to help marketers make informed decisions. To accomplish this, we focus on developing an effective automated system that evaluates various elements of a promotional offer and inform the marketer of the best-performing offer. As a result, helping the marketers make the best choices for their promotional offerings and what to think about next. The proposed system entails supervised machine learning algorithms to generate the output as the best performing promotional offer.

In view of above discussion, I have posed following objectives of this study:

* Review the impact of AI on marketing strategies. The following sub objectives are included in this main goal:
* To know about the concept of Artificial Intelligence and Marketing Strategy.
* To know how artificial intelligence is impacting the marketing strategy of a business.
* To know whether such impact of AI on the marketing strategy is either positive or negative.
* To know whether there exists any threat to the privacy of customers while doing marketing with the usage of artificial intelligence.

To achieve this, I will provide a thorough review of studies that implies the impact of AI in different areas of marketing in Chapter ‎2.

* Establish an effective framework to display implication of AI in real world marketing problem.

To achieve this goal, I will focus on developing a supervised machine learning-based predictive framework that will supply users with information on the top - performing promotional offers. In Chapter ‎3, we discuss the proposed methodology in more detail.

## Contribution

Considering the aims and objectives of this study, the primary contributions of this study are outlined:

* A thorough review of studies is conducted to provide necessary information regarding impact of AI on marketing. Chapter 2 covers the repercussions of AI, both positive and negative, as well as issues with data privacy.
* Developing an efficient supervised machine learning-based systematic model to determine the applicability of AI in marketing strategy.
* Using four state-of-the-art machine learning regressors including K Nearest Neighbours (KNN-R), Stochastic Grading Descendent (SGD-R), Logistic Regression (LR), and an artificial neural network (ANN-R).
* The algorithms are trained and tested using a dataset from Kaggle that is a well-known web platform containing benchmark datasets.
* Performance of the algorithms is evaluated using two parameters including R squared Error (R2), and negative Root Mean Square Error (RMSE).
* The performance of algorithms is validated using K-Fold cross validation with the value of

## Document Organization

This document is divided into five Chapters. I introduced the underlying study in Chapter 1 along with outlining the aims and objectives that will be covered in this study. Chapter 2 involves a thorough review of the studies that will enable me to provide sufficient information regarding Objective 1. In Chapter 3, I present the proposed supervised machine learning-based framework that will predict the best performing offer. The experimental results of the proposed framework and findings regarding Objective 2 will be presented in Chapter 4. Finally, I conclude my study in Chapter 5.

**CHAPTER 2**

This chapter attempts to give a thorough background on the research topic of this study. In this chapter, I've also discussed about how AI is impacting marketing and the privacy issues it raises. Objective 1 mentioned in Section ‎1.2 is accomplished in this Chapter.

# Literature Review

The impact of AI in marketing has been studied extensively and the results are surprising. AI can be applied to marketing in many different ways such as machine learning, predictive analytics, chatbots, etc. Marketers have very high expectations of AI and believe that it will change the way they market. However, they are not always sure what AI can do for them and are positive about new developments in the field (Calvert, 2008).

AI technology has been widely adopted by the marketer’s toolbox with many different applications such as predictive analytics, content creation, social media automation, etc. The use of machine learning algorithms has led to a rise in demand for data scientists in the marketer’s team. This is because machine learning algorithms can be applied to various fields such as finance, healthcare and retail which require a deep understanding of data science techniques (Gotto, 2007).

## Review of Marketing Strategy

This literature review is a summary of the most popular marketing strategies and how they are used in the industry today. It also introduces the future of marketing strategy and what it could be like in 20 years. This literature review will cover three main topics:

* Marketing Strategy: As consumers increase their digital consumption, marketers have to use brand ambassadors to reach them in this new digital world. This literature review will cover three main topics: marketing strategy, digital branding, and e-commerce (Carter, et al., 2010).
* Marketing Technology: In today's world, marketing technology is taking over the market. Marketing automation software has been developed that can help businesses automate their marketing processes and grow their company. It handles most of the marketing tasks by gathering customer data, running campaigns on social media, and sending emails to customers.
* Marketing Automation: The marketing automation industry is booming, with over 50% of marketers saying that they use marketing automation. It has become the go-to tool for marketing managers to help them make decisions and automate tasks. In order to succeed in a competitive market, companies need to be able to use this tool at scale and efficiently.

Let start the review with some basic questions and their answers.

*What is marketing strategy?*

Marketing strategy is the overall plan that a company uses to market its products and services. It includes developing goals, objectives, and strategies for reaching those goals. It also includes setting target markets, determining how best to reach them, and developing plans to track and evaluate results (Evans, 2006). The goal of marketing strategy is to create value for customers. To do this, marketers must understand the customer's needs and wants. They must also identify the competition and develop strategies to stay ahead of the competition. Marketing strategy involves four main steps:

1. Define your goals. What do you want to achieve with your marketing? What are your specific objectives? What factors will influence whether or not you achieve those objectives?
2. Identify your target markets. Who is your target audience? What are their needs and wants? How can you reach them most effectively?
3. Develop plans to reach your target markets. How will you spend your marketing resources (money, time, effort, etc.) to achieve your objectives? What are the risks and rewards associated with each approach?
4. Track and evaluate results. Are you achieving the desired results? How can you improve your strategies in order to reach even greater success?

*What are the key elements of a successful marketing strategy?*

The key elements of a successful marketing strategy include setting realistic goals, targeting your market correctly, and using effective advertising and promotion techniques. Additionally, a good marketing strategy should be flexible enough to change as new information becomes available (Engel, 1974).

Setting realistic goals is essential to any successful marketing strategy. It is important to remember that a company's budget and resources are finite, so it is important to set goals that will be achievable while still providing value to the company.

Targeting your market correctly is also critical to a successful marketing strategy. Companies should target their market by researching what type of customers they want to reach, what types of products or services they offer, and where their customers live and work.

Effective advertising and promotion techniques are also essential to a successful marketing strategy. Companies should use advertising that is relevant to their target market, using effective media such as print, radio, television, the internet, and mobile devices. They should also use promotional techniques such as coupons, special offers, and targeted mailing lists (Devers, et al., 2003).

A marketing strategy is a plan that outlines how a company will reach its target market. It includes the identification of the customer segment, the creation of a product or service that meets their needs, and the development of marketing plans and tactics to promote the offering. A marketing strategy should be tailored to the specific needs of a company and its target market. It should also be consistent with the company's overall business strategy. The most important part of a marketing strategy is the identification of the customer segment. This involves understanding who your target market is, what they want, and what they are willing to pay for. Once you have this information, you can develop a product or service that meets their needs.

The development of marketing strategies and techniques is necessary if you have a good product or service that appeals to your target market. Public relations, direct marketing, and advertising are possible components of these initiatives. There are many different types of marketing strategies, but most fall into one of two main categories: direct or indirect. *Direct marketing strategies* involve selling products or services directly to consumers. *Indirect marketing strategies* involve selling products or services through third parties, such as retailers or distributors (Group, 2016).

Direct marketing strategies are the most common type of marketing strategy. Direct marketing strategies can be divided into two main categories: direct mail and telemarketing. *Direct mail* is the oldest form of direct marketing, and it involves sending out letters or flyers to consumers. *Telemarketing* is a more recent form of direct marketing, and it involves contacting consumers by phone.

Indirect marketing strategies are the second most common type of marketing strategy. Indirect marketing strategies can be divided into two main categories: advertising and promotion. *Advertising* is the oldest form of indirect marketing, and it involves placing ads in newspapers, magazines, or on TV. *Promotion* is the newer form of indirect marketing, and it involves providing incentives (such as free products) to customers who buy products from a company.

There are many different types of advertising campaigns, but most fall into one of three main categories: product placement, brand awareness, and direct response advertising. *Product placement* is the oldest form of advertising, and it involves placing products in TV shows, movies, or commercials. *Brand awareness* is the most recent form of advertising, and it involves creating a name for a company that people know without ever having seen an advertisement for that company. *Direct response advertising* is the most common type of advertising campaign, and it involves contacting consumers directly to sell them products (Betancourt, et al., 2005).

There are many different types of promotion campaigns, but most fall into one of three main categories: customer retention campaigns, customer acquisition campaigns, and lead generation campaigns. *Customer retention campaigns* are designed to keep customers from leaving a company’s product or service line-up. *Customer acquisition campaigns* are designed to bring new customers into a company’s product or service line-up. *Lead generation campaigns* are designed to generate leads (people who have not yet bought a product from a company) for a company’s product or service line-up.

Direct marketing strategies can be classified according to how the company obtains information about its target market: through surveys, interviews, or focus groups; through direct mail; or through online channels such as websites or social media. Indirect marketing strategies can be divided into four main categories: advertising, public relations, direct sales, and distribution channels.

*Surveys* are the most common way to collect information about a target market. Companies use surveys to gather information about what products or services people want, how much they are willing to pay for them, and how likely they are to recommend them to others (Armstrong, 1982).

*Interviewing* is another way to collect information about a target market. Companies interview potential customers, clients, or employees to find out what products or services they would like to see offered, what their needs are, and how best to meet them.

*Focus groups* are a type of interviewing that allows companies to gather feedback from a group of people who have similar interests or experiences. Focus groups can be used to test new products or marketing campaigns before they are released to the public, and they can also be used to gather feedback on existing products or services.

*Direct mail* is another way in which companies collect information about their target markets. Direct mail is a form of advertising that uses letters, postcards, flyers, or other types of advertisements to reach potential customers.

*Online channels* such as websites and social media are also ways companies collect information about their target markets. Companies use websites and social media platforms to advertise their products and services, connect with potential customers, and gather feedback from users.

*Advertising* is the most common form of indirect marketing. It involves paying for space in newspapers, magazines, radio stations, and online platforms such as Google AdWords. *Public relations* are another common indirect marketing strategy. It involves creating positive publicity for a company by distributing favourable news stories about it to journalists and other influential people*. Direct sales* are another indirect marketing strategy. It involves selling products or services through independent retailers rather than through a company’s own website. *Distribution channels* are the fourth main type of indirect marketing strategy. They include such methods as word-of-mouth advertising, referral programs, and franchising (Carter, et al., 2010).

## Review of Artificial Intelligence

AI is a rapidly evolving technology that has the potential to change the way we work. The literature review of AI will provide a comprehensive understanding of what AI is and how it can be applied in different industries.

The literature review starts with an overview of the history of AI, followed by an introduction to machine learning, then discusses challenges and opportunities for AI in various industries. It also provides insights on how the development of AI is changing our society. The literature review concludes with recommendations for future research into AI.

Artificial intelligence is the process of designing computer programs that simulate intelligent behaviour. AI research is divided into two main branches: machine learning and artificial general intelligence. Machine learning is the study of how computers can be taught to improve their performance on specific tasks without being explicitly programmed. Artificial general intelligence is the goal of creating a computer program that can effectively solve problems in many different domains, including natural language processing, problem solving, and knowledge representation (Evans, 2006).

AI has been used in a number of fields, including computer vision, natural language processing, and machine learning. One of the most common uses of AI is in customer service. Customer service robots are used to help customers with their queries by providing information and answering questions. AI is also being used to create chatbots, which are computer programs that can communicate with humans in a natural way. AI research has been underway for centuries, but the field has recently seen a resurgence in interest and funding due to the potential benefits it offers. It has the potential to improve many areas of our lives, from health care to economic decisions (Masys, 2002).

Natural language processing (NLP), computer vision, and machine learning are three subfields of AI study. *NLP* is focused with the development of machines capable of comprehending human speech and written materials (Chowdhary, 2020).NLP enables machines to accurately and quickly comprehend speech signals and respond to it. As a result, machines are now better equipped to do jobs that call for human interaction, including customer service or automated decision-making. NLP techniques have been widely implemented in various areas such as in opinion mining (Sun, et al., 2017), bioinformatics (Zeng, et al., 2015), information retrieval (Strzalkowski, et al., 1999), and many more. *Computer vision* focuses on the ability of machines to identify objects and scenes in 3-D spaces (Shapiro, et al., 2001). With the growing technological interest of researchers, computer vision has been largely implemented to address many real-world problems such as colour assessment (Milovanovic, et al., 2021), welding inspection (Cruz, et al., 2020) etc. Recently, (Eyiokur, et al., 2022) developed a computer vision-based system that can detect face-to-hand interaction and facemasks to limit the spread of coronavirus.

*Machine learning* concerns the development of algorithms that can learn from experience and automatically improve their performance over time. This has led to machines becoming more intelligent and able to carry out tasks that were once thought to be beyond their capabilities (Evans, 2006). Different algorithms are used by machine learning to carry out predictive tasks. First, training data is given to the algorithms so they can become familiar with the patterns in correspondence to target predictive class in the input data (Saad, et al., 2021). The algorithms are then supplied a test set of unseen input data to evaluate their performance. In consideration of this, the following four categories of machine learning are listed:

* ***Supervised Learning***: Labelled training data is provided to the machine learning algorithms. The input data and output target variable of the algorithm are established in this type of machine learning. Supervised learning is further categorized as classification and regression. Classification is concerned with categorical data and regression deals with continuous data.
* ***Unsupervised Learning:*** This is the type of machine learning that trains the algorithms using unlabelled data. The machine learning algorithm processes the input data for meaningful correlations.
* ***Semi-supervised Learning:*** Semi-supervised machine learning combines the two previously stated types. Labelled input data are given to the machine learning algorithms so they can be trained. The algorithms could also independently analyse the input dataset and establish a conceptual understanding of the data.
* ***Reinforcement Learning:*** The algorithms are trained using a rewarding mechanism in this type of machine learning. The intended result is rewarded, and the unintended result is penalized. The reinforcement learning model typically acts and learns by incorporating the trial-and-error method.

Several machine learning algorithms such as decision tree (Kumar, 2020); (Nabipour, et al., 2020), support vector machines (Ma & Sun, 2020), topic models (Ansari, et al., 2018), deep neural networks (Zhang & Luo, 2022), tree ensembles (Zhang & Luo, 2022) among others, have been implemented in the field of marketing for predictive tasks and information gathering. A marketing decision-making system has been developed by researcher (Kumar, 2020) that can help with planning process and lighten the load on organizations. The system is a decision tree/artificial neural network hybrid. Another study (Nabipour, et al., 2020) predicted the trends in stock market using several machine learning algorithms including KNN, support vector classifier, extreme gradient boosting, adaptive boosting, random forest, decision tree, long-short term memory, artificial neural network, and recurrent neural network. Similarly, the authors (Cui & Curry, 2005) made use of support vector machine in marketing application for high efficacy.

AI can also assist us in solving many other complex problems. For instance, AI can help us design more efficient cars (Aulinas & Sjafrie, 2021) or medical devices (Chinzei, et al., 2018). AI can also help us identify and prevent crime (Quest, et al., 2018). AI has the potential to improve our lives in countless ways. Another key benefit of AI is that it can improve our cognitive abilities. For example, AI can help us learn new information faster or remember things better. AI also helps us understand and process information more effectively (Conklin, 2002).

*Threats:* Overall, AI offers great potential for positive impact, on the contrary, it comes with significant risks. It is important to weigh these risks and benefits carefully before making any decisions. One concern is that AI could be used to automate jobs that are considered “low-skilled” or “menial” work (Wajcman, 2017). This could lead to a decrease in the number of jobs available and an increase in unemployment. Another key concern is that AI could be used to create weapons that are able to target specific individuals or groups (Johnson, 2019). Additionally, AI can potentially enable robots to do dangerous things on their own. Moreover, AI could lead to massive economic shifts that could badly affect the working and living conditions of many people. Adopting AI, according to the authors (Bughin, et al., 2018) , may widen the differences between workers, companies, and nations.

*Challenges:* Despite these advances, there are still many challenges facing researchers working on AI. Making sure that the robots can learn from experience and acquire their own capabilities is one of the primary issues facing researchers working on AI. Developing machines that can efficiently communicate with humans is another difficult task. Another significant is the difficulty of creating machines that are able to learn from experience and develop their own skills. This is due in part to the fact that learning requires a high level of abstraction, which is not always easy for computers to achieve. Additionally, it is difficult for machines to generalize knowledge from one situation to another.

## Review of Impact of AI on Market Strategy

The technology has the potential to change the way we approach marketing. It can help businesses leverage new opportunities and be more effective in their activities. However, the technology is still in its infancy, and there are many factors that need to be considered before it can be used effectively in a business setting. This literature review will discuss how AI is impacting market strategy and what these impacts are (Blunt, ).

The impact of artificial intelligence (AI) on market strategy is still being studied, but there are some potential benefits and drawbacks that should be considered. The authors (Jarek & Mazurek, 2019) stated that primarily three areas of AI are implemented in the field of marketing including decision-making, text-recognition, and image-recognition. Moreover, AI is also employed for automated robots and voice recognition technologies in this field by big organizations such as Apple. However, the cost concerning the development of AI-driven marketing systems pose an uncertainty of practicing these innovations.

A rapid shift of marketing towards AI has been observed by (Siau & Yang, 2017). However, the authors stated that in the near future, AI will eventually replace some positions in sales and marketing by outperforming humans. AI robots will replace the salespeople, and web platforms can update autonomously relying on eye-tracking data. The authors urged advertisers to acquire skillset that cannot be duplicated by robots, such as invention, originality, and design.

Another research (Marinchak, et al., 2018) evaluated how AI will affect marketing in the long-term. Along with the growth in customers’ satisfaction, the authors reasoned about the growing ethical and security issues that AI agents are posing. The costs related to the development of new concepts and the uncertainty of the outcome of their implementation may affect the organization in putting these innovations into practice as well.

(Dimitrieska, et al., 2018) analysed the impact of AI on marketing. The authors stated that the way businesses operate has changed as a result of increased competition, productivity, and responsiveness. AI market integration can lead to more spontaneous and easy customer interactions. By utilising platforms like Google, Facebook, and YouTube, among others, AI enables businesses to understand the relevance and personalisation of their clients. Using these platforms, marketers can connect with up to billions of individuals according to their preferences.

The authors (Soni, et al., 2019) analyse the impact of AI on the marketing strategy employed by business models. The authors discussed two primary attributes that makes AI the fundamental technology for the automation. The authors assessed the growing accessibility of big data and hardware accelerators to be the main attributes. They also described the dark side of AI along with AI divide.

Similar to this, the authors (Shahid & Li, 2019) used a questionnaire to conduct a survey and examine the impact of AI on marketing. To determine the qualities that allow the AI to be incorporated in marketing, the authors conducted interviews with marketing experts. The authors emphasised the benefits of AI for marketing while also pointing out the difficult technological compatibility of current systems with AI.

The study (Overgoor, et al., 2019) illustrated the usage of Cross Industry Standard Process for Data Mining (CRISP – DM) for establishing AI-driven systems to address marketing issues. Conversely, the study conducted by (Annor Antwi & Al-Dherasi, 2019) concludes that AI integrated systems with other technology such as data analytics, big-data, and machine learning can provide more efficient outcomes. The authors also stated the drawbacks of AI such as the AI-driven systems being costly and their relevant privacy threats.

Reviewing the aforementioned studies has led us to the conclusion that the impact of AI on market strategy will differ based on the particular industry and company, hence there is no universally applicable solution. According to a study of over 400 AI application scenarios encompassing 19 sectors and nine enterprises, the sales and marketing areas deploy AI at the highest proportion (Michael, et al., 2018). Utilizing AI, marketing methods including the next-best promo for buyers (Davenport, et al., 2011) and algorithmic online advertisement purchasing (Parekh, 2018) are evaluated for their effectiveness.

AI might also speed up the process of finding new opportunities and improve decision-making for businesses. Improved decision-making speed and accuracy are further possible advantages of AI, which may result in more profitable and effective trading activities. AI could assist strategists in developing more customized marketing efforts by helping them better understand consumer preferences and behaviour. AI may also make it easier for businesses to find and seize fresh market possibilities. First off, AI could assist businesses in automating procedures related to customer care or marketing. For companies, this could result in time and resource savings that eventually increase profits. Additionally, AI might be used to detect and prevent fraud, which would be advantageous for businesses. Overall, AI could have a substantial impact on market strategy by assisting businesses in cost- and resource-saving operational optimization. The potential for AI to help businesses better understand their clients is another advantage. This might lead to increased revenue and profitability (Carter, et al., 2010). Businesses might design more specialised marketing and products by better understanding client preferences and behaviour. AI could also assist companies in monitoring client data and trends so they can make better judgments.

To determine whether or not this technology is a good fit for them, organisations need still weigh the possible dangers and advantages of it. AI might, for instance, increase the precision of market price and assist market participants in finding opportunities more rapidly. AI may also help marketers in data analysis and decision-making on trading strategies. However, there are concerns that could arise, such as the possibility of job loss and reliance on algorithms that might not always be reliable. Citibank conducted a survey that reported 47% of positions being at risk of unemployment in US due to emergence of AI. Similarly, it is 35% across UK, 57% in OECD, and 77% in China (Yang & Siau, 2018). Moreover, if AI is not properly managed, it could lead to serious financial system vulnerabilities and raise the possibility of cyber-attacks. Additionally, organisations risk missing crucial client signals if they overly rely on AI to make choices. Making ineffective financial decisions as a result could result in lower profitability.

In conclusion, the use of AI-driven marketing systems is growing, which is reducing marketer accountability as well as delegating work to machines. Moreover, AI's principal drawback is the loss of employment and resulting public unrest (Wang & Siau, 2019). The statistics relating to the outcomes attained through the application of AI Marketing are shown in Figure ‎2.1. AI adoption is highlighted in Figure ‎2.2 across several industries and fields.

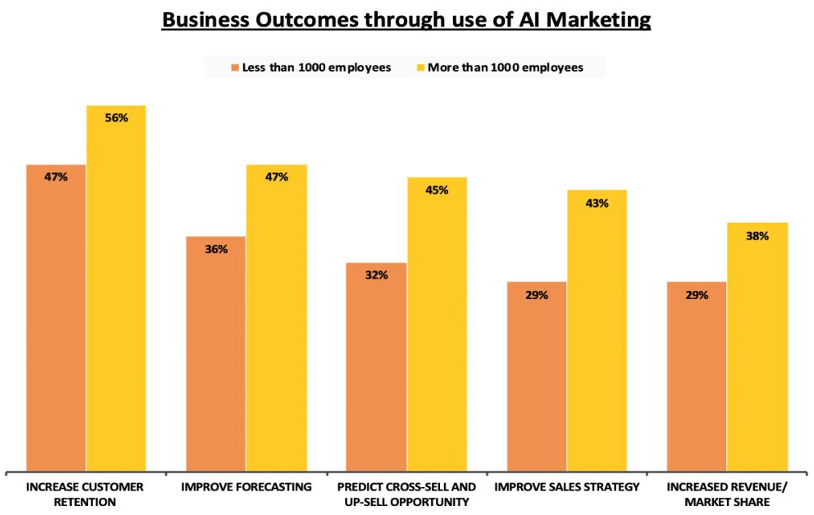


Figure ‎2.1-Impact of AI on the outcomes of businesses (Jain & Aggarwal, 2020).

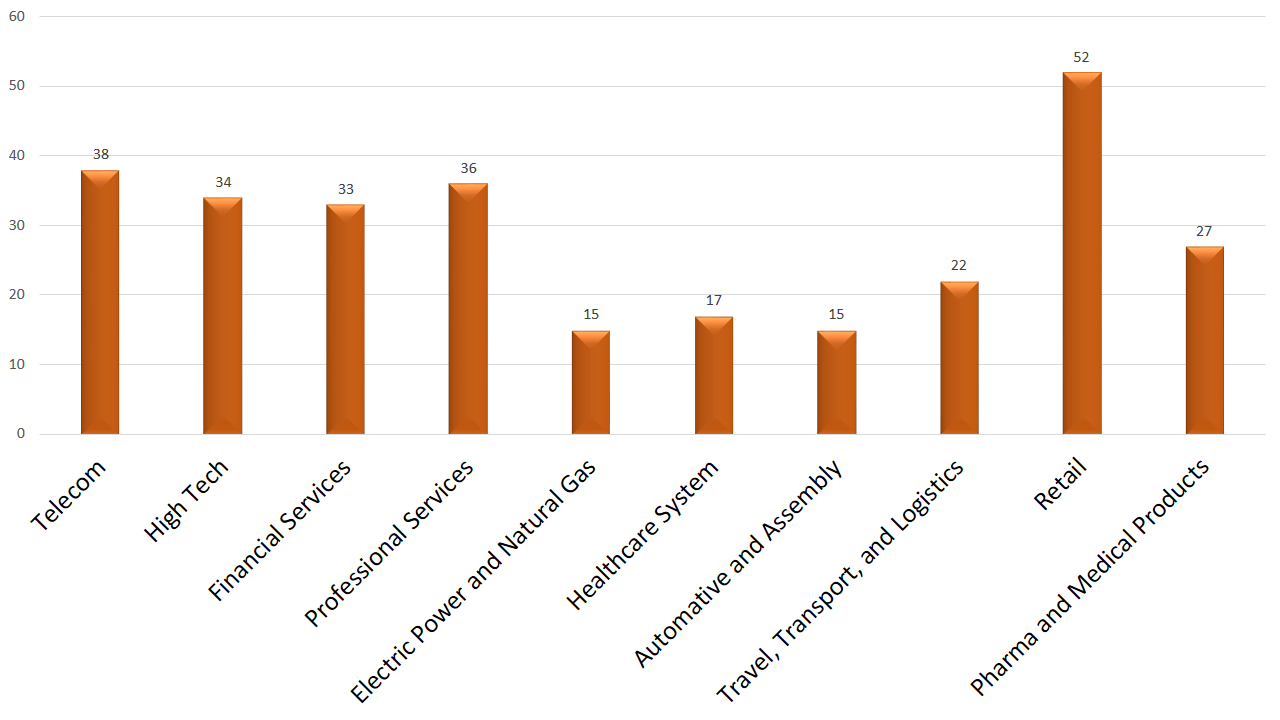


Figure ‎2.2-Implication of AI-based marketing in different sectors (Jain & Aggarwal, 2020).

*Data Privacy*: The author (Columbus, 2019) stated that Elon Musk and other experts consider AI to be "threatening". Because of the challenges it poses with regards to information confidentiality, computational inefficiencies, especially ethics, AI may well not strictly adhere to all of its promises. The author (Conklin, 2002) discussed the current state of the market, and how companies are integrating AI into their customer-organization relationship. First, AI can track customers’ online activity and preferences in order to target them with ads. Second, AI can monitor customer conversations in order to learn more about their preferences and interests. Third, AI can collect data about customers’ physical location in order to target them with marketing campaigns specific to their location. Finally, AI can collect data about customers’ social media activity in order to better understand their networks and preferences. The author also discussed some possible solutions that could be put in place to protect customers’ privacy. Another author (Heidi, 2018) quoted that the customers are forging their information to throw off targeted advertisements that use their information without their consent.

Businesses are using AI to understand the customers’ purchasing behaviour and their habits (Helbing, 2019). The public is becoming more informed regarding this activity that is viewed as growing concern and discomfort consequently, posing threats to buyers' confidentiality, freedom, as well as wellbeing. Additionally, the privacy concerns have escalated together with the rise of new vulnerabilities due to the growing big data generated by machine learning algorithms as a consequence of the greater internet traffic. AI integration in marketing could have negative effects on customer-business relationships (Vladeck, 2015). Some of the effects involve the possibility of fraud, whereby micro-targeted cognitive and behavioural marketing could persuade customers to make decisions despite their interests (Evans, 2006); (Grafanaki, 2016), and the refusal from some commodities toward certain potential consumers (Citron & Pasquale, 2014), and it may also be discriminating because of the widespread of adverts, tariffs, as well as exposed information (Barocas & Selbst, 2016).

Businesses as well as prospects for consumers may be adversely affected by anti-competitive behaviour and technological manipulation (Janka & Uhsler, 2018). Given the widespread customer monitoring and AI-based characterization, also the comparatively more "conventional" risks to confidentiality had also increased (Zuboff, 2015).

The study (Kabene, et al., 2006) elaborated that there are a number of ways in which AI could threaten privacy if used for marketing purposes. For example, if a company used AI to analyse customer data in order to personalise marketing content, this could potentially reveal sensitive information about individuals, such as their political views or personal finances. If this data was then shared with third parties without the consent of the individuals concerned, this would be an invasion of their privacy. In addition, if companies used AI to automatically generate targeted advertising based on customer data, this could be seen as intrusive and offensive. Finally, if companies used AI to predict which customers would be most likely to engage with marketing campaigns, this could lead to the profiling of individuals in a way that is not acceptable.

Additionally, AI might be used to monitor and evaluate our online behaviour in order to more effectively target advertisements and sales pitches. Furthermore, AI might be used to provide personalised recommendations for us based on our past behaviour, which might reveal sensitive data (e.g., personal health information, financial data, etc.). In order to alter product rankings and ratings, AI may also be used to generate "fake" reviews or ratings (Salminen, et al., 2022).

One way in which our privacy could be threatened is if our personal information is stolen by hackers. If our personal information is stolen, it could be used to identity us and to steal our money or other valuable assets. It is important that we take steps to protect ourselves from this risk, by using strong passwords and encryption software when we are online (Shenton, 2004).

Overall, it is important that we take steps to protect ourselves from the risks posed by the use of AI for marketing. By using caution when sharing personal information online, and by using protective measures such as password protection and encryption software, we can help ensure that our privacy remains protected.

**CHAPTER 3**

# Research Methodology

For this project, I created a system that indicates to marketers what promotional offers are most likely to increase sales and improve customer experience. The project was developed in the Anaconda environment using the Python language. Machine learning regression algorithms are deployed in this project. The algorithms are trained and validated using a dataset obtained from Kaggle. To yield the final model, three Jupyter notebooks have been generated and executed. Figure ‎3.1 showcases workflow of the proposed project. This chapter covers the methods used in this project.

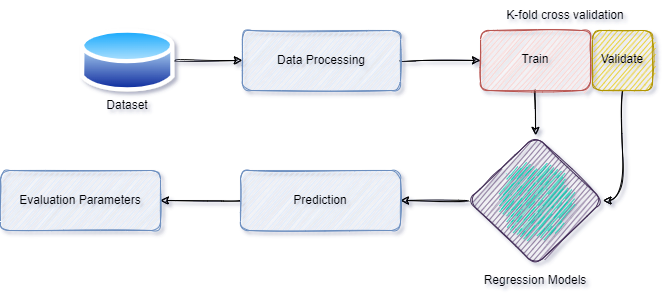


Figure ‎3.1- Architecture of devised model for prediction of best performing promotional offer

## Dataset Description & Processing

This project uses Starbucks capstone dataset obtained from Kaggle (Sabar, 2020). The dataset is comprised of three files in Jason format including “*portfolio.json”-*comprising promotional offer ids and related data such as type of offer, duration, etc., “*transcript.json”*-comprising of transaction records, received offers, viewed offers, and promotional offers that have been completed, and *“profile.json”-*comprising customer demographic data. The description of features involved in these three files is given in Table ‎3.1.

Table ‎3.1-Description of dataset files and their corresponding features.

|  |  |  |  |
| --- | --- | --- | --- |
| **File** | **Feature** | **Description** | **Type** |
| Portfolio.json | id | Unique identification number of offer | string |
| offer\_type | Corresponding type of the promotional offer such as information, discount, or BOGO. | string |
| difficulty | The minimal amount involved in completing an offer | int |
| reward | incentive after finishing an offer | int |
| duration | Timeframe in days in which the offer is live | int |
| channels | n/a | string |
| Profile.json | age | Customers’ age | int |
| became\_member\_on | Customer account creation date | int |
| gender | Customers’ gender | string |
| id | Customer identification number | string |
| income | Income of customer | float |
| Transcript.json | event | Description of record | str |
| person | Customer identification number | str |
| Time | Time duration in hours since the start of the test, beginning value of t=0 | int |
| value | Transaction amount or offer id based on the record | Dict of string |

To begin with experiments, the datasets are read into the Jupyter notebook using the code:

|  |
| --- |
| user\_data\_file = "./data/profile.json"  promo\_data\_file = "./data/portfolio.json"  transaction\_data\_file = "./data/transcript.json"  user\_data = pd.read\_json(user\_data\_file, lines=True)  promo\_data = pd.read\_json(promo\_data\_file, lines=True)  transaction\_data = pd.read\_json(transaction\_data\_file, lines=True) |

Then, the dataset is processed to extract the features that useful for this study. To do so, these three files were combined into a dataframe as displayed by the code below:

|  |
| --- |
| transactions = pd.DataFrame()  offer\_completed = pd.DataFrame()  offer\_received = pd.DataFrame()  data\_set = pd.DataFrame() |

Since, the model focuses on predicting the best performing promotional offer, I obtained all of the transaction type data instances. Furthermore, the record or data instances corresponding to completed events are extracted. The resulting dataFrame is saved as “data\_set”.

|  |
| --- |
| transactions = transaction\_data.loc[transaction\_data['event'] == 'transaction']  offer\_completed = transaction\_data.loc[transaction\_data['event'] == 'offer completed']  transactions = transactions.append(offer\_completed)  transactions = transactions.sort\_index()  data\_set = transactions |

Afterward, necessary steps were conducted to obtain only useful features and drop remaining of the features from the dataset. The extracted features include *'amount', 'gender', 'user\_income', 'offer\_type', 'offer\_diff', 'offer\_reward', 'web', 'mobile', 'social', and 'offer\_id'.*

|  |
| --- |
| # User Data  data\_set['gender'] = data\_set['person'].map(user\_data.set\_index('id')['gender'])  data\_set['user\_age'] = data\_set['person'].map(user\_data.set\_index('id')['age'])  data\_set['user\_income'] = data\_set['person'].map(user\_data.set\_index('id')['income'])  # Promo Data  data\_set['offer\_id'] = [d.get('offer\_id') for d in transactions.value]  data\_set['offer\_id'] = data\_set['offer\_id'].shift(periods=-1)  data\_set['amount'] = [d.get('amount') for d in transactions.value]  data\_set = data\_set[data\_set.event != 'offer completed']  data\_set['offer\_type'] = data\_set['offer\_id'].map(promo\_data.set\_index('id')['offer\_type']).astype('category')  data\_set['offer\_diff'] = data\_set['offer\_id'].map(promo\_data.set\_index('id')['difficulty'])  data\_set['offer\_reward'] = data\_set['offer\_id'].map(promo\_data.set\_index('id')['reward'])  offer\_channels = data\_set['offer\_id'].map(promo\_data.set\_index('id')['channels']).apply(pd.Series)  data\_set['web'] = offer\_channels[0].apply(lambda i: i if i == 0 else 1)  data\_set['email'] = offer\_channels[1].apply(lambda i: i if i == 0 else 1)  data\_set['mobile'] = offer\_channels[2].apply(lambda i: i if i == 0 else 1)  data\_set['social'] = offer\_channels[3].apply(lambda i: i if i == 0 else 1)  del data\_set['person']  del data\_set['value']  del data\_set['event'] |

Since, the machine learning models used in this study are the regression models that deal with continuous data, we converted the *‘gender’* feature string values to integers. Similarly, the string data type of ‘*offer\_type’* was converted to integer by replacing the string values by integer values. Finally, the resulting dataset was saved as a CSV file as shown in the code below:

|  |
| --- |
| data\_set['gender'] = data\_set['gender'].replace([np.nan, 'M', 'F', 'O'], [0, 1, 2, 3])  data\_set['offer\_type'] = data\_set['offer\_type'].replace([np.nan, 'bogo', 'discount', 'informational'], [0, 1, 2, 3])  data\_set = data\_set.fillna(0)  data\_set.to\_csv('./data/data\_set.csv') |

The resulting dataset comprised of 293,879 data instances or records. I further processed the *“offer\_id”* feature and integrated it with the created feature *“amount”.* This resulted in 9 values in the “*offer\_id*” feature that will be further treated as the target variables for the machine learning regression models. The target classes with “0” values amounted to 108,338 whereas, distribution of remaining of the offer ids is showcased in Figure ‎3.2. A sample of the resulting dataset in CSV format is presented in Figure ‎3.3.

Figure ‎3.2-Sample count of offer ids in the dataset.

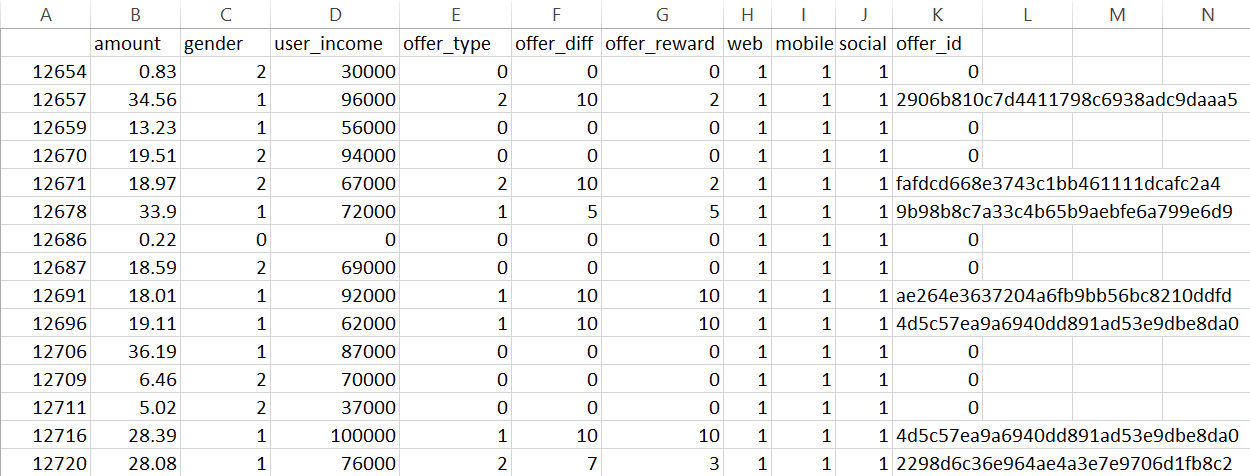


Figure ‎3.3-Sample of dataset after being subjected to processing steps.

## Training, Validation, & Evaluation

Before presenting the description of machine learning regression models used in the project, I will cover the model training and evaluation process. In this project, k-fold cross validation has been used to split the dataset into train and validation sets. Furthermore, the final performance of regression models is evaluated using R2 error and negative RMSE values.

### K-Fold Cross-Validation

K-fold cross validation has been used to effectively train and validate the performance of regression models. This technique enhances the generalizability of the machine learning algorithms and reduces the chances of over-fitting. This technique conducts k-consecutive splits on the dataset to divide the dataset into () train and () validation sets. Further, the machine learning model is trained on splits and validated on the split. For effective generalizability of the proposed model, I set the value of k to 10.

### R-Squared Error

R2 is the statistical measure that assesses the performance of a regression model. Its value ranges between 0 and 1. It can be computed by subtracting the ratio from 1.0. The performance of a regression model can be considered good if the value of R2 is closer to 1.0 and vice versa.

### Negative Root Mean Square Error

The standard deviation of the prediction errors is measured by RMSE. It offers details on the sample saturation around the best fit line. Negative RMSE is the value of RMSE that is negative and can be calculated as:

|  |  |
| --- | --- |
|  | (3.1) |

The mean of the calculated result is indicated by the line above the squared variances between the predicted and actual value.

## Regression Models

Regression primarily seeks to estimate two factors: *1.* the influence of variables on the projected target variable and *2.* features important for performing the predictive task. In this project four state-of-the-art machine learning regression models such as KNN-R, SGD-R, LR, and ANN-R have been implemented for predictive purposes. This section covers the details of these regression models.

### K-Nearest Neighbors

KNN-R is the regression model that adopts “feature similarity” in order to carry out predictions of the unseen data. KNN computes the average of target value of the k nearest neighbors. It further assigns the target value to the input variables based on the difference of values between the computed average value and other nearest data instances. In this project I have set n\_neighbors to 7. This will enable the regression model to compute and compare the average predictive value with 7 neighboring data instances. The code used to define KNN-R is provided below:

|  |
| --- |
| def KNN(self, X, Y, total\_neighbors = 7):  model = KNeighborsRegressor(n\_neighbors= total\_neighbors) # model  scores = cross\_validate(model, X, Y, scoring=self.cross\_scoring, cv=self.cv, n\_jobs=-1) # rmse and r2  prediction = model\_selection.cross\_val\_predict(model, X, Y, cv=self.cv, n\_jobs=-1) # prediction values  return scores, prediction |

### Stochastic Grading Descent

SGD is a gradient descent variation that uses an iterative process to minimize the loss function. In SGD-R, the regression model works by selecting one data instance in random manner from the dataset for every iteration, hence reducing computing time. The parameters used for SGD-R in this project involve “Huber” as the loss function and 0.00000001 value as the initial learning rate for the model. By shifting between squared and linear loss beyond the range of epsilon, *huber* alters R2 to place minimal emphasis on correctly identifying exceptions. The code executed to define the SGD-R regression model is stated below:

|  |
| --- |
| def descent(self, X, Y, folds, n\_chunks):  gmodel = SGDRegressor(loss="huber",eta0=0.00000001) # model -- this loss rate gave us the best accuracy, huber gave us results as well  # stochastic model -- was used previously  """ scores = cross\_validate(gmodel, X, Y, scoring=self.cross\_scoring, cv=self.cv, n\_jobs=-1)  prediction = model\_selection.cross\_val\_predict(gmodel, X, Y, cv=self.cv, n\_jobs=-1) """  # mini batch  x\_folds = np.array\_split(X, folds)  y\_folds = np.array\_split(Y, folds)  scores = {}  r2 = []  rmse = []  predicted\_values = []  #print(X) |

### Linear Regression

LR is a regression model that assumes the mapping between features and target variable to be linear. It can be computed as:

|  |  |
| --- | --- |
|  | 3.2 |

Where, is the predicted target variable for given feature. is the predicted value of given the value of to be 0 and is the expected change in the value of y with the increase in the value of Finally, is the estimated residual value or error. The python code for LR is shown below:

|  |
| --- |
| def lRegression(self, X, Y):  model = LinearRegression()  scores = cross\_validate(model, X, Y, scoring=self.cross\_scoring, cv=self.cv, n\_jobs=-1)  prediction = model\_selection.cross\_val\_predict(model, X, Y, cv=self.cv, n\_jobs=-1)  return scores, prediction |

### Artificial Neural Network

An ANN is a deep learning model that replicates the neurons present in the human brain. It has three layers: the input layer, which is made up of neurons that introduce the input data into the paradigm, the hidden layer, are often more than one and assigns weight values to neurons as well as regulates them via activation function, and the output layer, which is mainly composed of neurons that produce final predictions. Three hidden layers make up the ANN-R model employed in this project.

|  |
| --- |
| def create\_Network(self): # creates network for cross\_validate and cross\_val\_predict  network = models.Sequential()  network.add(layers.Dense(128, activation='relu', input\_shape=(self.features,)))  network.add(layers.Dense(64, activation='relu'))  network.add(layers.Dense(1, activation='linear'))  network.compile(optimizer='adam', loss='mse', metrics=[tf.keras.metrics.RootMeanSquaredError(), "mae"])  return network    def doNeuralNetwork(self, X, Y):  neural\_network = KerasClassifier(build\_fn = self.create\_Network,  epochs=20,  batch\_size=500,  verbose=1)  results = cross\_validate(neural\_network, X, Y, scoring=self.cross\_scoring, cv=self.cv)  prediction = model\_selection.cross\_val\_predict(neural\_network, X, Y, cv=self.cv, n\_jobs=-1)  return results, prediction |

## Proposed Approach

This study focuses on providing the practical implication to show the impact of AI on marketing strategy. To deliver my Objective 2 presented in Section ‎1.2 I have devised a systematic model that can effectively predict the best performing promotional offer based on several features discussed in Section ‎3.1. To do so, I first processed the dataset to drop unnecessary information to minimize the computation overhead of the regression models. I only worked with nine features from the three files of the dataset. Afterward, 10-fold cross validation is applied to train and validate the regression models (). I used 10-fold cross validation since it reduces over-fitting and maximizes the generalizability of the model. Furthermore, the outcome of each split in the k-fold corresponding to each is then aggregated. Then, I added the resulting aggregated score to the corresponding offer ids. Afterward, I implement function to yield the offer id with the maximum score which is the best performing promotional offer. The graphical visualization of the proposed approach is displayed in Figure ‎3.4. The performance of is evaluated using two parameters including R2 and negative RMSE.

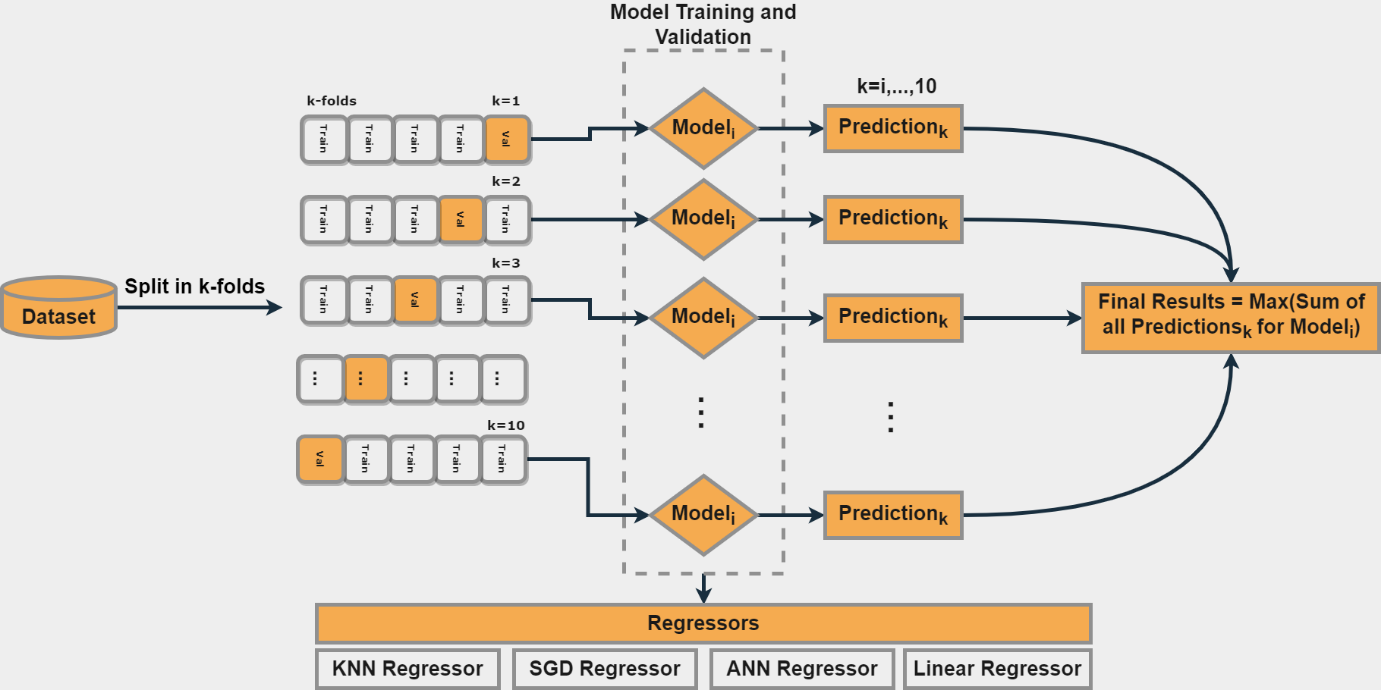


Figure ‎3.4-Architecture of the proposed approach

Implementation of the proposed approach in Python is given below:

|  |
| --- |
| def main():  datafile = 'newdata\_set.csv'  features = 9  # activate model with features  doML = Model(features)    # read in data  data\_set = pd.read\_csv(datafile) # read in data  data\_set = data\_set.sample(frac=1).reset\_index(drop=True) # shuffle  y\_list = data\_set.loc[:, 'amount'] # y = amount  X\_list = data\_set.drop(['amount', 'offer\_id'], axis=1) # x = rest of features aside from amount and offer id  # get values  Y = y\_list.values  X = np.array(X\_list)  # promotion id analytics -- save id  promotion\_id = []  promotion\_id.append("ae264e3637204a6fb9bb56bc8210ddfd")  promotion\_id.append("4d5c57ea9a6940dd891ad53e9dbe8da0")  promotion\_id.append("9b98b8c7a33c4b65b9aebfe6a799e6d9")  promotion\_id.append("0b1e1539f2cc45b7b9fa7c272da2e1d7")  promotion\_id.append("2298d6c36e964ae4a3e7e9706d1fb8c2")  promotion\_id.append("fafdcd668e3743c1bb461111dcafc2a4")  promotion\_id.append("f19421c1d4aa40978ebb69ca19b0e20d")  promotion\_id.append("2906b810c7d4411798c6938adc9daaa5")  promotion\_id.append("0")  # grab sums of each offer in transaction list  idsum = []  idsum.append(data\_set.loc[data\_set['offer\_id'] == "ae264e3637204a6fb9bb56bc8210ddfd", 'amount'].sum())  idsum.append(data\_set.loc[data\_set['offer\_id'] == "4d5c57ea9a6940dd891ad53e9dbe8da0", 'amount'].sum())  idsum.append(data\_set.loc[data\_set['offer\_id'] == "9b98b8c7a33c4b65b9aebfe6a799e6d9", 'amount'].sum())  idsum.append(data\_set.loc[data\_set['offer\_id'] == "0b1e1539f2cc45b7b9fa7c272da2e1d7", 'amount'].sum())  idsum.append(data\_set.loc[data\_set['offer\_id'] == "2298d6c36e964ae4a3e7e9706d1fb8c2", 'amount'].sum())  idsum.append(data\_set.loc[data\_set['offer\_id'] == "fafdcd668e3743c1bb461111dcafc2a4", 'amount'].sum())  idsum.append(data\_set.loc[data\_set['offer\_id'] == "f19421c1d4aa40978ebb69ca19b0e20d", 'amount'].sum())  idsum.append(data\_set.loc[data\_set['offer\_id'] == "2906b810c7d4411798c6938adc9daaa5", 'amount'].sum())  idsum.append(data\_set.loc[data\_set['offer\_id'] == "0", 'amount'].sum())  # grab indices of each ID post-shuffle  indices = []  index = data\_set.index  condition = data\_set["offer\_id"] == "ae264e3637204a6fb9bb56bc8210ddfd"  indices.append(index[condition])  condition = data\_set["offer\_id"] == "4d5c57ea9a6940dd891ad53e9dbe8da0"  indices.append(index[condition])  condition = data\_set["offer\_id"] == "9b98b8c7a33c4b65b9aebfe6a799e6d9"  indices.append(index[condition])  condition = data\_set["offer\_id"] == "0b1e1539f2cc45b7b9fa7c272da2e1d7"  indices.append(index[condition])  condition = data\_set["offer\_id"] == "2298d6c36e964ae4a3e7e9706d1fb8c2"  indices.append(index[condition])  condition = data\_set["offer\_id"] == "fafdcd668e3743c1bb461111dcafc2a4"  indices.append(index[condition])  condition = data\_set["offer\_id"] == "f19421c1d4aa40978ebb69ca19b0e20d"  indices.append(index[condition])  condition = data\_set["offer\_id"] == "2906b810c7d4411798c6938adc9daaa5"  indices.append(index[condition])  condition = data\_set["offer\_id"] == "0"  indices.append(index[condition])    for i in range(0,1): # repeat each experiment 3 times  print("Run # :"+str(i))  print("Doing KNN")  knn\_results, knn\_pred = doML.KNN(X, Y)  print("KNN done -- Doing LR")  LR\_results, LR\_pred = doML.lRegression(X, Y)  print("LR done -- Doing Gradient")  Gradient\_results, Gradient\_pred = doML.descent(X, Y, 10, 2600)  print("Gradient done -- Doing NN")  NN\_results, NN\_pred = doML.doNeuralNetwork(X, Y)  #calculate id sums  knn\_sums = []  LR\_sums = []  Gradient\_sums = []  NN\_sums = []  for i in range(0, len(indices)):  knn\_sums.append(knn\_pred[indices[i]].sum())  LR\_sums.append(LR\_pred[indices[i]].sum())  Gradient\_sums.append(Gradient\_pred[indices[i]].sum())  NN\_sums.append(NN\_pred[indices[i]].sum())  print("KNN results are: "+str(knn\_results))  print("Linear Regression results are: "+str(LR\_results))  print("Gradient results are: "+str(Gradient\_results))  print("NN results are: "+str(NN\_results))  print("Average of transactions = "+str(np.mean(idsum)))  print("Average of KNN = "+str(np.mean(knn\_sums)))  print("Average of Gradient = "+str(np.mean(Gradient\_sums)))  print("Average of LR = "+str(np.mean(LR\_sums)))  print("Average of NN = "+str(np.mean(NN\_sums)))  # show sum of predictions -- which offer type is most valuable?  print("Sums")  print("KNN Predictions")  for i in range(0, len(indices)):  print("Offer ID: " + str(promotion\_id[i])+ " sum = "+str(knn\_sums[i]))  print("Linear Regression Predictions")  for i in range(0, len(indices)):  print("Offer ID: " + str(promotion\_id[i])+ " sum = "+str(LR\_sums[i]))  print("Gradient Predictions")  for i in range(0, len(indices)):  print("Offer ID: " + str(promotion\_id[i])+ " sum = "+str(Gradient\_sums[i]))  print("NN Predictions")  for i in range(0, len(indices)):  print("Offer ID: " + str(promotion\_id[i])+ " sum = "+str(NN\_sums[i])) |

**CHAPTER 4**

# Results & Discussion

This chapter provides a thorough discussion of the experimental results. I conducted experiments with four regression models in this project. The performance of the models is evaluated in terms of R2 and negative RMSE. I will select the proposed regression model based on the optimum results. Table ‎4.1 presents the negative RMSE and R2 scores of regression models with respect to each of the 10 folds. I estimated the average negative RMSE and R2 scores of the regression models to provide a better understanding of the results. Since the closer the R2 is to one, the more efficient the model. The performance of the regression model also improves with higher negative RMSE scores.

Table ‎4.1-Experimental results of regression models in correspodance to each fold

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| k-folds | ***Negative RMSE Score*** | | | | ***R2 Score*** | | | |
| **KNN-R** | **SGD-R** | **LR** | **ANN-R** | **KNN-R** | **SGD-R** | **LR** | **ANN-R** |
| k=1 | -28.8664 | -25.091 | -24.9312 | -28.895 | -0.0622 | 0.07346 | 0.08522 | -0.2287 |
| k=2 | -33.8955 | -32.6871 | -32.5323 | -35.926 | -0.0196 | 0.051786 | 0.06074 | -0.145 |
| k=3 | -25.4294 | -23.3314 | -23.2449 | -27.271 | -0.0907 | 0.08184 | 0.08863 | -0.2543 |
| k=4 | -30.5272 | -29.092 | -28.9257 | -32.573 | -0.0359 | 0.059144 | 0.06986 | -0.1794 |
| k=5 | -30.4626 | -28.4985 | -28.36 | -31.985 | -0.07455 | 0.05954 | 0.068654 | -0.1846 |
| k=6 | -32.0216 | -30.0444 | -29.8654 | -33.502 | -0.07431 | 0.054265 | 0.065498 | -0.176 |
| k=7 | -35.844 | -34.2768 | -34.0948 | -37.539 | -0.04162 | 0.047468 | 0.05755 | -0.1424 |
| k=8 | -32.3378 | -30.9491 | -30.7387 | -34.265 | -0.03495 | 0.052028 | 0.064871 | -0.162 |
| k=9 | -30.9778 | -29.2319 | -29.04 | -32.767 | -0.05414 | 0.061326 | 0.073611 | -0.1794 |
| k=10 | -30.5059 | -28.7623 | -28.6349 | -32.182 | -0.05838 | 0.059143 | 0.067461 | -0.1778 |
| **Average** | **-31.0868** | **-29.1965** | **-29.0368** | **-32.6903** | **-0.05464** | **0.06** | **0.07021** | **-0.18295** |

The negative RMSE scores with respective to each fold indicate the superior performance of LR in terms of predicting the target variable. It yielded highest average negative RMSE score of -29.0368 consequently outperforming other regression models used in this project. LR considers linear mapping between the features and target variable. Since the underlying dataset used in this study is linearly separable, it provided sufficient predictive information to the LR regression model to carry out predictions. On the contrary, ANN-R is used to address the limitation of linear mapping in LR however, it did not perform well in the current project. Along with the poor performance of ANN-R, it takes more computational time making it inefficient for the underlying domain. Contrarily, SGD-R showed somewhat better performance with an average negative RMSE score of -29.1965 that is slightly lower than that of LR. The main advantage of SGD-R is that it adopts mini batch in its iteration to train the model and optimizes the performance of the working model. KNN-R is sensitive to the data size, thus showing poor performance in this project.

Figure ‎4.1-Negative RMSE scores of regression models with respect to each fold.

Similar trend is observed with R2 score of the regression models. LR yielded R2 score closest to 1 therefore, showing its superior performance over other regression models. Along the lines of this, SGD-R acquired 0.06 R2 score that shows it is also a better performing model however, it comes in second with LR remaining in the top. On the contrary, ANN-R and KNN-R showed poor performance with -0.18295 and -0.05464 R2 scores respectively. The negative value of R2 score depicts that the regression models were not able to predict the data instances accurately.

The negative RMSE and R2 scores are also shown in the form of bar charts in Figure ‎4.1 and Figure ‎4.2 respectively. It is evident that throughout the 10 splits, LR and SGD-R showed better performance as compared to ANN-R and KNN-R.

Figure ‎4.2-R-squared score of regression models corresponding to each fold.

**CHAPTER 5**

# Conclusion

This study is a two-fold as it attempts to focus on two primary objectives: theoretical review of impact of AI on marketing strategy, and practical implication to display the impact of AI on marketing strategy. To do so, I first conducted an extensive review of the studies that presented benefits, drawbacks, and challenges of integrating AI with marketing. Then, I advocated a systematic regression model to display the usage of AI in the marketing domain. This chapter will conclude the two objectives that were focused in this project.

The review study confirmed the high impact of AI in several marketing areas. Three technologies of AI such as decision-making systems, text, and image recognition have been widely used by the marketing strategists to exploit ways to maximize profits and enhance business-to-customer relationship. AI offers technologies to automate the product review and popularity analysis for better customer experience. The review of several existing studies indicated an unquestionable contribution of AI in the domain of marketing such as: real-time customer support, highly customised services, easier and efficient purchasing, and the flexibility of making an informed decision. More importantly, this demands for introducing the roles as well as competencies to advertising agencies, — in other words, those with the adequate information of AI, knowledge engineering, or credentials in the planning and execution of technical approaches. A paradigm shift in collaboration with the organisations providing cutting-edge artificial intelligence applications must be managed in order to achieve a symbiosis between AI as well as other activities. This study also weighed the risk associated with integrating AI into market such as threat to data privacy, confidentiality, computational inefficiencies, and especially ethics. Overall, the thorough review revealed that although AI has a considerable impact on marketing, it also has some pitfalls. Researchers must therefore consider data security, confidentiality, and integrity in marketing domain.

The practical implication of AI was provided by devising a machine learning-based model that uses regression algorithms to carry out predictive tasks. The model is designed to predict the best performing promotional offer with high efficacy that can be leveraged in the marketing domain for maximum profitability. The model is trained and validated using 10-fold cross-validation and further evaluated using R2 and negative RMSE scores. Several regression algorithms such as KNN-R, SGD-R, LR, and ANN-R were utilized in this project. Experimental results revealed LR to be the best performing model with -29.0368 RMSE score and 0.07 R2 score.

## Future Work Recommendations

This study conducted a review of “Impact of AI on marketing strategies” however, I consider using a limited number of studies in the review to be the limitation of this study. In future, I plan on investigating a systematic literature review study to analyze and investigate the impact of AI on marketing strategy.

Secondly, the proposed model is trained on highly imbalanced dataset. The machine learning algorithms have the tendency to learn the patterns regarding majority class consequently, producing ambiguous results related to the minority classes which are the promotional offers in this case. Addressing the imbalanced distribution of dataset can be another future direction as well. Moreover, limited number of machine learning regression models were investigated in this study, in future study, I will focus on integrating more effective algorithms in the model along with deep neural networks.

References

Annor Antwi, A. & Al-Dherasi, A. A. M., 2019. Application of Artificial Intelligence in Forecasting: A Systematic Review. *SSRN 3483313.*

Ansari, A., Li, Y. & Zhang, J. Z., 2018. Probabilistic topic model for hybrid recommender systems: A stochastic variational Bayesian approach. *Marketing Science,* 37(6), pp. 987--1008.

Armstrong, J. S., 1982. Strategies for implementing change: an experiential approach. *Group & Organization Management,* , 7(4), pp. 457-475.

Aulinas, J. & Sjafrie, H., 2021. AI for cars.

Barocas, S. & Selbst, A. D., 2016. Big data's disparate impact. *Calif. L. Rev.,* Volume 104, p. 671.

Betancourt, J. R., Green, A. R., Carrillo, J. E. & Park, E. R., 2005. Cultural Competence And Health Care Disparities: Key Perspectives And Trends. *Health Affairs,* , 24(2), pp. 499-505.

Blunt, L., . Types of Marketing Channels. *Chron,* , (), p. .

Br, K., 1987. Push and pull strategies: applications for health care marketing.. *Health care strategic management,* , 5(8), p. 13.

Bughin, J. et al., 2018. Notes from the AI frontier: Modeling the impact of AI on the world economy. *McKinsey Global Institute,* Volume 4.

Calvert, S. L., 2008. Children as Consumers: Advertising and Marketing. *The Future of Children,* , 18(1), pp. 205-234.

Carter, N. et al., 2010. The Role of Nursing Leadership in Integrating Clinical Nurse Specialists and Nurse Practitioners in Healthcare Delivery in Canada. *Nursing leadership,* , 23(), pp. 167-185.

Chinzei, K. et al., 2018. Regulatory science on AI-based medical devices and systems. *Advanced Biomedical Engineering},* Volume 7, pp. 118--123.

Chowdhary, K., 2020. Natural language processing. *Fundamentals of artificial intelligence,* pp. 603--649.

Citron, D. K. & Pasquale, F., 2014. The scored society: Due process for automated predictions. *Wash. L. Rev.,* Volume 89, p. 1.

Columbus, L., 2019. *10 Charts That Will Change Your Perspective Of AI In Marketing.* [Online]   
Available at: https://www.forbes.com/sites/louiscolumbus/2019/07/07/10-charts-that-will-change-your-perspective-of-ai-in-marketing/amp/  
[Accessed 24 11 2022].

Conklin, T. P., 2002. Health Care in the United States: An Evolving System. *Michigan Family Review,* , 07(1), p. 5.

Coye, M. J. & Kell, J., 2006. How Hospitals Confront New Technology. *Health Affairs,* , 25(1), pp. 163-173.

Cruz, Y. J. et al., 2020. Computer vision system for welding inspection of liquefied petroleum gas pressure vessels based on combined digital image processing and deep learning techniques. *Sensors,* 20(16), p. 4505.

Cui, D. & Curry, D., 2005. Prediction in marketing using the support vector machine. *Marketing Science,* 34(4), pp. 595--615.

Davenport, T. H., Dalle Mule, L. & Lacker, J., 2011. Know What Your Customers Want Before They Do-Advances in IT, data gathering, and analytics make it possible to create highly customized offers—Sometimes in milliseconds.. *Harvard Business Review,* p. 84.

Devers, K. J., Brewster, L. R. & Casalino, L. P., 2003. Changes in Hospital Competitive Strategy: A New Medical Arms Race?. *Health Services Research,* , 38(), pp. 447-469.

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

Engel, J. F., 1974. Understanding the Market For Christianity - a Report on Asian Christian Strategy. *ACR North American Advances,* , (), p. .

Evans, W. D., 2006. How social marketing works in health care. *BMJ,* , 332(7551), pp. 1207-1210.

Eyiokur, F. I., Ekenel, H. K. & Waibel, A., 2022. Unconstrained face mask and face-hand interaction datasets: building a computer vision system to help prevent the transmission of COVID-19. *Signal, Image and Video Processing,* pp. 1--8.

Gotto, A. M., 2007. Statin Therapy and the Elderly SAGE Advice. *Circulation,* , 115(6), pp. 681-683.

Grafanaki, S., 2016. Autonomy challenges in the age of big data. *Fordham Intell. Prop. Media & Ent. LJ,* Volume 27, p. 803.

Group, W. B., 2016. Ranking of economies – Doing Business – World Bank Group. *,* , (), p. .

Heidi, V., 2018. *You Weren’t Born in 1905? Why People Lie to Facebook.* [Online]   
Available at: https://www.wsj.com/articles/you-werent-born-in-1910-why-people-lie-to-facebook-1522682361  
[Accessed 24 11 2022].

Helbing, D., 2019. Societal, economic, ethical and legal challenges of the digital revolution: from big data to deep learning, artificial intelligence, and manipulative technologies. In: *Towards digital enlightenment.* s.l.:Springer, pp. 47--72.

Higgins, L. F., Rn, K. W. & Rn, K. A., 1998. Winning Market Positioning Strategies for Long Term Care Facilities. *Health Marketing Quarterly,* , 15(1), pp. 33-47.

Inglis, A., 2005. Quality Improvement, Quality Assurance, and Benchmarking: Comparing two frameworks for managing quality processes in open and distance learning. *The International Review of Research in Open and Distributed Learning,* , 6(1), p. .

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

Janka, S. F. & Uhsler, S. B., 2018. Antitrust 4.0-the rise of Artificial Intelligence and emerging challenges to antitrust law. *European Competition Law Review,* 39(3), pp. 112--123.

Jarek, K. & Mazurek, G., 2019. Marketing and Artificial Intelligence.. *Central European Business Review,* 8(2).

Johnson, J., 2019. Artificial intelligence \& future warfare: implications for international security. *Defense \& Security Analysis,* 35(2), pp. 147--169.

Kabene, S. M. et al., 2006. The Importance of Human Resources Management in Health Care: A Global Context. *Human Resources for Health,* , 4(1), pp. 20-20.

Kumar, T. S., 2020. Data mining based marketing decision support system using hybrid machine learning algorithm. *Journal of Artificial Intelligence,* 2(3), pp. 185--193.

Lai, A. W., 1995. Consumer Values, Product Benefits and Customer Value: a Consumption Behavior Approach. *ACR North American Advances,* , (), p. .

Ma, L. & Sun, B., 2020. Machine learning and AI in marketing--Connecting computing power to human insights. *International Journal of Research in Marketing,* 37(3), pp. 481--504.

Ma, L. & Sun, B., 2020. Machine learning and AI in marketing--Connecting computing power to human insights. *International Journal of Research in Marketing,* 37(7), pp. 481--504.

Marinchak, C. L. M., Forrest, E. & Hoanca, B., 2018. The impact of artificial intelligence and virtual personal assistants on marketing. *Encyclopedia of Information Science and Technology, Fourth Edition,* pp. 5748--5756.

Masys, D. R., 2002. Effects Of Current And Future Information Technologies On The Health Care Workforce. *Health Affairs,* , 21(5), pp. 33-41.

Michael, et al., 2018. *An analysis of more than 400 use cases across 19 industries and nine business functions highlights the broad use and significant economic potential of advanced AI techniques..* [Online]   
Available at: https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning  
[Accessed 24 11 2022].

Miller, K. E., 1975. A Situational Multi-Attribute Attitude Model. *ACR North American Advances,* , (), p. .

Milovanovic, B. et al., 2021. Colour assessment of milk and milk products using computer vision system and colorimeter. *International Dairy Journal,* Volume 120, p. 105084.

Mutolo, A. & Jl, M., 2011. The Influence of Social, Political, and Economic Factors on the Development and Form of Zulu Religious Activity in the 19th and 20th Centuries. *Inquiries Journal,* , 3(07), p. .

Nabipour, M. et al., 2020. Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis. *IEEE Access,* Volume 8, pp. 150199--150212.

Omachonu, V. K., 2010. Innovation in Healthcare Delivery Systems: A Conceptual Framework. *The Innovation Journal,* , 15(1), pp. 1-12.

Overgoor, G., Chica, M., Rand, W. & Weishampel, A., 2019. Letting the computers take over: Using AI to solve marketing problems. *California Management Review,* 61(4), pp. 156--185.

Pannucci, C. J. & Wilkins, E. G., 2010. Identifying and avoiding bias in research.. *Plastic and Reconstructive Surgery,* , 126(2), pp. 619-625.

Pantouvakis, A., Chlomoudis, C. I. & Dimas, A. G., 2010. Marketing Strategies in Port Industry: An Exploratory Study and a Research Agenda. *American Journal of Economics and Business Administration,* , 2(1), pp. 64-72.

Parekh, J., 2018. *Why Programmatic provides a better digital marketing landscape..* [Online]   
Available at: https://www.adweek.com/programmatic/why-programmatic-provides-a-better-digital-marketing-landscape/  
[Accessed 24 11 2022].

Pitta, D. A., Fung, H.-G. & Isberg, S. C., 1999. Ethical issues across cultures: Managing the differing perspectives of China and the USA.. *Journal of Consumer Marketing,* , 16(3), pp. 240-256.

Prakash, A., 2002. GREEN MARKETING, PUBLIC POLICY AND MANAGERIAL STRATEGIES. *Business Strategy and The Environment,* , 11(5), pp. 285-297.

Pryde, M. & Muller, M., 1995. A marketing strategy for a nursing college. *Curationis,* , 18(3), pp. 53-58.

Quest, L., Charrie, A. & Roy, S., 2018. The risks and benefits of using AI to detect crime. *Harv. Bus. Rev. Digit. Artic,* pp. 2--5.

Rauen, C. A., 2004. Simulation as a Teaching Strategy for Nursing Education and Orientation in Cardiac Surgery. *Critical Care Nurse,* , 24(3), pp. 46-51.

Saad, E. et al., 2021. Determining the efficiency of drugs under special conditions from users’ reviews on healthcare web forums. *IEEE Access,* Volume 9, pp. 85721--85737.

Sabar, L., 2020. *Starbucks Capstone Challenge.* [Online]   
Available at: https://towardsdatascience.com/starbucks-capstone-challenge-b95b0931bab4  
[Accessed 24 11 2022].

Salminen, J. et al., 2022. Creating and detecting fake reviews of online products. *Journal of Retailing and Consumer Services,* Volume 64, p. 102771.

Sankrusme, S., 2011. Marketing Strategy Competition among Beer Companies before Liquor Liberalization. *Journal of Marketing Development and Competitiveness,* , 5(6), pp. 65-82.

Shahid, M. Z. & Li, G., 2019. Impact of artificial intelligence in marketing: A perspective of marketing professionals of Pakistan. *Global Journal of Management and Business Research.*

Shapiro, L. G., Stockman, G. C. & others, 2001. Computer Vision. Volume 3.

Sharp, B., 1991. Competitive Marketing Strategy: Porter Revisited. *Marketing Intelligence & Planning,* , 9(1), pp. 4-10.

Shenton, A. K., 2004. Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information,* , 22(2), p. .

Siau, K. & Yang, Y., 2017. Impact of artificial intelligence, robotics, and machine learning on sales and marketing. Volume 48, pp. 18--19.

Soni, N., Sharma, E. K., Singh, N. & Kapoor, A., 2019. Impact of artificial intelligence on businesses: from research, innovation, market deployment to future shifts in business models. *arXiv preprint arXiv:1905.02092.*

Strzalkowski, T., Lin, F., Wang, J. & Perez-Carballo, J., 1999. Evaluating natural language processing techniques in information retrieval. In: *Natural language information retrieval.* s.l.:Springer, pp. 113--145.

Sun, S., Luo, C. & Chen, J., 2017. A review of natural language processing techniques for opinion mining systems. *Information fusion,* Volume 36, pp. 10--25.

Taneja, G., Girdhar, R. & Gupta, N., 2012. MARKETING STRATEGIES OF GLOBAL BRANDS IN INDIAN MARKETS. *Researchers World,* , 3(3), p. 71.

Tan, L. & Ong, K., 2002. The Impact of Medical Technology on Healthcare Today. *Hong Kong Journal of Emergency Medicine,* , 9(4), pp. 231-236.

Vladeck, D. C., 2015. Consumer protection in an era of big data analytics. *Ohio NUL Rev.,* Volume 42, p. 493.

Wajcman, J., 2017. Automation: is it really different this time?. *The British journal of sociology,* 68(1), pp. 119--127.

Wang, W. & Siau, K., 2019. Artificial intelligence, machine learning, automation, robotics, future of work and future of humanity: A review and research agenda. *Journal of Database Management (JDM),* 30(1), pp. 61--79.

Yang, Y. & Siau, K. L., 2018. A qualitative research on marketing and sales in the artificial intelligence age.

Zeng, Z., Shi, H., Wu, Y. & Hong, Z., 2015. Survey of natural language processing techniques in bioinformatics. *Computational and mathematical methods in medicine,* Volume 2015.

Zhang, M. & Luo, L., 2022. Can consumer-posted photos serve as a leading indicator of restaurant survival? Evidence from yelp. *Management Science.*

Zuboff, S., 2015. Big other: surveillance capitalism and the prospects of an information civilization. *Journal of information technology,* 30(1), pp. 75--89.

Ahangar, M.N., Ahmed, Q.Z., Khan, F.A. and Hafeez, M., 2021. A survey of autonomous vehicles: Enabling communication technologies and challenges. *Sensors*, *21*(3), p.706.

Pettigrew, S., Worrall, C., Talati, Z., Fritschi, L. and Norman, R., 2019. Dimensions of attitudes to autonomous vehicles. *Urban, Planning and Transport Research*.

Sarker, A., Shen, H., Rahman, M., Chowdhury, M., Dey, K., Li, F., Wang, Y. and Narman, H.S., 2019. A review of sensing and communication, human factors, and controller aspects for information-aware connected and automated vehicles. *IEEE transactions on intelligent transportation systems*, *21*(1), pp.7-29.

Jiang, Q., Zhang, N., Ni, J., Ma, J., Ma, X. and Choo, K.K.R., 2020. Unified biometric privacy preserving three-factor authentication and key agreement for cloud-assisted autonomous vehicles. *IEEE Transactions on Vehicular Technology*, *69*(9), pp.9390-9401.

Abou-Zeid, H., Pervez, F., Adinoyi, A., Aljlayl, M. and Yanikomeroglu, H., 2019. Cellular V2X transmission for connected and autonomous vehicles standardization, applications, and enabling technologies. *IEEE Consumer Electronics Magazine*, *8*(6), pp.91-98.

Guo, Q., Li, L. and Ban, X.J., 2019. Urban traffic signal control with connected and automated vehicles: A survey. *Transportation research part C: emerging technologies*, *101*, pp.313-334.

Khayatian, M., Mehrabian, M., Andert, E., Dedinsky, R., Choudhary, S., Lou, Y. and Shirvastava, A., 2020. A survey on intersection management of connected autonomous vehicles. *ACM Transactions on Cyber-Physical Systems*, *4*(4), pp.1-27.

Pisarov, J.L. and Mester, G., 2021. The use of autonomous vehicles in transportation. *Tehnika*, *76*(2), pp.171-177.

Alexakos, C., Katsini, C., Votis, K., Lalas, A., Tzovaras, D. and Serpanos, D., 2021. Enabling digital forensics readiness for internet of vehicles. *Transportation Research Procedia*, *52*, pp.339-346.

Manfreda, A., Ljubi, K. and Groznik, A., 2021. Autonomous vehicles in the smart city era: An empirical study of adoption factors important for millennials. *International Journal of Information Management*, *58*, p.102050.

Dixit, P. and Silakari, S., 2021. Deep learning algorithms for cybersecurity applications: A technological and status review. *Computer Science Review*, *39*, p.100317.

Wang, Y., Masoud, N. and Khojandi, A., 2020. Real-time sensor anomaly detection and recovery in connected automated vehicle sensors. *IEEE transactions on intelligent transportation systems*, *22*(3), pp.1411-1421.

# Appendix

## Data Processing

|  |
| --- |
| 1. import pandas as pd 2. import time 3. import numpy as np 4. start\_time = time.time() 5. print("Start Time: ", start\_time) 6. user\_data\_file = "./data/profile.json" 7. promo\_data\_file = "./data/portfolio.json" 8. transaction\_data\_file = "./data/transcript.json" 9. user\_data = pd.read\_json(user\_data\_file, lines=True) 10. promo\_data = pd.read\_json(promo\_data\_file, lines=True) 11. transaction\_data = pd.read\_json(transaction\_data\_file, lines=True) 12. transactions = pd.DataFrame() 13. offer\_completed = pd.DataFrame() 14. offer\_received = pd.DataFrame() 15. data\_set = pd.DataFrame() 16. # Get all 'transaction' type rows 17. transactions = transaction\_data.loc[transaction\_data['event'] == 'transaction'] 18. # Get all completed offer events 19. offer\_completed = transaction\_data.loc[transaction\_data['event'] == 'offer completed'] 20. transactions = transactions.append(offer\_completed) 21. transactions = transactions.sort\_index() 22. data\_set = transactions 23. # User Data 24. data\_set['gender'] = data\_set['person'].map(user\_data.set\_index('id')['gender']) 25. data\_set['user\_age'] = data\_set['person'].map(user\_data.set\_index('id')['age']) 26. data\_set['user\_income'] = data\_set['person'].map(user\_data.set\_index('id')['income']) 27. # Promo Data 28. data\_set['offer\_id'] = [d.get('offer\_id') for d in transactions.value] 29. data\_set['offer\_id'] = data\_set['offer\_id'].shift(periods=-1) 30. data\_set['amount'] = [d.get('amount') for d in transactions.value] 31. data\_set = data\_set[data\_set.event != 'offer completed'] 32. data\_set['offer\_type'] = data\_set['offer\_id'].map(promo\_data.set\_index('id')['offer\_type']).astype('category') 33. data\_set['offer\_diff'] = data\_set['offer\_id'].map(promo\_data.set\_index('id')['difficulty']) 34. data\_set['offer\_reward'] = data\_set['offer\_id'].map(promo\_data.set\_index('id')['reward']) 35. offer\_channels = data\_set['offer\_id'].map(promo\_data.set\_index('id')['channels']).apply(pd.Series) 36. data\_set['web'] = offer\_channels[0].apply(lambda i: i if i == 0 else 1) 37. data\_set['email'] = offer\_channels[1].apply(lambda i: i if i == 0 else 1) 38. data\_set['mobile'] = offer\_channels[2].apply(lambda i: i if i == 0 else 1) 39. data\_set['social'] = offer\_channels[3].apply(lambda i: i if i == 0 else 1) 40. del data\_set['person'] 41. del data\_set['value'] 42. del data\_set['event'] 43. data\_set = data\_set[['amount', 'gender', 'user\_income', 'offer\_type', 'offer\_diff', 44. 'offer\_reward', 'web', 'mobile', 'social', 'offer\_id']] 45. data\_set['gender'] = data\_set['gender'].replace([np.nan, 'M', 'F', 'O'], [0, 1, 2, 3]) 46. data\_set['offer\_type'] = data\_set['offer\_type'].replace([np.nan, 'bogo', 'discount', 'informational'], [0, 1, 2, 3]) 47. data\_set = data\_set.fillna(0) 48. data\_set.to\_csv('./data/data\_set.csv') 49. end\_time = time.time() - start\_time 50. print("Runtime: ", end\_time) 51. # EOF |

## Defining Regression Models

|  |
| --- |
| 1. from sklearn import model\_selection 2. from sklearn.neighbors import KNeighborsRegressor 3. from sklearn import metrics 4. from sklearn.model\_selection import GridSearchCV, KFold, cross\_validate, learning\_curve 5. from sklearn.linear\_model import LinearRegression, SGDRegressor 6. from keras import models, layers, optimizers 7. from keras.wrappers.scikit\_learn import KerasClassifier 8. import numpy as np 9. import pandas as pd 10. import tensorflow as tf 11. import matplotlib.pyplot as plt 12. class Model: 13. def \_\_init\_\_(self, n\_features): 14. self.features = n\_features # feature count 15. self.cv = KFold(n\_splits=10, shuffle=False) # shuffle = false because data is shuffled in program.py 16. self.cross\_scoring = ['neg\_root\_mean\_squared\_error', 'r2'] # return r2 and RSME (negative) 17. def KNN(self, X, Y, total\_neighbors = 7): 18. model = KNeighborsRegressor(n\_neighbors= total\_neighbors) # model 19. scores = cross\_validate(model, X, Y, scoring=self.cross\_scoring, cv=self.cv, n\_jobs=-1) # rmse and r2 20. prediction = model\_selection.cross\_val\_predict(model, X, Y, cv=self.cv, n\_jobs=-1) # prediction values 21. return scores, prediction 22. def find\_best\_N(self, x\_train, y\_train): # will output best N 23. params = {'n\_neighbors':[2,3,4,5,6,7,8,9]} 24. knn = KNeighborsRegressor() 25. model = GridSearchCV(knn, params) 26. model.fit(x\_train, y\_train) 27. return model.best\_params\_ 28. def descent(self, X, Y, folds, n\_chunks): 29. gmodel = SGDRegressor(loss="huber",eta0=0.00000001) # model -- this loss rate gave us the best accuracy, huber gave us results as well 30. # stochastic model -- was used previously 31. """ scores = cross\_validate(gmodel, X, Y, scoring=self.cross\_scoring, cv=self.cv, n\_jobs=-1) 32. prediction = model\_selection.cross\_val\_predict(gmodel, X, Y, cv=self.cv, n\_jobs=-1) """ 33. # mini batch 34. x\_folds = np.array\_split(X, folds) 35. y\_folds = np.array\_split(Y, folds) 36. scores = {} 37. r2 = [] 38. rmse = [] 39. predicted\_values = [] 40. #print(X) 41. for i in range(0, folds): 42. # split into test and training --> current i value will be the test set 43. list\_o\_nums = [] # this is used to combine the training set into one np array (it is currently an array of 9 arrays) 44. for k in range(0, folds): 45. if k != i: 46. list\_o\_nums.append(k) 47. y\_test = y\_folds[i] 48. x\_test = x\_folds[i] 49. # combine test data 50. for j in range(0, len(list\_o\_nums)): 51. if j == 0: 52. x\_train = x\_folds[list\_o\_nums[j]] 53. y\_train = y\_folds[list\_o\_nums[j]] 54. else: 55. x\_train = np.concatenate((x\_train, x\_folds[list\_o\_nums[j]])) 56. y\_train = np.concatenate((y\_train, y\_folds[list\_o\_nums[j]])) 57. # split training data into n\_chunks for mini-batch 58. x\_chunk = np.array\_split(x\_train, n\_chunks) 59. y\_chunk = np.array\_split(y\_train, n\_chunks) 60. # mini batch 61. for counter in range(0, n\_chunks): 62. gmodel.partial\_fit(x\_chunk[counter], y\_chunk[counter]) 63. #once done get values 64. y\_pred = gmodel.predict(x\_test) 65. rmse.append(metrics.mean\_squared\_error(y\_test, y\_pred, squared=False)) # root mean squared error 66. r2.append(metrics.r2\_score(y\_test, y\_pred)) 67. predicted\_values.append(y\_pred) 68. #plt.plot(x\_test, y\_pred, "r-") 69. #plt.plot(x\_test, y\_test, "b.") 70. #plt.show() 71. flat\_list = [item for sublist in predicted\_values for item in sublist] # predictions is list of lists --> turn into one list 72. prediction = np.array(flat\_list) # np array 73. scores["r2"] = r2 # save r2 in scores 74. scores["rmse"] =rmse # save rmse ins cores 75. return scores, prediction 76. def lRegression(self, X, Y): 77. model = LinearRegression() 78. scores = cross\_validate(model, X, Y, scoring=self.cross\_scoring, cv=self.cv, n\_jobs=-1) 79. prediction = model\_selection.cross\_val\_predict(model, X, Y, cv=self.cv, n\_jobs=-1) 80. return scores, prediction 81. def create\_Network(self): # creates network for cross\_validate and cross\_val\_predict 82. network = models.Sequential() 83. network.add(layers.Dense(128, activation='relu', input\_shape=(self.features,))) 84. network.add(layers.Dense(64, activation='relu')) 85. network.add(layers.Dense(1, activation='linear')) 86. network.compile(optimizer='adam', loss='mse', metrics=[tf.keras.metrics.RootMeanSquaredError(), "mae"]) 87. return network 89. def doNeuralNetwork(self, X, Y): 90. neural\_network = KerasClassifier(build\_fn = self.create\_Network, 91. epochs=20, 92. batch\_size=500, 93. verbose=1) 94. results = cross\_validate(neural\_network, X, Y, scoring=self.cross\_scoring, cv=self.cv) 95. prediction = model\_selection.cross\_val\_predict(neural\_network, X, Y, cv=self.cv, n\_jobs=-1) 96. return results, prediction |

## Performing Predictions

|  |
| --- |
| 1. from models import Model 2. import pandas as pd 3. import numpy as np 4. def main(): 5. datafile = 'newdata\_set.csv' 6. features = 9 7. # activate model with features 8. doML = Model(features) 9. # read in data 10. data\_set = pd.read\_csv(datafile) # read in data 11. data\_set = data\_set.sample(frac=1).reset\_index(drop=True) # shuffle 12. y\_list = data\_set.loc[:, 'amount'] # y = amount 13. X\_list = data\_set.drop(['amount', 'offer\_id'], axis=1) # x = rest of features aside from amount and offer id 14. # get values 15. Y = y\_list.values 16. X = np.array(X\_list) 17. # promotion id analytics -- save id 18. promotion\_id = [] 19. promotion\_id.append("ae264e3637204a6fb9bb56bc8210ddfd") 20. promotion\_id.append("4d5c57ea9a6940dd891ad53e9dbe8da0") 21. promotion\_id.append("9b98b8c7a33c4b65b9aebfe6a799e6d9") 22. promotion\_id.append("0b1e1539f2cc45b7b9fa7c272da2e1d7") 23. promotion\_id.append("2298d6c36e964ae4a3e7e9706d1fb8c2") 24. promotion\_id.append("fafdcd668e3743c1bb461111dcafc2a4") 25. promotion\_id.append("f19421c1d4aa40978ebb69ca19b0e20d") 26. promotion\_id.append("2906b810c7d4411798c6938adc9daaa5") 27. promotion\_id.append("0") 28. # grab sums of each offer in transaction list 29. idsum = [] 30. idsum.append(data\_set.loc[data\_set['offer\_id'] == "ae264e3637204a6fb9bb56bc8210ddfd", 'amount'].sum()) 31. idsum.append(data\_set.loc[data\_set['offer\_id'] == "4d5c57ea9a6940dd891ad53e9dbe8da0", 'amount'].sum()) 32. idsum.append(data\_set.loc[data\_set['offer\_id'] == "9b98b8c7a33c4b65b9aebfe6a799e6d9", 'amount'].sum()) 33. idsum.append(data\_set.loc[data\_set['offer\_id'] == "0b1e1539f2cc45b7b9fa7c272da2e1d7", 'amount'].sum()) 34. idsum.append(data\_set.loc[data\_set['offer\_id'] == "2298d6c36e964ae4a3e7e9706d1fb8c2", 'amount'].sum()) 35. idsum.append(data\_set.loc[data\_set['offer\_id'] == "fafdcd668e3743c1bb461111dcafc2a4", 'amount'].sum()) 36. idsum.append(data\_set.loc[data\_set['offer\_id'] == "f19421c1d4aa40978ebb69ca19b0e20d", 'amount'].sum()) 37. idsum.append(data\_set.loc[data\_set['offer\_id'] == "2906b810c7d4411798c6938adc9daaa5", 'amount'].sum()) 38. idsum.append(data\_set.loc[data\_set['offer\_id'] == "0", 'amount'].sum()) 39. # grab indices of each ID post-shuffle 40. indices = [] 41. index = data\_set.index 42. condition = data\_set["offer\_id"] == "ae264e3637204a6fb9bb56bc8210ddfd" 43. indices.append(index[condition]) 44. condition = data\_set["offer\_id"] == "4d5c57ea9a6940dd891ad53e9dbe8da0" 45. indices.append(index[condition]) 46. condition = data\_set["offer\_id"] == "9b98b8c7a33c4b65b9aebfe6a799e6d9" 47. indices.append(index[condition]) 48. condition = data\_set["offer\_id"] == "0b1e1539f2cc45b7b9fa7c272da2e1d7" 49. indices.append(index[condition]) 50. condition = data\_set["offer\_id"] == "2298d6c36e964ae4a3e7e9706d1fb8c2" 51. indices.append(index[condition]) 52. condition = data\_set["offer\_id"] == "fafdcd668e3743c1bb461111dcafc2a4" 53. indices.append(index[condition]) 54. condition = data\_set["offer\_id"] == "f19421c1d4aa40978ebb69ca19b0e20d" 55. indices.append(index[condition]) 56. condition = data\_set["offer\_id"] == "2906b810c7d4411798c6938adc9daaa5" 57. indices.append(index[condition]) 58. condition = data\_set["offer\_id"] == "0" 59. indices.append(index[condition]) 60. for i in range(0,1): # repeat each experiment 3 times 61. print("Run # :"+str(i)) 62. print("Doing KNN") 63. knn\_results, knn\_pred = doML.KNN(X, Y) 64. print("KNN done -- Doing LR") 65. LR\_results, LR\_pred = doML.lRegression(X, Y) 66. print("LR done -- Doing Gradient") 67. Gradient\_results, Gradient\_pred = doML.descent(X, Y, 10, 2600) 68. print("Gradient done -- Doing NN") 69. NN\_results, NN\_pred = doML.doNeuralNetwork(X, Y) 70. #calculate id sums 71. knn\_sums = [] 72. LR\_sums = [] 73. Gradient\_sums = [] 74. NN\_sums = [] 75. for i in range(0, len(indices)): 76. knn\_sums.append(knn\_pred[indices[i]].sum()) 77. LR\_sums.append(LR\_pred[indices[i]].sum()) 78. Gradient\_sums.append(Gradient\_pred[indices[i]].sum()) 79. NN\_sums.append(NN\_pred[indices[i]].sum()) 80. print("KNN results are: "+str(knn\_results)) 81. print("Linear Regression results are: "+str(LR\_results)) 82. print("Gradient results are: "+str(Gradient\_results)) 83. print("NN results are: "+str(NN\_results)) 84. print("Average of transactions = "+str(np.mean(idsum))) 85. print("Average of KNN = "+str(np.mean(knn\_sums))) 86. print("Average of Gradient = "+str(np.mean(Gradient\_sums))) 87. print("Average of LR = "+str(np.mean(LR\_sums))) 88. print("Average of NN = "+str(np.mean(NN\_sums))) 89. # show sum of predictions -- which offer type is most valuable? 90. print("Sums") 91. print("KNN Predictions") 92. for i in range(0, len(indices)): 93. print("Offer ID: " + str(promotion\_id[i])+ " sum = "+str(knn\_sums[i])) 94. print("Linear Regression Predictions") 95. for i in range(0, len(indices)): 96. print("Offer ID: " + str(promotion\_id[i])+ " sum = "+str(LR\_sums[i])) 97. print("Gradient Predictions") 98. for i in range(0, len(indices)): 99. print("Offer ID: " + str(promotion\_id[i])+ " sum = "+str(Gradient\_sums[i])) 100. print("NN Predictions") 101. for i in range(0, len(indices)): 102. print("Offer ID: " + str(promotion\_id[i])+ " sum = "+str(NN\_sums[i])) 103. if \_\_name\_\_ == "\_\_main\_\_": 104. main() |