UNIVERSITY OF HERTFORDSHIRE

School of Computer Science

M.sc Data science and Analytics

7COM1039- Advanced Computer Science Masters Project

Final Project Report

Topic

Analysing customer’s feedback Using Machine Learning (Sentiment Analysis)

Student’s Name: DEEPIKA MORA

Student’s ID: 19059957

Date:

Declaration

This report is submitted in partial fulfilment of the requirement for the degree of Master of Science in Data Science and Analytics at the University of Hertfordshire (UH).

It is my own work except where indicated in the report.

I did not use human participants in my MSc Project.

I hereby give permission for the report to be made available on the university website provided the source is acknowledged.

Table of Contents

Chapter 1

1.1 Introduction

According to Gerber (2015), successful business leaders have a knack for adapting to the ever-changing ways of conducting business in a global environment. Currently, business environments face numerous challenges such as changing customer needs and demands, growth in customer power, expanding global markets, and the blooming electronic markets on the internet. These difficulties require organisations to be more agile and adaptive and seek innovative ways to cope. Artificial intelligence technologies are at the core of responding to these business pressures and supporting decision-making. That said, artificial intelligence has not been fully leveraged in global decision support in organisations (Haenlein & Kaplan, 2019).

The exponential growth of the use of technology and its infrastructure and the dramatic increase in our capacity to store data have resulted in the rapid growth of big data, which is a precious asset. Big data is characterised by its vast volume, a variety of formats, high speeds of production with respect to the ability to process, and an extent of ambiguity and uncertainty (Chan, 2013). It is possible to gain new, valid, and useful insights that can shape the way decisions are made in organisations using artificial intelligence techniques such as machine learning, natural language processing, and text mining. Sentiment analysis, otherwise called opinion mining, is a useful technique yet to be fully leveraged. Machine learning employs algorithms that can be used to extract sentiments from data.

According to Zhang et al. (2018), sentiment analysis is an approach to natural language processing used to help identify emotional tone in data and determine whether it is positive, negative or neutral. A lexicon-based approach calculates the number of positive and negative sentiments related to the data analysed. Buyers and traders over the internet and social media leave a trail of unstructured information that can help businesses gain insights into product sentiments and monitor their brand from their customers' feedback. For instance, social media is a great avenue for pursuing opinion mining since they preserve a considerable deal of data given by online users. Social media allows access to any information about product service, place or event (Beigi et al., 2016).

This research study aims to discover the extent to which sentiment analysis has been applied in organisations and how it can be fully leveraged to help businesses better understand their customer needs to boost their profit margins. The paper will focus on social media comments and reviews on video games, specifically 'Call of Duty' using Twitter's application programming interface, Tweepy, to mine tweets reviewing video games.

1.2 Research Background

As aforementioned, a massive amount of unstructured data is left behind by users over the internet. User-contributed reviews are critical in helping businesses that sell goods and services online improve their marketing strategies and support decision-making. According to Jang et al. (2012), information searching is a vital step in a buyer's decision-making processing where they get to evaluate alternatives. The information search process entails searching for information about a product or service, paying attention to friends, colleagues, adverts and other alternative sources of information in their surroundings. Online shoppers tend to consult user reviews before making purchase decisions because of their apparent trust in their peers (Yogesh & Yesha, 2014).

It is possible for a human to manually inspect, extract, and classify sentiments in a review. Nonetheless, this process can be too costly and with low effectiveness and efficiency due to the rapidly growing volume of online reviews. Sentiment analysis provides a more effective way to gain useful insights into the unstructured users' information on the internet which, according to Rusu et al. (2013), accounts for ninety per cent of data produced globally. Sentiment analysis of online customer reviews encompasses a combination of text classification and text mining methods which allow extraction of information from user comments and sentiments of product reviews. Sentiments are ideally feelings or emotions the comment writer had towards a product, service or someone. On the other hand, an opinion is a sentence that contains the writer's opinion.

The research aims to find out how sentiment analysis can be fully leveraged to extract user opinions using video games as a case study. Video games have largely grown over the past recent years, especially due to the Covid-19 pandemic. The video game industry has grown to become a multibillion-dollar industry for a reason. These companies have put heavy investments to help them keep up with the growing demands and desire for their customers' quality. Video game developers are keen on captivating their fans with news stories and features on every franchise release. The franchise owners use video trailers and video snippets of the actual game to tease their fans. These trailers normally get a lot of criticism and hype from ardent followers.

For the game developers, these criticism or feedback is critical since it can allow them to ascertain whether the previews and trailers met their targeted customers' expectations. Additionally, social media comments are a great way to evaluate the success of a campaign. For instance, more positive reviews should mean that the customer reception was great and predict that the formal release is likely to become a huge success (Păvăloaia et al., 2019). Furthermore, the early release of these games allows users to make pre-orders out of enthusiasm and generate buzz that companies can take up and obtain valuable information through sentiment analysis.

According to Das and Kumar (2013), businesses globally largely derive their business intelligence from their well-structured data, which only accounts for twenty per cent of the data generated. Therefore, organisations must capture and process unstructured data for better results. The research will focus on finding an appropriate model to automate sentiment analysis in the gaming sector by using text mining techniques such as natural language processing on tweets. The research aims at discovering a way to fully leverage opinion mining and emotion detection in users' data.

1.3 Problem Statement

According to Ye and Nov (2013), the proliferation of user-contributed content on the internet has made it difficult for search engines to index everything and provide real-time searching capabilities. However, users still demand up-to-date information to help them perform various tasks and responsibilities better. Recently, the internet has become a great source for obtaining customer feedback, for instance, through business-to-consumer electronic commerce platforms. In addition to allowing users to explore catalogues and fulfil purchases, these platforms provide features that allow users to access other consumers' opinions. Furthermore, with online review systems' services like ServiceMagic, Epinions, and Yelp, users and consumers can leave and read reviews of products and services. Users can express their ideas and opinions in various ways, ranging from open text fields, lists of pros and cons, point scales, and polls (Cruz et al., 2010).

Most online customer review websites provide designated input areas for positive and negative ratings. As a result, the sentiment orientation is made clear, but the actual features and details that ought to be discussed and shared are left out (Rana & Singh, 2016). Additionally, the kind of rigour typically applied to the terminologies in professional papers such as corporate reports, news documents, or journal articles is not present in the text used in social media. The language used on social media platforms does not normally adhere to established linguistic and grammatical conventions. Comments on social media may lack whole sentences. Furthermore, many grammatical mistakes and words do not exist in dictionaries (Baldwin et al., 2013). Term-based approaches to sentiment analysis may not be sufficient or helpful for accurate sentiment classification because they require advanced natural language processing tools to analyse and discern the users' intentions in opinion sentences.

Moreover, many new terms such as acronyms, slang, and emoticons exist in the analysis of social media data. These terms include phrases that are not official and name entities like emojis. These factors make it challenging for businesses to extract insights from their users' opinions effectively. This research project proposes to develop a strategy for effectively classifying evaluations detailed in social media posts using machine learning and text mining. The onset of the Covid-19 pandemic resulted in a dramatic growth in e-commerce and working from home. This global growth in computer users needs to be fully leveraged by businesses. This study will endeavour to understand what sentiment analysis is, the extent and has been applied by organisations, and how it can be implemented to effectively detect opinions and emotions from user feedback to support marketing decisions.

1.4 Proposed Solution

Companies need to use sentiment analysis to gain important information and insights to improve their customer service and satisfaction. Using sentiment scores is a great method that can be used to evaluate sentiments. Sentiment scores are scale ratings that reflect the emotional intensity of the feelings expressed in a comment. A lexicon approach to sentiment analysis is proposed for the sentiment analysis of Twitter social media data for gaming insights. A lexicon approach entails using a pre-prepared sentiment lexicon where a document is scored by obtaining an aggregate of the sentiment scores of all the words in the text (Gitari et al., 2015). The Twitter data is labelled to allow for supervised learning.

1.5 Project Rationale and Significance

The research study aims to find out how sentiment analysis can be applied to customers' feedback to aid organisations in discerning their emotions in their marketing campaigns. Analysing customers' feedback is a necessary tool when it comes to customer loyalty, product acceptance, customer experience and satisfaction, advertising, and marketing promotions campaigns (Gallagher et al., 2019). Computer infrastructure has grown immensely, and processing and storage capabilities have drastically escalated. Additionally, powerful artificial intelligence techniques and tools can be used to analyse big data for useful insights. Brands have increasingly become interested in learning about their customers' pain points. There is also a sharp increase in chatbots, which offer a powerful and engaging way to interact with customers using computers.

All these avenues call for a need for a proper way to obtain novel information useful for decision-making (Chen et al., 2021).

1.6 Research Questions

The following are the research questions that will be used to guide the research study and narrow down the research scope.

i. To what extent has sentiment analysis been taken up by business entities, and how has it been used?
ii. What are the sources of information that can be used for sentiment analysis?
iii. What are the ways and techniques in which sentiment analysis can be used to analyse customer feedback effectively?
iv. What constituents of customer purchase behaviour, and how can it be capitalised to improve businesses' bottom line?
v. What machine learning techniques, technologies, and tools can be used to leverage opinions from customers' feedback?

1.7 Research Aim

This research study aims to discover and discern the role of sentiment analysis in supporting decision-making and the machine learning techniques and tools that organisations globally can use to extract useful sentiments, attitudes, and emotions from their customer's feedback data.

1.8 Research Objectives

The following are the report objectives:

- To find out and understand what sentiment analysis is.
- To find out machine learning and other artificial intelligence techniques that can be used to undertake sentiment analysis effectively.
- To find out the extent to which sentiment analysis has been used by organisation entities and how it is used.
- To discover and understand customer purchase behaviour and how it can be leveraged to precisely make decisions on marketing and other strategic decisions.
- To design and develop a classification model using appropriate machine learning algorithms to categorise sentiments from tweet data.

Chapter 2

2.0 Literature Review

2.1 Sentiment Analysis

According to Dang et al. (2020), sentiment analysis is a technique used in natural language processing that involves finding out people's opinions, attitudes, and emotions towards a particular entity. NLP is a subfield in linguistics that is concerned with the manner in which computers interact with human language (Chowdhary, 2020). Computers are programmed to process and analyse large sets of natural language data. Opinion mining emphasises extracting people's opinion towards an entity whereas sentiment analyses focus on identifying sentiments expressed on opinions and classifying their polarity. Polarity levels include positive, neutral, or negative. Suppose the polarity precision of an organisation is important. In that case, it is critical to consider expanding the polarity categories to more levels of positive and negative, for instance, very positive or very negative (Guerini et al., 2013).

Sentiment analysis is categorised into document level, sentence level, and aspect level. The document level spans a whole and aims to classify it as either expressing positive or negative sentiments (Behdenna et al., 2018). Sentence-level sentence analysis classifies sentiments in a sentence. It is important first to determine whether a sentence is subjective or objective. If it is subjective, the sentence-level sentiment analysis finds out whether it expresses positive or negative opinions (Arulmurugan et al., 2019). The aspect level sentiment analysis is the most effective level. At this level, sentiment analysis aims at classifying the sentiments with regard to the specific aspects of the entities. The two other levels do not offer sufficient detail in the needed opinions on all aspects of the entity that is much needed by applications.

According to Raghuvanshi and Patil (2016), there are four most common types of sentiment analysis, fine-grained, emotion detection, multilingual sentiment analysis, and aspect-based sentiment analysis. Fine-grained entails using a scale to categorise comments and statements. It is applied on the sub-sentence level to identify a target topic of a statement. Fine-grained sentiment analysis breaks down phrases or clauses, and the resultant parts are analysed in connection to the other (Guzman & Maalej, 2014). Emotion detection opinion mining aims to extract specific emotions such as happiness, irritation or anger from text data. Emotion detection software and tools normally use lexicons or a list of words that denote specific emotions. According to Garcia-Garcia et al. (2017), this procedure can be inefficient as some words may convey several different meanings. For instance, the word kill might represent sadness but may also represent happiness or approval. For example, "You're killing it at your job" conveys approval.

Emotion detection is made possible by the use of emotion models, which are its building block and define how emotions should be defined. According to PS and Mahalakshmi (2017), when undertaking an emotion detection analysis, it is essential to define the model of emotion to use. Emotion models are divided into two distinct groups, discrete emotion models and dimensional emotional models. Discrete models involve placing emotions into clear-cut classes. According to Scherer and Ekman (2014), prominent categories include the Paul Ekman model, which distinguishes emotions based on six classes. This model asserts a theory that six fundamental emotions originate from the neural system as a result of how the experiencer perceives the situation. They include sadness, disgust, surprise, happiness, anger, and fear. This model is supported by other models such as Robert Plutchik and Orthony, Clore, and Collins (OCC) model.

On the other hand, multilingual analysis involves performing sentiment analysis to obtain sentiments from different languages. According to Dashtipour et al. (2016), this type of sentiment analysis is often difficult, as it involves numerous pre-processing and resources. It is more efficient when machine translation can pick up cultural subtleties, colloquialisms and other references within their context in social comments and online reviews. It is necessary for organisations with an international customer base to adopt sentiment analysis for more than just the English language. According to Lo et al. (2017), English accounts for only twenty-six per cent of internet users. The diversity in languages and cultures globally greatly influences social analytics; thus, sentiment analysis in only the English language is not sufficient.

2.2 Applications of Sentiment Analysis

Sentiment analysis has been applied considerably by organisations and research projects. Organisations use sentiment analysis to analyse qualitative data they gather from various channels to improve their service delivery. Sentiment analysis is used in Voice of the customer programs (VoC), where customer feedback data is collected, organised, and shared internally across the organisation (Jaworski & Kohli, 2014). Customer feedback is the most reliable way to understand how a product is performing, and the bottle necks stalling its growth. Using sentiment analysis and opinion mining, organisations can understand the rationale behind customer feedback, their pain points, or what they need and want, and provide synthesised feedback that can be shared in a way to improve and strengthen different areas of the business (McColl-Kennedy et al., 2019).

Net promoter scores surveys are a good way to evaluate customer feelings (Zaki et al., 2016). These surveys can help organisations have better customer retention in good performances. Net promoter scores are an example of Voice of customer tools. Feedback collected from such grow exponentially and becomes profoundly difficult to make sense of. Sentiment analysis can help organisations quickly identify reported negative experiences and rectify the issues before it affects their profit margins. Moreover, sentiment analysis can be used to improve the enhance the customer experience of users. According to Nobar and Rostamzadeh (2018), customers need to know that their queries and problems are dealt with haste, efficiently, and professionally.

Customer service and experience are critical for any organisation's success. It is possible to analyse customer data, find comments or posts that show that customers are particularly frustrated, and rectify them using sentiment analysis. Sentiment analysis allows for a priority based solving of issues which ultimately reduces processing times and increases efficiencies by ensuring queries are direct to the responsible entities. Shirdastian et al. (2019) looked into why people have positive and negative sentiments about the authenticity of a brand. They used a database of 2,204 coded tweets for the sentiment analysis. They developed a quantitative model for predicting brand attributes and authenticity with their corresponding emotional polarity. They also classified tweets according to their quality and categorised them based on their historical significance and originality. Furthermore, the terms often sued in each category were discovered through the application of latent semantic analysis (LSA). As a result, the brand authenticity dimensions and the sentiment polarity associated with such phrases can be accurately predicted.

Nemes and Kiss (2021) used a Recurrent Neural Network and Natural Language Processing to look into how Twitter users felt and acted based on the most common trends; in this case, the word 'covid' and words related to the Covid-19 pandemic were the most popular keywords. Their trained model performed relatively better when determining the polarity of 'contemporary', which are often confusing tweets. The model reduced the margin of error because it received further detailed instructions. Strelluf (2019) looked into the steps of finding data, gathering its, and training to solve challenges encountered during sentiment analysis data discovery.

When gathering data for analysis, it is necessary first to determine how much data is expected and which part of the research is the most important. Thirdly, it is essential to build an infrastructure to handle the amount and format of data and figure out how to get structured information out of unstructured data. Liang et al. (2015) look at how customer reviews are important and how numerous online shoppers rely on product reviews to identify commodities that suit their preferences. Parsing and analysing sentences and comments currently use natural language processing. Writing styles used in internet reviews are informal; thus, some grammatical and semantic errors in opinion expressions may not be covered in any dictionary. Class association rules and Naïve Bayes Classifiers are used to arrange opinion statements into relevant product feature classes and create a summary of customer testimonials.

Priyantina and Sarno (2019), in their research, came up with the idea of a system that uses sentiment analysis to sort reviews that customers leave for hotels. Their study describes a process that can be used to get a semantically relevant domain-specific lexicon of terms from a given corpus. The result of the study was creating a lexicon which helped the sentiment analysis used to classify the reviews. The new system performed better than the baseline that had been set. Li et al. (2021) conducted a study on sentiment information included in customer reviews and looked at how it could be used to improve forecasting of hotel demand. In the study, four h0gh-end hotels were chosen, and their reviews from two popular websites were used. The Short-Term Long model, a form of deep learning was used to get information about how the customers felt. Three sentiment indices were made and looked at the bullish, average, and variance indexes. An autoregressive integrated moving average with exogenous variables was used to find out more about how useful the sentiment indices were. The results of this study show that a piece of information is “actionable” if a company can use it to improve its product or service.

Data analytics and machine learning algorithms can monitor social media and recognise customers’ perspectives on luxury hotels from beginning to end (Giglio et al., 2020). Their study suggests that managers of comfort hotels can use the new visual data analysis to develop a more effective brand management strategy. Based on social media, they created new features and established a machine learning model that can anticipate how active each company is on social media. They used Twitter data, and their results concluded that deep learning was the most accurate method for predicting the level of involvement and showed that the number of likes and retweets a business tweet gets is critical to tell whether their marketing strategy is working or not.

2.3 Consumer decision process & Its relation to Sentiment Analysis

Understanding consumer behaviour is important for any business that wants to succeed. According to Schiffman et al. (2013), consumer behaviour is an analysis of how customers make their decisions about purchases, what motivates them, and their expectations for products. Consumer behaviour is a crucial aspect of marketing concepts. Organisations must develop a portfolio and marketing mix that satisfies customers' needs and wants, given their changes in behaviour and patterns (Gull & Pervaiz, 2018). Gaining an understanding of factors and issues that affect consumer behaviour makes marketing strategies more precise, and predictions on how customers may respond to marketing strategies can be made easier. Furthermore, understanding consumer behaviour reduces the risks involved in launching new services or products.

According to Sarathy and Patro (2013), buying decision behaviour can be categorised into major classes, high involvement and low involvement. The purchase decision and process vary depending on the item to be purchased. High involvement behaviour is categorised into complex buying and dissonance-reducing behaviour. Complex buying behaviour entails a situation where there is a high consumer involvement with significant brand differences. The buyer here is aware of significant differences in brands. It mainly occurs in situations where a product is expensive, infrequently bought, or highly expressive (Delafrooz et al., 2014). Organisations can use sentiment analysis and text mining to find out what information their consumers need before making a purchase decision and use this to ensure they avail concise and appropriate information.

Dissonance reducing buyer behaviour is also a high involvement purchase behaviour, but the brands are quite similar with few perceived differences. The customer notices few differences but takes their time since the product might be expensive and infrequently bought (Sharma, 2014). However, the consumer ends up quickly making the purchase by responding to the best price offered, after which they might experience post-purchase dissonance when they notice a perk or hear information about a product not purchased. Sentiment analysis can be used to augment their belief in their purchase by ensuring all feedback is well analysed for better reviews on comments in future.

Low involvement behaviour includes habitual buying behaviour where there are few perceived differences and little loyalty involved. The buyer decision-making process comprises five main stages: problem recognition, information searching, evaluation of alternatives, purchase decision, and post-purchase behaviour (Puspitasari et al., 2018). After recognising a need, a consumer will search for information about a product depending on his drive, the value placed on additional information, and the satisfaction they will receive at the end of the search. Sentiment analysis is applied to ensure an organisation performs well to capture customers as they evaluate alternatives.

Chapter 3

3.1 Research Methodology

The research study aims to use a quantitative methodology to conduct the research. The study will entail the collection of secondary data already available from the Twitter application programming interface. The study's research design will be a non-experimental observational design where data will be collected from online sources. The research will take a stance using a pragmatic school of thought. Pragmatism is an epistemology based on the notion that truth is the final outcome of an ideal scientific closely related to usefulness. According to Russ (2014), epistemology is the study of knowledge and how people perceive different aspects of the world as they try to understand it. The research-based its ontology on the current state of reality or truth, that sentiment analysis has not been fully leveraged in organisations globally.

The study will use comments and reviews from Twitter regarding the 'Call of duty' promotional video and social campaign on the platform. Since data will have noise and errors, a data cleaning process and preparation will be conducted to ensure the data is clean and high-quality. Python libraries will be used to manipulate the data, and descriptive statistics for categorical data will be computed to draw insights. Furthermore, bar charts and word clouds will be used to display common words and the distribution of the sentiments. The data will be analysed using a neural network-based model that will run on the Keras application programming interface to classify fan reviews into three categories of sentiments, negative, positive, and neutral. Furthermore, eighty per cent of the data will be used in training while the remaining twenty for testing. The established mathematical model will be built using tensor flow to predict the sentiments of customers' feedback from the Twitter social platform.

An application programming interface is required to obtain tweets from Twitter. According to Ofoeda et al. (2019), an API is a software that allows two applications to share information without needing user intervention. The research will incorporate Tweepy, which is a Twitter API for the study. The API will require tokens and keys from the users that will be used as credentials. Pre-processing of the tweets will be done in addition to natural language processing. Text cleaning will be used to make processing and prepare the data for easier analysis. Stop words which are regarded as noise will be removed as part of the text data preparation process Vijayarani et al. (2015). Additionally, the text will be further normalised using lemmatisation and stemming. Text visualisation tools will be used to gain insights into various sentiments, where a word cloud will be used to display most of the commonly used words in tweets. Furthermore, Matplotlib and Seaborn libraries will be used to create charts, histograms, boxplots, and line graphs.

Additionally, the study will employ the CRISP-DM methodology to help fulfil the research objectives. CRISP-DM stands for Cross Industry Standards Process for data mining and is a model used to execute data science projects (Huber et al., 2019). This model provides a robust and well-proven methodology that is powerful, practical and has great flexibility and usefulness when used to solve tough business issues using data analytics. The model will follow a sequence of events or steps (Martínez-Plumed et al., 2019). However, in practice, many tasks can be done in different sequences, but it is always necessary to backtrack to certain previous tasks and repeat some of them. Additionally, this model does not try to capture all possible routes through the use of data mining procedures.

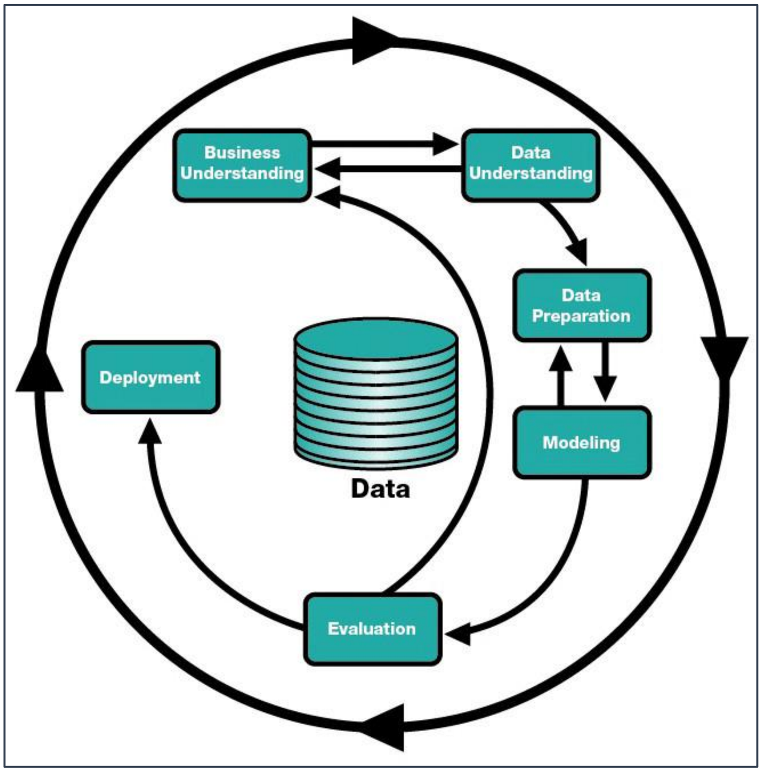

Figure 1 Crisp-dm methodology

The CRISP-DM methodology will encompass six sequential phases: business understanding, data understanding, data preparation, modelling, evaluation, and deployment (Schröer et al., 2021). The first stage of business understanding will entail understanding the project's objectives and requirements. This stage involves determining business objectives and what is the desired accomplishment of the project undertaking. The availability of resources required for the project and an assessment of risk and contingencies help in carrying out the cost-benefit analysis. It will be critical also to define what success looks like from a technical point of view. For instance, the success criteria of this project will depend on the model's ability to achieve above an eighty per cent accuracy. For the project to be put into production, it is necessary for it to make economic sense; therefore, the model accuracy should supersede the human ability to earn sufficient justification to be put into production.

The next stage is data understanding which involves adding to the foundation of the business understanding process by identifying and collecting necessary data sets that will help fulfil the project goals. This stage entails collecting the initial data, examining the data to gain an understanding of its properties, such as format, and digging deeper into the data by processing, querying, and finding relationships in it. Identifying the quality of data and how clean the data is is a necessary procedure in the CRISP-DM model. The third stage in the model is the data preparation stage, also called data munging. It entails selecting the perfect datasets and their reasons for inclusions and exclusion, followed by cleaning the data. New attributes can be derived from the data to increase its usefulness (Shafique & Qaiser, 2014). The data is then integrated by combining it from multiple sources, after which it is reformatted as necessary.

The modelling phase is the fourth stage of the CRISP-DM model, which will be critical in this study. In this stage, various models are developed using different modelling techniques. In this stage, it will be important to choose the best modelling techniques and algorithms to use, for instance, neural networks or regression. Before building the model, the generation of test designs where data may need to be split into validation, training, and test tests. It is important to iterate the model-building process until the best possible model is obtained. The fifth stage entails evaluating the model to find out whether it meets the business needs to be determined in phase one. Three key tasks should be done. These tasks include evaluating the results of the model by ascertaining whether the models meet the business criteria. Additionally, the work should be reviewed to determine whether any aspects were overlooked and depending on how successful the evaluation is, it is determined whether to proceed to deployment. The final stage involves deployment, which is done depending on the requirements.

As aforementioned, the study will use secondary data that Twitter will provide through their Tweepy application programming interface. It will be possible to query data using the API depending on its recency, for instance, the most recent tweets, which should make it possible to view tweets from the past seven days. The search query for the purpose of the research will be ‘Call of Duty’, which will be used as a case study whilst conducting the research. The main reasoning technique used in the study will be a deductive mental process in which the research question will be broken down into sub-questions which will ultimately help in obtaining the solution.

The success criteria of this project will depend on the model's ability to achieve above 80 percent accuracy. If the project is to be put into production, it would need to make economic sense. Hence, the model accuracy should supersede human ability so that it can earn sufficient justification to be put into production. Feedback from this project stands to benefit video game developers as it provides feedback on what users enjoyed and did not enjoy in their products. Feedback is essential, especially the improvements of future iterations, as developers would narrow down what users want or dislike.

3.4 Data source

Some social media services, such as Facebook and Twitter, made publicly collected data available through their Application Programming Interface (API). Tweepy provided statistical data such as user profiles, whereas Twitter REST API provided information such as tweets (Client — tweepy 4.10.0 documentation, 2022). Snapchat and Facebook have made Tencent API5 and Facebook Graph API4 available. These APIs also assisted in obtaining posts and other data from their site, and they have been thoroughly investigated.

As a prerequisite, the Twepy library must be installed with 'pip3 install Tweepy' and imported with 'import Tweepy" company. "Tweepy" has a feature called "recent search tweets." It can be used to find Tweets that were posted within the last seven days (Client — tweepy 4.10.0 documentation, 2022). The search query was the "game's video Twitter username," for instance, ‘CallOfDuty,' and so on. Each Tweet appears in your response, along with its "Tweet ID" and Text. By default, the response to the request retrieves ten tweets. For more than 10 Tweets per request, the "max results" parameter request is used to specify the number of Tweets to receive. This study set the 'max results' at 10,000 tweets.

The tweets were mined on Thursday 8th June 2022. During that time 'Call of Duty" had realised a trailer for Modern Warfare II and a had generated a lot of excitement among the fans of the Franchise. The trailer was also a campaign tool to urge followers make pre-orders for the beta version. In less than a week the post had attracted 4 million views and over two thousand replies. The subsequent posts were snippets of the video game just to keep momentum of the campaign. In addition, Thursday June 9June 9 2022 marked the start of the Summer Game Fest. Summer Game Fest began in 2020 to fill the void left by a cancelled E3 due to the pandemic and global lockdowns (Henderson, 2022). It ran from May to August and included publisher announcements, in-game events, and an opening livestream jam-packed with game information and trailers. The main event of Summer Game Fest began on Thursday with a live kick-off show. Therefore, it would be interesting to find out what fans of the 'Call of Duty franchise had to say about them realise. The feedback from this event would tell whether the campaign was a success through the number of positive reviews.

3.2 Legal and Ethical Consideration

The study generally did not involve collecting and storing personal data; hence, the study will not violate data privacy laws. The sole purpose of the data will be to train a machine learning algorithm to recognise and classify sentiments. Additionally, the data to be used in the study will be collected from public forums and will not involve personal information on individuals. The research will adhere to data privacy laws to access data from twitter which regulate access and consumption of data, ensuring access to twitter data is organised. Moreover, the participants have no risks of financial or legal liabilities since the data is readily available in public forums.

Chapter 4

Data analysis/Implementation

4.1 Chapter overview

This chapter delves into the implementation of the study objectives. The report starts by undertaking an understanding of the data used in the study. The research design adopted follows the CRISP-DM methodology, so you should expect to find a significant portion of this chapter dedicated to implementing it. The comprehension of data is essential because it enables one to better understand the data and determine the types of data that need to be collected (Smart Vision Europe, n.d.). Exploring the data, comprehending the variables and the relationships between them, and locating any issues that may exist with the data are all components of understanding the data. Other critical steps that will be covered in the chapter include the data preparation process and modeling. Data preparation should involve data cleaning, text pre-processing, feature engineering, data labeling, and extracting insights. The modeling part will involve setting model parameters, model training, and evaluation.

4.2 About the dataset

The data set used in the study was mined from Twitter between the 8th and 9th of June, 2022. The restriction on the timelines arose by default due to the capping of max results (maximum number of tweets). The data collection dates coincided with the launch of Summer Game Fest that kicked off on Thursday 9th. The Tweepy API uses key search words to extract tweets and for our case the key search word was ‘Call of Duty’. As a result, the tweets revolved around “Call of Duty “, a series of first-person shooter games made by Activision.

4.2.1 Dataset description

The dataset obtained using the Tweepy API included 10,000 tweets with three dimensions; 'Text,' 'location,' and 'Date.’ The data types of the features include; 'Text': objects, 'location': object, 'Date' and datetime64[ns]. The ‘Text’ attribute contained the actual tweets shared on the Twitter platform. The “Text” was stored in “object” form. ‘Location’ attribute gives details as to where the tweet was generated from. The locations as well were stored as ‘object’. The Tweepy API used countries and cities were used to determine location. The final attribute was “date”; the attribute gave details on the date and time the tweet was posted. ‘Date’ data was stored in ‘object’ for although later converted to the relevant datatype. The dataset had a total of 6,269 locations. The number of unique items in every variable was as follows; Text: 3,827, Location: 3,180, and Date: 8,447. The Location had a total of 3,731 missing locations.

4.3 Resources and Libraries used

An Application Programming Interface (API) is required to obtain tweets from Twitter. In addition, pre-processing of tweets would necessitate the use of Natural Language Processing Tools. Text cleaning technology will make processing and preparing data for analysis easier. Text visualization tools would be required to gain insights into the various sentiments. Additionally, a lexicon-based tool was used to label the tweets. Valence Aware Dictionary and Sentiment Reasoner (VADER) instead of humans was used to label the tweets. The study employs the use of a word cloud to display the most commonly used words in tweets. For better context, the regularly used terms would be displayed depending on their sentiments; positive, negative, and neutral. To create charts, histograms, boxplots, and line graphs, Matplotlib and Seaborn libraries have been used. Data visualization improves the understanding and utilization of data. The study also utilized a neural network to build the sentiment classifier. A neural network based on TensorFlow which runs on Keras API will be used to train the model.

4.4 Data pre-processing

4.4.1 Data cleaning

The dataset for this study had missing values for the location variables and was replaced by imputing "unknown" instead of deleting the records. Of the 10,000 records, only 3,826 records of the variable 'Text' were unique. In addition, there were only 1,238 locations and 3,614 unique dates. Consequently, the data was all the duplicate values for text were deleted while retaining the first value in the list of duplicates. At the end of the data cleaning, we remained with 3,826 records.

According to Chu et al. (2016), "data cleaning" refers to the process of preparing data for analysis by removing or modifying data that is inaccurate, incomplete, irrelevant, duplicated, or improperly formatted. This is accomplished by sorting through the data and identifying and removing any of these issues as they are found. This data is typically neither required nor helpful when analyzing it because it may slow down the process or provide inaccurate results. As a result, this data is typically ignored. There are a few different ways to clean data, and the ones used depend on how the data is stored and the questions being asked.

In this breakdown, a variety of data cleaning strategies have been used. Data quality assessment outcomes highly influence data cleaning approaches. Data quality assessment focuses on ensuring that the data meets consistency, uniqueness, validity, timeliness, and accuracy.

The records with missing locations were not deleted. The Location variable has an inconsequential impact on the model; as a result, we replaced the missing values with 'Unkown" values. Deleting records with missing values would have reduced the dataset, yet the affected feature is insignificant. It is important to note that an essential feature, 'Sentiment' that contains labels of the tweets was missing.

4.4.2 Text pre-processing

Text processing was performed on the dataset as part of the data cleaning process. The raw data that arrive from a variety of sources need to be organised and cleaned up before they can be analysed. The pre-processing phase involves splitting the tokens and getting rid of the word stops is one of the more well-known processes in the process. Other processes, such as tagging, extraction, and representation, are standard practices. A sentence was broken up into graphics or other essential markers using a variety of approaches, including highlighting, taking out words and expressions, separating the tokens from the words, and separating the tokens from one other. Among these methods, highlighting was the most common. Stop words were not included in the study; instead, the pre-processing processes were used to remove them once they had been identified. Various administrators were required to return different portions of speech and words to their basic structures, which ultimately resulted in the procedure. When tagging content, using the many parts of speech used in dialect discourse to discover different components is highly significant.

The quality of a dataset can be improved by using pre-processing techniques. The dataset can be "cleaned" of "noise" using the pre-processing techniques such as correction of spelling errors, minimizing of repeated characters, and clarification of ambiguous acronyms). Additionally, pre-processing procedures include removing stopwords, eliminating punctuation marks, word stemming, and word lemmatization. In some circumstances, both stemming and lemmatization can increase the dataset's quality for sentiment analysis tasks.

For this study, variety of text pre-processing methods were used. First, the tweets were tokenized using the natural language tool kit's function word tokenize(). Tokenization is a process that divides large amounts of text into smaller pieces known as tokens. This process is performed on the text. These tokens are an essential component of the stemming and lemmatization processes, as well as an extremely helpful tool for discovering patterns. The replacement of sensitive data elements with non-sensitive data elements is another benefit that can be gained through tokenization. Secondly, the tweets were all converted into lowercase using the function lower ().

Thirdly, using the regular expression function; HTML tags, punctuation marks, and multiple sets of emoticons were removed. Fourthly we used the function ‘set (stopwords. words('English))' to get rid of stopwords. Stopwords are words that are used frequently throughout the entire language (Harris, n,d). These frequent words are words that are used frequently in a particular field. For any problem statement in a certain field, frequent words can be ignored because they don't convey much information. Finally, we performed stemming and lemmatization on the tweets. Lemmatisation is the computational process of identifying a word's lemma based on its intended meaning. Instead of just identifying the word's designated part of speech and meaning within a phrase, lemmatization additionally considers the surrounding context, such as the sentences immediately preceding or following it or even the entirety of the document. When it comes to stemming, inflected words are reduced to their root form, whereas derived words are reduced to their inflected form.

4.4.3 Text Labelling

A library known as VADER is used for the sentiment analysis of the text. Emotional polarity (positive/negative) and intensity (strong) are both considered by this model. Human raters and the collective wisdom of internet users are used in the VADER text sentiment analysis, which combines qualitative research with empirical validation. Using the dictionary and five heuristics, it is possible to link lexical traits to emotional intensity and encode how contextual elements affect the mood of a piece of Text (Calderon, 2017). Lexical features are linked to the intensity of an emotional response in a dictionary.

The goal of lexical approaches is to map emotions to words by constructing a lexicon, also known as a "dictionary of sentiment." We don't have to look anywhere else because this dictionary will be able to tell us how words and sentences make us feel regardless of what else we read. In lexical techniques, the sentiment category or score of each word in the sentence is considered separately before combining the results to determine the overall sentiment category or score of the sentence.

Sentiment analysis by VADER relies heavily on the use of lexicons that connect lexical features to emotional intensity ratings (these ratings are called sentiment scores). It is possible to determine the sentimental value of a text by adding up the emotional weight of each word in the text itself (Hutto & Gilbert, 2014). There are negative connotations that some terms bear and may or may not bother someone else, but for somebody, they do. This was addressed by enlisting many human raters and then averaging their scores for each word to create the VADER library (Calderon, 2017). Many people believe that group consensus is more reliable than an individual's, which is why this approach is known as the "wisdom of the crowd.". So, a score from -1 to 1 is used to indicate how positive or negative the VADER sentiment analysis results are. To calculate a sentence's sentiment score, the VADER lexicon's unique sentiment scores for each word in the sentence are added together (Hutto & Gilbert, 2014). Social media platforms are best suited for using this methodology, but it has also been found to be useful in evaluating the sentiment of film reviews and other opinion articles.

4.4.4 Data Labelling

This section focused on feature engineering. The feature engineering involved the development of polarity scores using the lexicon-based tool VADER. The tweets were subjected for sentiment scoring. Utilizing the “vaderSentiment” module, the SentimentIntensityAnalyzer function was imported. The function scores the tweets and stores them in the 'compound' variable. The polarity scores were stored using the ['polarity_score'] variable as a new feature in our dataset, dataset_2. As mentioned earlier, the polarity scores range from -1 to 1. +1 represents the extreme positive sentiment, while -1 is the extremely negative and 0 is neutral. A custom function was used to extract sentiments based on an arbitrary scale where a polarity score of above -0.05 was classified as positive, a polarity score of between -0.05 and 0.05 were classified as neutral, and a polarity score of less than -0.05 was negative. As a result, the analysis was able to develop a label for the tweets and store them under the variable sentiment.

4.4.5 Converting texts to a numeric matrix

The predictor variable in this study was the ‘Text’ attribute that was stored as strings. However, the modelling part ingests numeric data, hence we converted the text into numeric data. To achieve the numeric conversion, TfidfTransformer was utilised. It was necessary to first develop a CountVectorize to allow count the number of words, determine the size limit of the vocabulary, implement stop words, and so on. The tf insta.fit_transform (X train) function was used to calculate the values of the term-frequency inverse document frequency (TF-IDF). The word counts that were calculated before were transformed into an array using array(). The lowest IDF values will be seen in words with a high phrase frequency. Words with a lower IDF value are less likely to be found in a single text. Below is the implementation.

```
#Getting the term-frequency inverse document frequency (IDF) scores
from sklearn.feature_extraction.text import TfidfTransformer
tf_insta = TfidfTransformer()
X_train_filter = tf_insta.fit_transform(X_train).toarray() #get IDF scores and converting them into numeric array
```

Figure 2: Getting term-frequency inverse document frequency (IDF) scores

4.5 Modelling

The research aimed to develop a classification model capable of distinguishing tweets into one of three categories based on their sentiment: positive, negative, or neutral. Keras and TensorFlow were used alongside each other in order to implement the neural network. The architectural design of a neural network consists of the input, hidden, and output layers. At this stage, the model was configured by setting the parameters. The model function has input parameters which include; input_dim representing the input shape and output representing the output nodes. The input layers were set to 102 notes with the activation function as Rectified Linear Unit (‘relu’). The hidden layer consists of 3 layers with their respective nodes stated above; All the hidden layers were set to 'relu' as the activation function.

On the other hand, the output layer was set to 'softmax' as the activation function with three nodes. The three nodes represent the target values. The model as well adopted the 'categorical_crossentropy' loss function, since the problem at hand is a classification model. The cross-entropy loss quantifies how well a predicted set of classes coincides with the true set of classes. The cross-entropy loss increases when the anticipated classes diverge from the true classes. The evaluation metric for this study used accuracy, while "adam” was set as an optimizer. The model_params’s function returns the model configuration. When fitting the model, the data set was split into; 80 percent for training and 20 percent for validations. Ten epochs were used during the training. Below is the implementation.

```
#configuring the Neural network model using Keras and TensorFlow
def model_params(input_dim, output):
  model = tf.keras.models.Sequential([
      tf.keras.layers.Dense(512, input_dim=input_dim, activation='relu'),
      tf.keras.layers.Dense(256, activation='relu'),
      tf.keras.layers.Dense(102, activation='relu'),
      tf.keras.layers.Dropout(0.165),
      tf.keras.layers.Dense(128, activation='relu'),
      tf.keras.layers.Dense(64, activation='relu'),
      tf.keras.layers.Dense(output, activation='softmax')
  ])
  model.compile(
        loss='categorical_crossentropy',
        optimizer='adam',
        metrics=['accuracy']
    )
  return model
```

Figure 3: Configuring the neural Network Model

4.5.1 How a neural network works

A recurrent neural network for classification works by taking in a data point, and then using that data point to update the weights of the network. The network then takes in the next data point, and updates the weights again. This process continues until all data points have been used. Data points are fed into the recurrent neural network where they flow through the Hidden layer. The weights and bias of the Hidden layer are then updated using the back-propagation algorithm. The final layer is then used to classify the data point by propagating the output back through the previous layers (the Hidden and Input layers). The final output value is compared to the desired output value, and adjust the weights and biases accordingly. This process occurs on each new data point and continues until all data points have been outputted. In a feed forward neural network, the data points are just fed as input to the Hidden layer, meaning they don’t affect the weights at all. In a recurrent neural network, on the other hand, the weights are continually updated as the data points are fed into the network. A recurrent neural network for classification involves designing the architecture in such a way that the Hidden layer of the network learns to classify. This part is the critical aspect of a recurrent neural network used for classification.

5.0 Chapter

Results and Discussions

This chapter will focus on the outcome of the previous chapter. After the data collection the following dataset was obtained. The data displayed is a preview of the first 5 data points.

```
#Preview of the cleaned dataset
data_2.head()
```

	Text	location	Date
0	rt callofduti new era come tweet first receiv ...	Puerto Rico	2022-06-09 06:30:49
1	rt callofduti see pt summergamefest	roblox sex simulator	2022-06-09 06:30:47
2	beenoxcodpc callofduti pmsl would take away ad...	Glasgow	2022-06-09 06:30:32
3	rt callofduti new era come tweet first receiv ...	NaN	2022-06-09 06:30:24
4	callofduti know need	Cincinnati, OH	2022-06-09 06:30:08

Figure 4: Sample dataset records

5.1 Exploratory data analysis results

To be able to draw more insights from the data, additional features were developed. Using the Date variable that contained the dates when the tweets were posted, the days of the week and the month were derived. Generally, the outcome revealed that June was the only month, and Wednesday and Thursday were the only days of the week in the period timelines.

	Text	location	day_of_week	Month	Sentiment
count	3826	3826	3826	3826	3826
unique	3826	1238	2	1	3
top	RT @CallofDuty: A new era is coming. #ModernWa...	Uknown	Thursday	June	Positive
freq	1	1579	2008	3826	1691

Figure 5: Descriptive statistics

The descriptive statistics show that most of the tweets were from unknown locations. The 'unknown" is more a result of the imputation of missing Locations which were quite a number. Most tweets collected were tweeted on Thursday, and a majority of them were positive. June was the most popular month; well, it is the only month that we retrieved the tweets.

However, it was not possible to the rank location where the tweets were sent from. The outcome was as follows,

United States	195
Los Angeles, CA	80
California, USA	79
Texas, USA	66
Houston, TX	54
Chicago, IL	52
Atlanta, GA	48
Florida, USA	42
New York, USA	39
Las Vegas, NV	33

Table 1: The tweets distribution based on location

A huge fan base of the 'Call of Duty" franchise is located in the United States; the tweets posted within the mentioned period (8th and 9th June 2022) were mostly from the United States. The observation could also be due to the realise time of the trailer favoured fans from these regions. According to Fortune Business Insights (2021), the market for video games may be broken down into several distinct regions, including North America, South America, Europe, Asia Pacific, Africa, and the Middle East. It is anticipated that the Asia Pacific region will hold the largest share of the global market for video games, particularly in nations such as China, India, and Japan. Due to the presence of important company participants in the market in China, such as Tencent Corp., NetEase, Perfect World, and Shanda, amongst others, it is anticipated that China will hold the bulk of the share of the market in the Asia Pacific region.

Since 'Call of Duty' is American, and the country uses English, we should not expect to find tweets from non-English speaking nations like the Indian-Pacific countries. Besides that, the language parameter in the Tweepy API was set in English; hence, we got tweets in the English language only. This also explains where most of the tweets came from the United States. Presumably, the second-largest gaming region globally after the Asia-Pacific region is the United States because of the existence of Gaming players like Microsoft, Cave Digger, and many others.

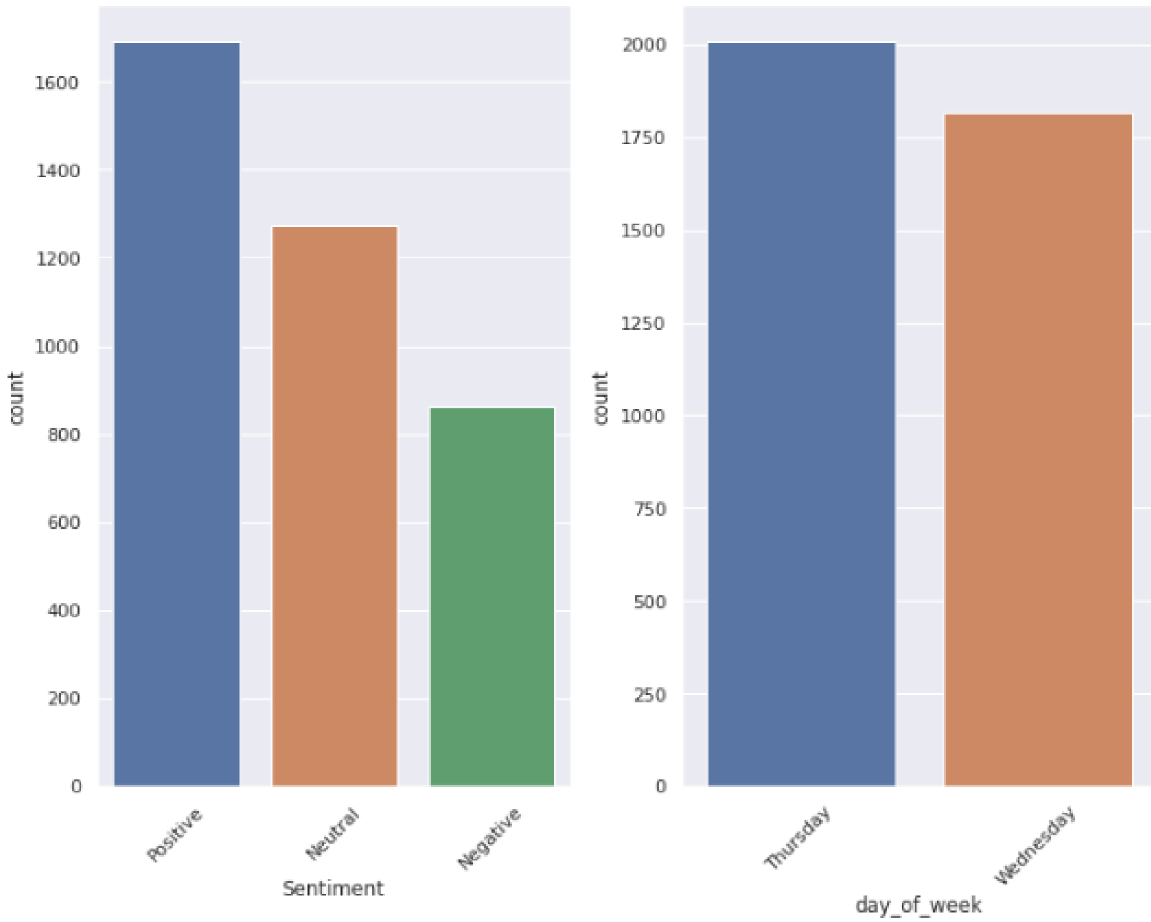

Figure 6: The sentiment distribution and Tweets distribution

The figure above reveals that the tweets obtained were mostly positive, and the neutral tweets came second, while the least was the negative tweets. The revelation should mean that the reception of trailers was good, and the fans are truly looking forward to enjoying the game. Also, most tweets were posted on Thursday, followed by Tuesday. The number of tweets from both days marginally varies. The high number of tweets on Thursday can be attributed to Kick off of Summer Game Fest, a huge event in the gaming calendar. The Summer Game Fest is a video game event that takes place entirely online and features several important events during June. During these events, video game publishers broadcast trailers of upcoming video games.

WordClouds

This section investigates the primary issue and source of concern brought on raised by 'Call of Duty' lovers. In order to accomplish this section's goal, word clouds were used to depict the frequently used and important words that were mentioned in the tweets. A "word cloud" is a collection of words shown in various font sizes. When it comes to the importance of a word, its length and complexity are directly correlated to the frequency with which it appears in a given text (Heimerl, Lohmann, Lange, and Ertl, 2014). Word clouds effectively isolate the most significant aspects of a message. Using a custom function, we separated tweets based on their sentiment. The dataset was segmented into three distinct subgroups, and as a result, we obtained three unique subsets of tweets: positive, neutral, and negative.

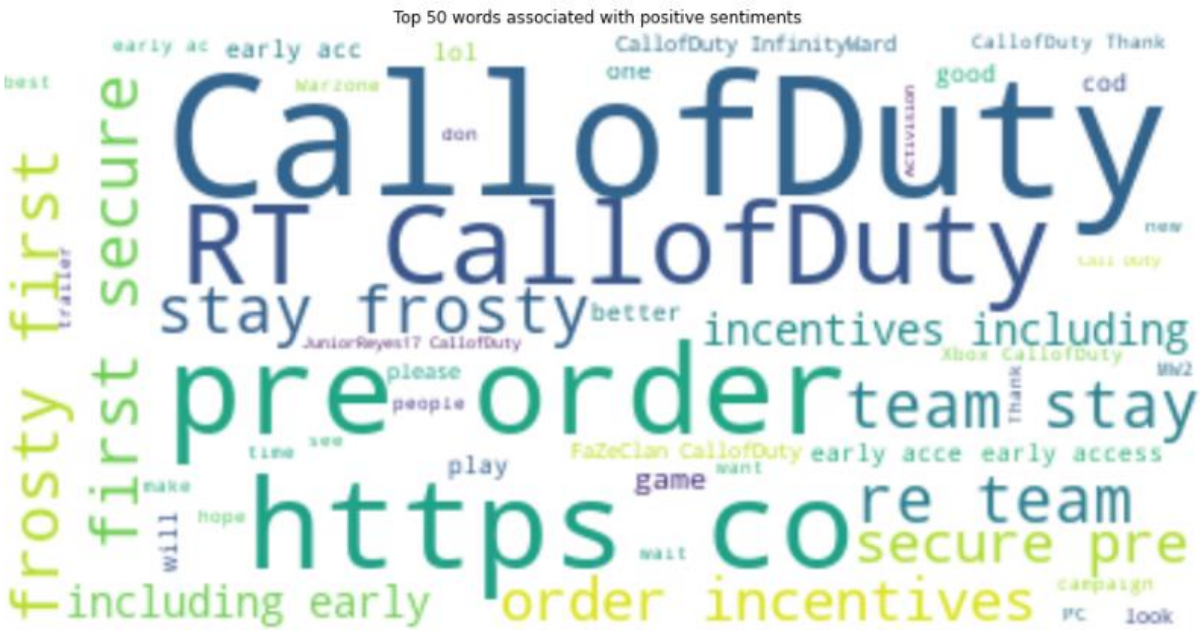

Figure 7: The word cloud of 50 positive words

The word cloud in figure above reveals the common top 50 words that are featured in the positive sentiments. The above revelation shows that dominant words were the Franchise's sequels, such as ‘Call of Duty Infinity ward’. In addition, words such as 'pre-order' and 'order incentives' dominated. This observation was so due to the campaign by the game publisher to urge fans to make pre-orders for the COD Modern Warfare 2. Pre-orders come with incentives of reduced prices.

The 2019 version of Call of Duty: Modern Warfare was rebooted as Call of Duty: Modern Warfare 2 in 2022. According to Fischer (2022), significant advancements include an Artificial Intelligence (A.I.) system that is more intelligent, receptive, and flexible and possesses advanced awareness. Because the A.I. will respond in a manner that is consistent with bullet wounds, calibre, momentum, and the physics of the environment, the fact that this game was in development for a longer period of time than any other previous “Call of Duty” release ought to be reflected in the level of polish and the quality of the content included in the game. Therefore, explaining the more positive reviews that the trailer received.

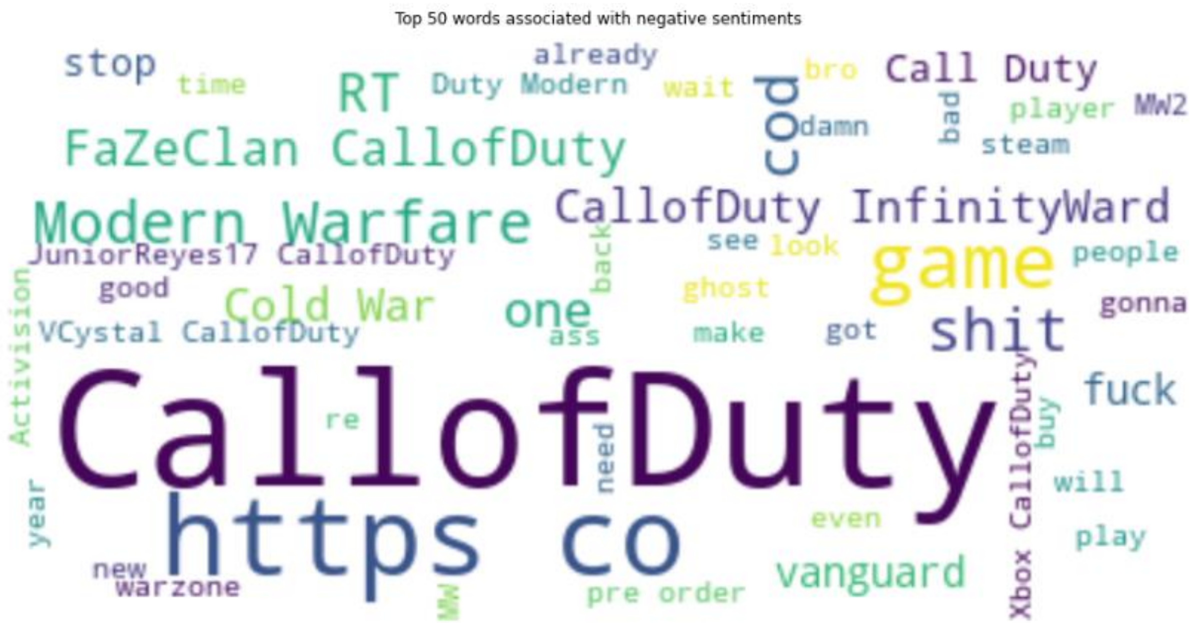

Figure 8: The top 50 words featured in negative tweets

The figure above reveals a word cloud of the top 50 negative tweets. The dominant words that are featured contain the Franchise's sequels, such as 'Call of Duty,' 'Infinity ward,' 'Modern Warfare,' and ‘Warzone.' From the figure above, there is no clear issue that calls for concern about the game. Technical issues about the game may be difficult to tell from the trailer since users have not really interacted with the game from their consoles, and hence it is difficult to find something of concern from the word cloud. The criticism from a trailer may touch on the graphics, soundtrack, and features, which can also be limited since a trailer can be shorter than three minutes.

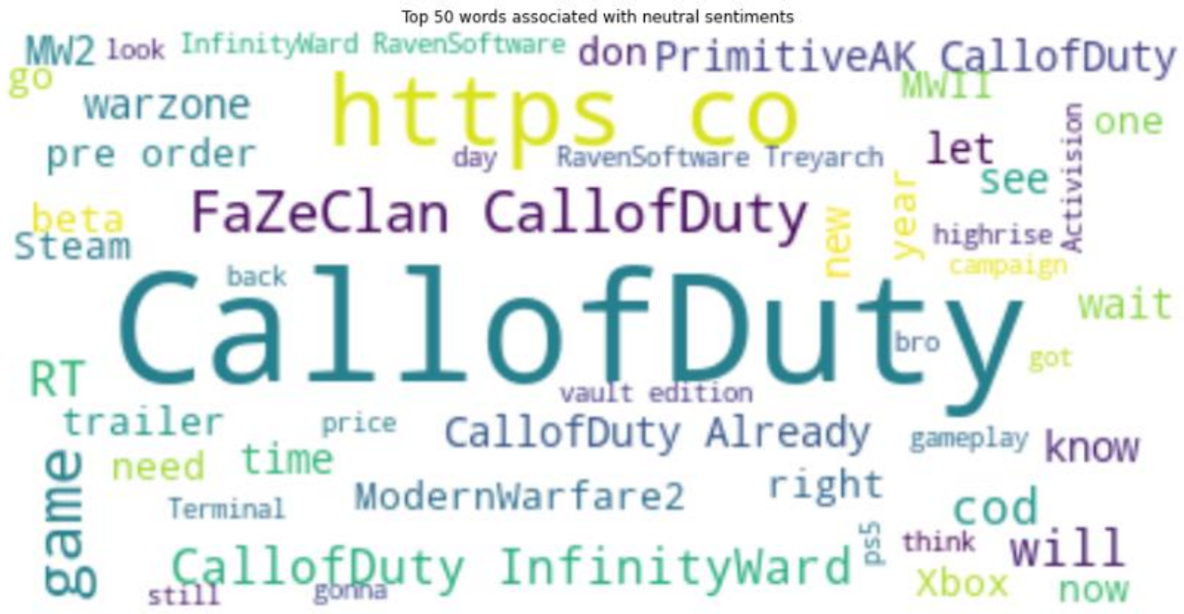

Figure 9: The top 50 words associated with neutral sentiments

The figure above reveals common words associated with neutral tweets. The dominant words featured in this category include; 'Call of Duty Infinity ward,' Modern Warfare, Warzone, Infinity Ward, Modern Warfare 2' and 'Activision'. Neutral tweets do not offer any insight as they bear neither good nor bad news.

Activision is the company that is responsible for publishing the video game franchise, Call of Duty. When it first launched in 2003, its primary focus was on video games that were set during the Second World War (Gish, 2010). Over the course of its existence, the Franchise has produced video games that take place in a variety of settings, including the middle of the Cold War, futuristic worlds, and outer space. Infinity Ward was the company that created the games in the beginning. Call of Duty: Vanguard, the most recent sequel, was made available for purchase on November 5, 2021. The following sequel, titled Call of Duty: Modern Warfare II, is scheduled to be made available for purchase on October 28, 2022. These video games are well-known for the intense first-person shooter action they feature, in addition to the captivating narratives they feature. Every new iteration in the series introduces something fresh to the table, whether it be a different environment, set of gameplay mechanics, or cast of characters.

Converting texts to a numeric matrix

Converting the sentences into a matrix is a necessary step that must be completed before our model can be fed the dataset. We are taking raw text and transforming it into a numerical vector representation of the words and n-grams using a tool called CountVectorizer. Because of this, it is very simple to use this format directly as features in Machine Learning tasks such as text classification and clustering because it is so easily accessible. Therefore, in order to translate them into a matrix, we made use of the CountVectorizer and the TfidfTransformer.

5.2 Modelling Observations

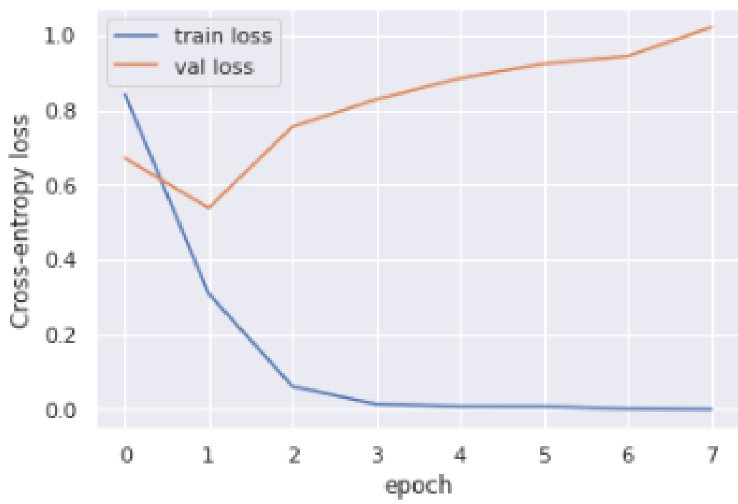

Figure 10: The 'categorical cross entropy' loss

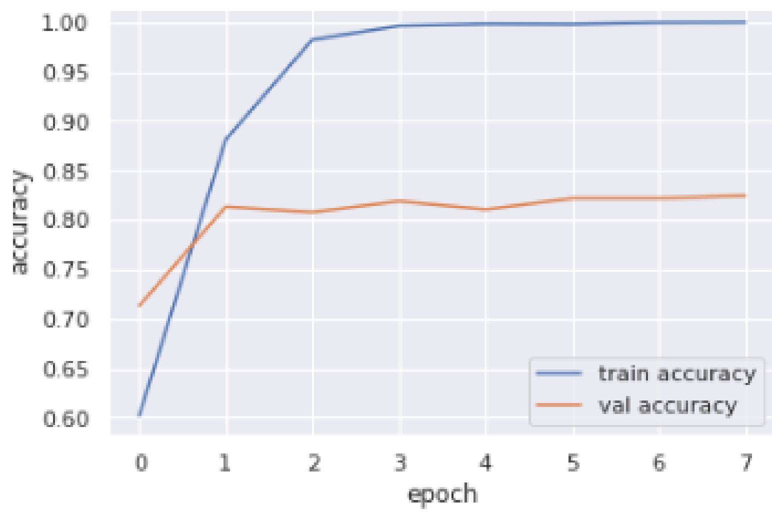

Figure 11: The model's accuracy

The figure 11 above reveals the training performance of the model. The training accuracy attained 99 percent, while the validation accuracy was significantly low at 82 percent. The figure compares the training loss alongside the validation loss. The training loss is approximately 0.08, while the validation loss was 1.1, based on the categorical cross entropy loss function. The model accuracy showed encouraging results; however, the cross-entropy loss showed increasing entropy. In an ideal setting, the entropy should be dropping. An increasing cross-entropy loss means that the model is not performing as well as it could be.

The training accuracy was encouraging, but the validation accuracy was significantly lower. This suggests that the model is overfitting the training data. To improve the model, I recommend using regularization techniques such as dropout or weight decay. Overfitting refers to a situation in which the model, fitted on an example set, reproduces that specific set too closely, but fails to generalize to other examples. It indicates an inability of model to capture the underlying law of the problem and is a result of model complexity. An overfitted model will be merely encoding the specificities of the dataset and not learning any general patterns. Regularization techniques exist in machine learning for one purpose, to prevent overfitting and to encourage the model to focus on generalizable features instead of memorizing the training example examples.

TensorFlow and Keras were employed in the experiment. TensorFlow, a free and open-source software framework, uses data flow graphs to do numerical computations. The network edges represent the multidimensional data arrays (tensors) exchanged between the nodes, while the graph nodes represent mathematical operations. It was designed with the primary goal of facilitating rapid experimentation in mind during development. One of the most important aspects of good research is the ability to go quickly and efficiently from a notion to a result. TensorFlow allows you to quickly develop and test a wide variety of models and architectures. The flexible design allows users to distribute computing between one or more CPUs or GPUs on a desktop, server, or mobile device.

Chapter 6

6.1 Conclusion

From this research study, it is evident that it is possible for organisations to effectively conduct sentiment analysis from their customer feedback to improve their service delivery. Customer satisfaction and experience are a great way to retain customers and acquire some more through means like word of mouth. According to Ovchinnikov et al. (2014), it is five times more expensive to acquire a new customer, and a sixty to seventy per cent chance of selling to a customer one already has compared to the five to twenty per cent success chance of selling to a new customer. Therefore, it is critical to pay attention to customer service, quality of products and services, and creating a customer-friendly environment. Ensuring that customers provide positive and excellent feedback is critical, especially on social media platforms, as people look for their peers' comments and reviews to inform their purchase decision.

The study found that mining sentiments from online reviews are possible using lexicon, rule-based, and machine learning tools. However, the lexicon method has shortcomings as it does not consider the context. The use of VADER provided an improved tool that can recognise punctuations, emoticons, and slang words. This stool is highly effective and suitable for scoring social media posts or reviews. The research study resulted in a model that achieved the success criteria of attaining accuracy of eighty per cent or above. Nonetheless, the results of the evaluation showed a grim picture. A good model should have high accuracy with a dropping cross-entropy loss. A low cross-entropy loss means that the predicted probabilities are close to the true labels.

Sentiment analysis and opinion mining have numerous perks for organisations and can offer a great way to upsell opportunities. Extracting sentiments from customer feedback makes it possible to easily identify the happiest customers, helping to steer the direction of sale pitches to try and make them spend more. Additionally, it can aid business enterprises from upsetting disgruntled customers with further unwelcome sale pitches. Moreover, sentiment analysis can help to better train chatbots which are key in customer service delivery recently. Chatbots can be trained to respond to and recognise certain customer moods and cues. Sentiment analysis can help detect whether a conversation in chatbots needs to be escalated to a human agent or route a potential client to a sales team or agent (Dang et al., 2020).

Sentiment analysis can come in handy in various use cases for organisations. For instance, social media monitoring and brand management can be done more effectively using models trained to collect positive and negative feedback. Product analysis is also a necessary use case for sentiment analysis where a company closely monitors its launched products and how customers relate to them. Therefore, extraction of sentiments and opinion mining is an essential process that each organisation looking to improve their bottom line should adopt to have more satisfied customers.

6.2 Recommendations

The targets of the tweets were obtained from fairly accurate tools; therefore, the labels might not have been as accurate as required. Ideally, in sentiment analysis, humans would rate themselves as more accurate since they better understand the context and informal language. Additionally, humans have the ability to distinguish grammatical errors and make appropriate judgements. In this case, the accuracy of the VADER used in generating tweet labels cannot be ascertained; therefore, it may ultimately affect the developed model's accuracy. Therefore, a recommendation for the provision of adequate resources is put in place to prepare an accurate training dataset. The following are further recommendations to organisation entities and future researchers:

- Organisation entities should ensure that they provide avenues where customers can make comments and reviews about their purchases. Product review features should be at the core of any company's website selling products online.
- Organisations should be sure to analyse and extract customer feedback from different avenues and use it to make better-informed decisions in their marketing campaigns. Data is valueless unless processed well to gain useful insights to improve service delivery and business processes.
- Businesses online should incorporate the use of social media in their business processes and leverage the power of communities. Social platforms belonging to companies should be well maintained to ensure that customers' queries and issues are handled quickly, professionally, and efficiently.
- Future researchers should seek to improve the accuracy by applying pre-trained models.
- Businesses globally should avail more areas of interactions with customers to find out and understand their pain points better and quicker. Tools like chatbots and other customer

relationship management tools should be adopted to increase efficiency and uptime when customers can get answers to their queries. Sentiment analysis should be adopted in conjunction with chatbots to make them more intelligent and efficient.

- Further research should be done on how to conduct sentiment analysis on video and audio data, as a lot of information and feedback can be obtained from such avenues. Furthermore, image and speech recognition should be incorporated into sentiment analysis and opinion mining.
- Companies should emphasise the quality of data they collect and ensure data correctness from the point of data collection to ensure training datasets are more accurate, resulting in more accurate models.

References

Arulmurugan, R., Sabarmathi, K., & Anandakumar, H. (2019). Classification of sentence level sentiment analysis using cloud machine learning techniques. *Cluster Computing*, *22*(1), 1199-1209.

Baldwin, T., Cook, P., Lui, M., MacKinlay, A., & Wang, L. (2013). How noisy social media text, how diffrnt social media sources? Proceedings of the Sixth International Joint Conference on Natural Language Processing,

Behdenna, S., Barigou, F., & Belalem, G. (2018). Document level sentiment analysis: a survey. *EAI Endorsed Transactions on Context-aware Systems and Applications*, *4*(13).

Beigi, G., Hu, X., Maciejewski, R., & Liu, H. (2016). An overview of sentiment analysis in social media and its applications in disaster relief. *Sentiment analysis and ontology engineering*, 313-340.

Chan, J. O. (2013). An architecture for big data analytics. *Communications of the IIMA*, *13*(2), 1.

Chen, J.-S., Tran-Thien-Y, L., & Florence, D. (2021). Usability and responsiveness of artificial intelligence chatbot on online customer experience in e-retailing. *International Journal of Retail & Distribution Management*, *49*(11), 1512-1531.

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

Cruz, F. L., Troyano, J. A., Enríquez, F., Ortega, F. J., & Vallejo, C. G. (2010). A knowledge-rich approach to feature-based opinion extraction from product reviews. Proceedings of the 2nd international workshop on Search and mining user-generated contents,

Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment analysis based on deep learning: A comparative study. *Electronics*, *9*(3), 483.

Das, T. K., & Kumar, P. M. (2013). Big data analytics: A framework for unstructured data analysis. *International Journal of Engineering Science & Technology*, *5*(1), 153.

Dashtipour, K., Poria, S., Hussain, A., Cambria, E., Hawalah, A. Y., Gelbukh, A., & Zhou, Q. (2016). Multilingual sentiment analysis: state of the art and independent comparison of techniques. *Cognitive computation*, *8*(4), 757-771.

Delafrooz, N., Taleghani, M., & Nouri, B. (2014). Effect of green marketing on consumer purchase behavior. *QScience Connect*, *2014*(1), 5.

Gallagher, C., Furey, E., & Curran, K. (2019). The application of sentiment analysis and text analytics to customer experience reviews to understand what customers are really saying. *International Journal of Data Warehousing and Mining (IJDWM)*, *15*(4), 21-47.

Garcia-Garcia, J. M., Penichet, V. M., & Lozano, M. D. (2017). Emotion detection: a technology review. Proceedings of the XVIII international conference on human computer interaction,

Gerber, D. J. (2015). Adapting the role of economics in competition law: a developing country dilemma. In *The Economic Characteristics of Developing Jurisdictions* (pp. 248-264). Edward Elgar Publishing.

Giglio, S., Pantano, E., Bilotta, E., & Melewar, T. (2020). Branding luxury hotels: Evidence from the analysis of consumers'"big" visual data on TripAdvisor. *Journal of Business Research*, *119*, 495-501.

Gitari, N. D., Zuping, Z., Damien, H., & Long, J. (2015). A lexicon-based approach for hate speech detection. *International Journal of Multimedia and Ubiquitous Engineering*, *10*(4), 215-230.

Guerini, M., Gatti, L., & Turchi, M. (2013). Sentiment analysis: How to derive prior polarities from SentiWordNet. *arXiv preprint arXiv:1309.5843*.

Gull, M., & Pervaiz, A. (2018). Customer behavior analysis towards online shopping using data mining. 2018 5th International Multi-Topic ICT Conference (IMTIC),

Guzman, E., & Maalej, W. (2014). How do users like this feature? a fine grained sentiment analysis of app reviews. 2014 IEEE 22nd international requirements engineering conference (RE),

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California management review*, *61*(4), 5-14.

Huber, S., Wiemer, H., Schneider, D., & Ihlenfeldt, S. (2019). DMME: Data mining methodology for engineering applications–a holistic extension to the CRISP-DM model. *Procedia Cirp*, *79*, 403-408.

Jang, S., Prasad, A., & Ratchford, B. T. (2012). How consumers use product reviews in the purchase decision process. *Marketing letters*, *23*(3), 825-838.

Jaworski, B., & Kohli, A. K. (2014). Co-creating the Voice of the customer. In *The service-dominant logic of marketing* (pp. 127-135). Routledge.

Li, X., Law, R., Xie, G., & Wang, S. (2021). Review of tourism forecasting research with internet data. *Tourism Management*, *83*, 104245.

Liang, T.-P., Li, X., Yang, C.-T., & Wang, M. (2015). What in consumer reviews affects the sales of mobile apps: A multifacet sentiment analysis approach. *International Journal of Electronic Commerce*, *20*(2), 236-260.

Lo, S. L., Cambria, E., Chiong, R., & Cornforth, D. (2017). Multilingual sentiment analysis: from formal to informal and scarce resource languages. *Artificial Intelligence Review*, *48*(4), 499-527.

Martínez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernández-Orallo, J., Kull, M., Lachiche, N., Ramírez-Quintana, M. J., & Flach, P. (2019). CRISP-DM twenty years later: From data mining processes to data science trajectories. *IEEE Transactions on Knowledge and Data Engineering*, *33*(8), 3048-3061.

McColl-Kennedy, J. R., Zaki, M., Lemon, K. N., Urmetzer, F., & Neely, A. (2019). Gaining customer experience insights that matter. *Journal of Service Research*, *22*(1), 8-26.

Nemes, L., & Kiss, A. (2021). Social media sentiment analysis based on COVID-19. *Journal of Information and Telecommunication*, *5*(1), 1-15.

Nobar, H. B. K., & Rostamzadeh, R. (2018). The impact of customer satisfaction, customer experience and customer loyalty on brand power: empirical evidence from hotel industry. *Journal of Business Economics and Management*, *19*(2), 417-430.

Ofoeda, J., Boateng, R., & Effah, J. (2019). Application programming interface (API) research: A review of the past to inform the future. *International Journal of Enterprise Information Systems (IJEIS)*, *15*(3), 76-95.

Ovchinnikov, A., Boulu-Reshef, B., & Pfeifer, P. E. (2014). Balancing acquisition and retention spending for firms with limited capacity. *Management science*, *60*(8), 2002-2019.

Păvăloaia, V.-D., Teodor, E.-M., Fotache, D., & Danileţ, M. (2019). Opinion mining on social media data: sentiment analysis of user preferences. *Sustainability*, *11*(16), 4459.

Priyantina, R. A., & Sarno, R. (2019). Sentiment analysis of hotel reviews using Latent Dirichlet Allocation, semantic similarity and LSTM. *Int. J. Intell. Eng. Syst*, *12*(4), 142-155.

PS, S., & Mahalakshmi, G. (2017). Emotion models: a review. *International Journal of Control Theory and Applications*, *10*(8), 651-657.

Puspitasari, N. B., WP, S. N., Amyhorsea, D. N., & Susanty, A. (2018). Consumer's buying decision-making process in E-commerce. E3S Web of Conferences, Raghuvanshi, N., & Patil, J. (2016). A brief review on sentiment analysis. 2016 International Conference on Electrical, Electronics, and Optimisation Techniques (ICEEOT),

Rana, S., & Singh, A. (2016). Comparative analysis of sentiment orientation using SVM and Naive Bayes techniques. 2016 2nd International Conference on Next Generation Computing Technologies (NGCT),

Russ, R. S. (2014). Epistemology of science vs. epistemology for science. *Science Education*.

Rusu, O., Halcu, I., Grigoriu, O., Neculoiu, G., Sandulescu, V., Marinescu, M., & Marinescu, V. (2013). Converting unstructured and semi-structured data into knowledge. 2013 11th RoEduNet International Conference,

Sarathy, P. S., & Patro, S. K. (2013). The role of opinion leaders in high-involvement purchases: An empirical investigation. *South Asian Journal of Management*, *20*(2), 127.

Scherer, K. R., & Ekman, P. (2014). *Approaches to emotion*. Psychology Press.

Schiffman, L., O'Cass, A., Paladino, A., & Carlson, J. (2013). *Consumer behaviour*. Pearson Higher Education AU.

Schröer, C., Kruse, F., & Gómez, J. M. (2021). A systematic literature review on applying CRISP-DM process model. *Procedia Computer Science*, *181*, 526-534.

Shafique, U., & Qaiser, H. (2014). A comparative study of data mining process models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research*, *12*(1), 217-222.

Sharma, M. K. (2014). The impact on consumer buying behaviour: Cognitive dissonance. *Global Journal of Finance and Management*, *6*(9), 833-840.

Shirdastian, H., Laroche, M., & Richard, M.-O. (2019). Using big data analytics to study brand authenticity sentiments: The case of Starbucks on Twitter. *International Journal of Information Management*, *48*, 291-307.

Strelluf, C. (2019). Anymore, It's on TwitterPositive Anymore, American Regional Dialects, and Polarity Licensing In Tweets. *American Speech*, *94*(3), 313-351.

Vijayarani, S., Ilamathi, M. J., & Nithya, M. (2015). Pre-processing techniques for text mining-an overview. *International Journal of Computer Science & Communication Networks*, *5*(1), 7-16.

Ye, C., & Nov, O. (2013). Exploring user contributed information in social computing systems: quantity versus quality. *Online Information Review*.

Yogesh, F., & Yesha, M. (2014). Effect of social media on purchase decision. *Pacific Business Review International*, *6*(11), 45-51.

Zaki, M., Kandeil, D., Neely, A., & McColl-Kennedy, J. R. (2016). The fallacy of the net promoter score: Customer loyalty predictive model. *Cambridge Service Alliance*, *10*, 1-25.

Zhang, L., Wang, S., & Liu, B. (2018). Deep learning for sentiment analysis: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *8*(4), e1253.