SENTIMENT ANALYSIS OF AMAZON REVIEWS USING MACHINE LEARNING MODEL

OMAR MOHAMMED ALI ALBAAGARI

UNIVERSITI TEKNOLOGI MALAYSIA

PSZ 19:16 (Pind. 1/13)

UNIVERSITI TEKNOLOGI MALAYSIA

DECLARATION OF THESIS / UNDERGRADUATE PROJECT REPORT AND COPYRIGHT

Author's full name : OMAR MOHAMMED ALI ALBAAGARI

Date of Birth : 13/04/2000

Title : SENTIMENT ANALYSIS OF AMAZON REVIEWS USING MACHINE LEARNING MODEL

Academic Session :

I declare that this thesis is classified as:

☐	**CONFIDENTIAL**	(Contains confidential information under the Official Secret Act 1972)*
☐	**RESTRICTED**	(Contains restricted information as specified by the organization where research was done)*
✓	**OPEN ACCESS**	I agree that my thesis to be published as online open access (full text)

1. I acknowledged that Universiti Teknologi Malaysia reserves the right as follows:
2. The thesis is the property of Universiti Teknologi Malaysia
3. The Library of Universiti Teknologi Malaysia has the right to make copies for the purpose of research only.
4. The Library has the right to make copies of the thesis for academic exchange.

Certified by:

SIGNATURE OF STUDENT	**SIGNATURE OF SUPERVISOR**
MCS41021	
MATRIX NUMBER	**NAME OF SUPERVISOR**
Date: 17 JAN 2024	Date: 17 JAN 2025

NOTES : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentiality or restriction

“I hereby declare that we have read this thesis and in my
opinion this thesis is suffcient in term of scope and quality for the
award of the degree of Master of Data Science ”

BAHAGIAN A - Pengesahan Kerjasama*

Adalah disahkan bahawa projek penyelidikan tesis ini telah dilaksanakan melalui kerjasama antara Click or tap here to enter text. dengan Click or tap here to enter text.

Disahkan oleh:

Tandatangan : Tarikh :

Nama :

Jawatan :

(Cop rasmi)

** Jika penyediaan tesis atau projek melibatkan kerjasama.*

BAHAGIAN B - Untuk Kegunaan Pejabat Sekolah Pengajian Siswazah

Tesis ini telah diperiksa dan diakui oleh:

Nama dan Alamat Pcmeriksa Luar :

Nama dan Alamat Pcmeriksa Dalam :

Nama Penyelia Lain (jika ada) :

Disahkan oleh Timbalan Pendaftar di SPS:

Tandatangan : Tarikh : 15JULAI 2018

Nama :

SENTIMENT ANALYSIS OF AMAZON REVIEWS USING MACHINE LEARNING MODEL

OMAR MOHAMMED ALI ALBAAGARI

A reportsubmitted in partial fulfilment of the
requirements for the award of the degree of
Master of Data Science

School of Computing
Faculty of Engineering
Universiti Teknologi Malaysia

JAN 2025

DECLARATION

I declare that this thesis entitled "SENTIMENT ANALYSIS OF AMAZON REVIEWS USING MACHINE LEARNING MODEL" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature	:	*ALBAAGARI*
Name	:	Omar Mohammed Ali Albaagari
Date	:	17 JANUARY 2025

DEDICATION

This thesis is dedicated to my father, who taught me that the best kind of knowledge to have is that which is learned for its own sake. It is also dedicated to my mother, who taught me that even the largest task can be accomplished if it is done one step at a time.

ACKNOWLEDGEMENT

ABSTRACT

Recently, Amazon has seen swift advancement. Consequently, online purchasing has increased, resulting in a rise in online customer reviews of products. The implicit sentiments in customer reviews significantly impact purchasing decisions, as consumers' perceptions of a product are shaped by the recommendations or grievances of other buyers. This research analyses the Amazon Reviews dataset and examines sentiment classification using several machine learning methodologies. Initially, the reviews were sanitised and subsequently reprocessed, followed by the execution of exploratory data analysis in this research. Ultimately, the VADER and Roberta models were created for sentiment analysis. The data frame for each model has been obtained.

ABSTRAK

Baru-baru ini, amazon telah melihat kemajuan pantas. Akibatnya, pembelian dalam talian telah meningkat, mengakibatkan peningkatan dalam ulasan pelanggan dalam talian terhadap produk. Sentimen tersirat dalam ulasan pelanggan memberi kesan ketara kepada keputusan pembelian, kerana persepsi pengguna terhadap sesuatu produk dibentuk oleh cadangan atau rungutan pembeli lain. Penyelidikan ini menganalisis set data ulasan amazon dan mengkaji klasifikasi sentimen menggunakan beberapa metodologi pembelajaran mesin. Pada mulanya, ulasan telah dibersihkan dan kemudiannya diproses semula, diikuti dengan pelaksanaan analisis data penerokaan dalam penyelidikan ini. Akhirnya, model vader dan roberta dicipta untuk analisis sentimen. Kerangka data bagi setiap model telah diperolehi..

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Problem Background

Emotions are present in every single situation in which people engage with one another (De Saa and Ranathunga, 2020). In a variety of contexts, they have the ability to mould an individual's view of an experience, a subject, an issue, and so on. Through a number of channels, Message boards, comments, and reviews are all examples of places where we could get feedback. and opinions about a wide range of items, both online and offline. These feedback and ideas may be presented in the form of text, video, image, and other forms. Each and every kind of feedback has a certain degree of emotion, such as whether the experience as a whole was positive, negative, or neutral."

In the twenty-first century, the internet has grown into a technology that is indispensable to our everyday life (Ripa et al., 2021). People are purchasing things from a large number of e-commerce websites in the current period, and it is more likely that they would first assess the products before purchasing them [Rathor et al., 2018]. In this modern era, all commercial enterprises have turned their focus to the internet technology. These days, the majority of websites that are dedicated to e-commerce now have a specific part where customers may submit evaluations of products or services

Sentimental analysis is one of the machine learning processing techniques that helps detect feelings (Rajat et al., 2021). This approach enables business owners to collect information about the perspectives of their customers via various online media, such as social media, questionnaires, and analyses of websites that allow for online shopping. A better understanding of the factors that contribute to the deterioration of the commodity will be possible as a result of this information. In this analysis, two things are taken into consideration: the line ”Apple Iphone 15 battery life is good and speakers’ quality is not good” is an example of sentiment analysis. Sentiment analysis represents the behavior of the consumer with regard to the product, as well as the reputation of the company. When it comes to the quality of the speakers, there is a negative opinion, but the battery life is great (Gope et al., 2022).

1.2 Problem Statement

Customer ratings and reviews reveal the buyer’s judgment on the product. It might be positive, negative, or neutral. When it comes to a product, customers could give it four or five stars, indicating contentment, while others may give it one or two stars, indicating discontent. When it comes to sentiment analysis, this does not provide any kinds of challenges. Other individuals, on the other hand, have given it three stars, despite the fact that they have definitely expressed their overall delight with it. As a result, this causes confusion among other consumers as well as among businesses that are interested in learning their genuine opinion. As a consequence of this, both consumers and businesses have challenges when it comes to assessing evaluations and comprehending the level of happiness experienced by customers. Therefore, the three-star rating does not genuinely indicate a neutral mood. This is due to the fact that in actuality, individuals who give a product or service a rating of three stars do not always imply that they are completely balanced in their judgment between both positive and negative opinion.

Based on the issue that mentioned above, the purpose of this research is to do a sentiment analysis applied to the Amazon office products dataset reviews in order to forecast the opinions of customers. This research will use real dataset from Amazon using the VADER (Valence Aware Dictionary and sEntiment Reasoner) and Reberta (A Robustly Optimized BERT Pretraining Approach) models in order to do the sentiment analysis.

1.3 Research Questions

This thesis seeks to illustrate consumer behaviour trends based on their purchases and reviews. Utilising the two distinct models reveals the consumers' reviews. A series of procedures must be undertaken to visualise the pattern of customer evaluations. Consequently, several research questions have been delineated for the experiment.

The research questions are:

(a) What preprocessing steps to carry out for the analysing sentiment analysis from office product dataset?

(b) What relevant keywords can be identified and retrieved by VADER and Roberta from the review's dataset?

(c) What conclusions may be derived from the customers purchases?

1.4 Research Objectives

(a) To conduct a preprocessing of the office products reviews datasets for sentiment analysis.

(b) To train a machine learning model that is capable of sorting customer evaluations into three unique sentiment categories, namely positive, neutral, and negative categories

(c) To develop a dashboard that summarize the analysis and making conclusion of their behaviour.

1.5 Scope of the Research

The scopes of the research are:

(a) The data will be collected from Amazon reviews’2023 Repository.

(b) The programming language that is chosen is Python.

(c) Implementing the VADER and Roberta models for the sentiment analysis

(d) Sorting the customer reviews into positive, neutral, and negative categories.

(e) Building a dashboard using PowerBi for the behaviour of the customers

CHAPTER 2

LITERATURE REVIEW

2.1 Summary

This chapter’s objective is to survey recent studies and reviews of relevant literature on sentiment analysis. An introduction to the many sentiment analysis levels follows, including topics such as data collecting, data pre-processing, sentiment analysis approaches, and lastly, challenges in the field. This chapter gives a good foundation for the sentiment analysis by covering it in detail.

2.2 Degree of the Sentiment Analysis

Several distinct degrees of investigation have been conducted on the subject of sentiment analysis. It is largely possible to identify sentiments and viewpoints at the level of the text, phrase, or aspect (Do et al., 2019). An illustration of the degrees of sentiment analysis may be seen in Figure 2.1. The first two levels are very great to go through, but they are also really challenging. In spite of this, the third level is more challenging than the levels that came before it since it demands a more comprehensive examination. (Cambria et al., 2017). Below you will find a brief summary of each level being discussed.

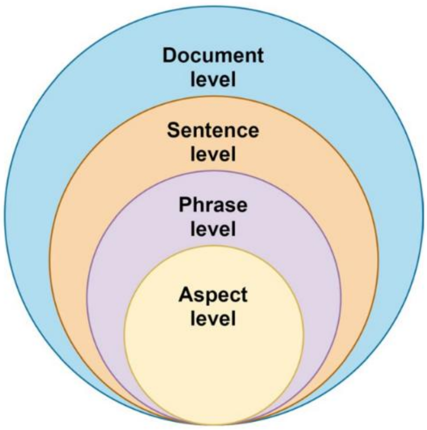

Figure 2.1 Various levels of emotional analysis (Wankhade et al., 2022).

2.2.1 Documents for sentiment analysis

This particular Part, trying to figure out if the overall attitude or viewpoint is favourable or negative. (Alqaryouti et al., 2020). Every text is classified based on the overarching opinion expressed by the author towards a certain object (such a product, for instance). When a single individual is responsible for its creation, categorization at the document level exhibits the highest level of success, and it isn't suited for publications that evaluate or compare many entities. There have been a lot of different approaches proposed for doing sentiment analysis at the document level. For the purpose of conducting document-level sentiment analysis, a framework that is independent of domains was designed (Zhao et al., 2017). This framework makes use of weighting criteria that are taken from Rhetorical Structure Theory (RST). After creating rhetorical structure trees, the authors analyzed the papers by calculating the emotion scores of phrases using two existing lexicons. This technique was used to analyze the articles. They compiled the scores of the sentences in accordance with the weighting criteria that had been developed in order to ascertain the condition of polarity of the emotion that was expressed throughout the article. The use of analysis in sentiment is very advantageous across a wide range of application areas; yet, documents may sometimes include competing sentiments that have the potential to impact the final judgment.

When doing analysis of sentiment at the document part, it's necessary to examine that the complete content. This type of sentiment analysis is defined by assigning a single polarity to the entire text. This specific sentiment analysis is hardly used. According to the information that it gives, this tool has the capability of categorising the chapters or pages of a book as either positive, negative, or neutral. Currently, both supervised and unsupervised learning approaches may be employed for material categorisation (Bhatia et al., 2015). The analysis of sentiment across several domains and the analysis of sentiment across numerous languages are the two most essential challenges within the arena of document-level sentiment analysis. According to (Saunders, 2021), When it comes to the domain in issue, domain-specific sentiment analysis is very accurate and highly sensitive to the language used. In the context of this discussion, the feature vector is a collection of words that in addition to being restricted in scope, must also be particular to the subject at hand.

2.2.2 Analysis of Opinion at the Sentence Level

During this stage, the phrase is the primary focus of attention. The major purpose of this analysis is to ascertain if the language transmits a positive, negative, or neutral mood (Liu, 2022). In order to achieve this aim, the statement must be classified as either objective, which conveys information that is true, or subjective, which reflects views and opinions.

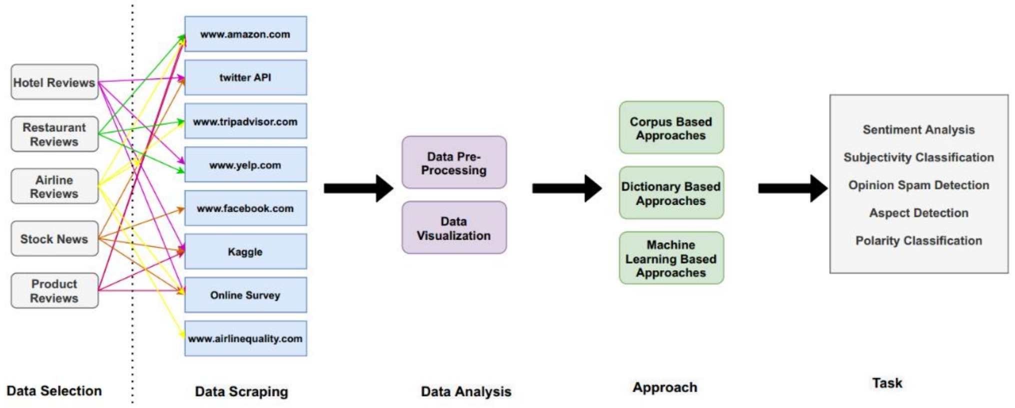

Figure 2.2 Procedures that are often used in sentiment analysis (Wankhade et al., 2022)

2.2.3 At the phrase level, an analysis of the sentiment

There is also the possibility of doing analysis of sentiment at the phrase part, which involves the extraction and that which is classified as decision-making words. It is possible for each phrase to include a single element or a handful of items. It is worth noting that a single aspect is stated in a sentence (Thet et al., 2010), which may prove to be advantageous for product evaluations that span many lines. Recently, this has been a popular topic of inquiry that has been being conducted. Given that. It is more helpful to do sentence-level analysis rather than text-level analysis when analysing a document that includes both positive and negative statements. Analysis of the text-level focuses on categorizing the whole content as either favourably or adversely subjective. The word is the basic unit of language, and the polarity of a word is strongly tied to the subjectivity of the sentence or text in which it is used. There is a high probability that a statement that contains an adjective is subjective (Fredriksen-Goldsen and Kim, 2017). In addition, the phrase that was selected for expression is representative of the demographic features of individuals, such as their gender and age, as well as their aspirations, social status, and personality, in addition to an assortment of other psychological and social factors. (Flek, 2020). In light of this, the phrase serves as the foundation for the interpretation of text sentiment.

2.2.4 An Examination of Sentiment at the Aspect Level

This level conducts an in-depth examination to determine the sensations that are associated with the different characteristics of entities. Take, for example, the sentence that states," The screen resolution of the iPhone 11 is exceptional."" screen," which is a feature of the entity" iPhone 11," is the subject of this review, and the evaluation is positive. Therefore, because of this, the task at this level makes it easier to precisely identify the preferences and aversions of people (Indurkhya and Damerau, 2010). It does not focus on analysing the mood of paragraphs or words but rather places an emphasis on the properties of objects (for example, the attributes of a product). Among the essential tasks of sentiment analysis is the extraction of aspects, which encompasses both implicit and explicit aspects, as stated in (Tubishat et al., 2018). An analysis of implicit aspect extraction methodologies was presented by the authors, who looked at the topic from a variety of angles.

2.3 The Gathering of Information and the Selection of Features

2.3.1 Data Gathering

As seen in figure 2.2, the process of gathering information from the internet may be achieved using a wide range of techniques. These techniques include online scraping, social media, news channels, e-commerce websites, forums, weblogs, and a multitude of other websites. When doing sentiment analysis, the first step is to gather data, which is the beginning of the process.

(a) Web scraping is the act of automatically getting data from websites, which may hold a substantial number of important information such as product details. This information may be obtained via the process of web scraping. The process is referred to as" Internet scraping," which is the word used to describe it. On the other hand, this information may be used for a variety of reasons, one of which is the examination of behavioral states that are characterized by psychological states. Among the many internet scraping programs or services that are available for free download, ParseHub is only one example among many others (Birjali et al., 2021).

(b) The process of creating data or annotations may be outsourced via the use of crowdsourcing technique. It is possible to create a significant quantity of data in a short period of time, which may be of tremendous benefit. This can be accomplished rather quickly. Amazon Mechanical Turk is a well-known site that is used for crowdsourcing, as pointed out by (Birjali et al., 2021) in his detailed article.

(c) Punctuation marks, which are often commonly referred to as exclamation marks, serve this function particularly when they are used to emphasis the enthusiasm of a positive or negative message. The question mark and the apostrophe are two more punctuation marks that are used in the same manner as is described above.

(d) phrases that are used in slang, include” lol” and” rofl,” amongst others. They are a method that is used rather often when one is aiming to infuse a sense of humour into a statement. It is fair to assume that a slang phrase in the text gives evidence of sentiment analysis. This is because opinion tweets are of a certain sort, and it is reasonable to conclude that this term provides proof. In order to better convey their meaning, the definition of the slang term should be modified (Wankhade et al., 2022).

(e) The goal of punctuation marks, which is comparable to the function of exclamation marks, is to emphasise the strength of a statement, whether it be a good or negative one. Not only is the apostrophe included in this category, but the question mark and the apostrophe are also included as supplementary punctuation marks (Wankhade et al., 2022)

2.3.2 Data Input

As a consequence of the growth of Web 2.0, a number of different types of data that were previously unavailable are now available. For the purpose of achieving enhanced sentiment categorization, it is conceivable for research domains that concentrate on sentiment analysis to take use of this version. As a result, the input of a SA system is a collection of papers or media files that are in a range of forms (Poria et al., 2017). Some examples of these formats include the following:

(a) A plain text file, often known as a TXT file, is a file that contains text that has not been properly formatted

(b) Comma-separated values, often known as CSV files, are plain text files that use commas to provide a separation between data.

(c) Extended Markup Language, abbreviated as HML A hierarchical text format that differentiates content via the use of individualized tags is known as a language file.

(d) Serialization and transmission of structured data via a network connection are both possible with the help of the JASON data format.

(e) Markup language, sometimes known as HTML, is a language that is used to define the structure of a web page

2.4 Data Pre-processing

When taken as a whole, the information that is compiled from a wide range of sources, most notably social media, is not organised. Furthermore, it is probable that the unprocessed form of this data has a significant quantity of noise, in addition to a considerable number of typographical and grammatical problems (Liu, 2022). Taking this into consideration, it is of the utmost importance to clean and pre-process the text before undertaking any type of analysis. Preprocessing is a procedure that is performed with the intention of enhancing analysis and decreasing the complexity of the data that is being entered. Eliminating superfluous words that do not contribute to the overall polarity of the text is the method by which this objective might be accomplished. Articles, prepositions, punctuation, and special characters are all examples of words that fall under this category. In Table 1, you will find a selection of tools that are accessible to the general public and may be utilised for a wide range of preprocessing and natural language processing operations. A comparative analysis of a number of natural language processing toolkits was carried out by (Pinto et al., 2016). This analysis was conducted within the context of both formal and social media writings. The entire process consists of a variety of actions that are performed on a regular basis, including the following:

(a) The process of tokenisation Tokens are the smaller components that are created by this process. For example, a document may be broken down into phrases, and a sentence can be broken down into words.

(b) The removal of stop words refers to the elimination of phrases (such as ”the,” ”for,” and ”under”) that do not often contribute to or improve analysis. As a result, these terms are removed in advance. act of identifying various structural components of a text, including as verbs, nouns, adjectives, and adverbs, is referred to as part-of-speech (POS) tagging. You may also hear this technique referred to as POS tagging.

(c) Lemmatisation is the process of reducing a word to its most basic form. This process is also known as” dialectical reduction.” Lemmatisation is a process that is comparable to stemming; however, it preserves word-related information, such as identifiers for parts of speech, in its whole.

The kind of data also plays a role that is being supplied, the pre-processing step could be different from one particular instance. Two examples of the extra processing and cleaning methods that are required for some formats are the extension of abbreviations and the elimination of repeated letters, such as the ”I” in ”liiiiiiike”. Both of these procedures are essential for certain formats. According to what was said before, textual data may be rather noisy; hence, in order to do sentiment analysis more effectively, it is necessary to follow two fundamental processes. This process consists of two steps: the extraction of features and the selection of features, both of which will be detailed in the next paragraph for your comprehension (Birjali et al., 2021).

2.5 Sentiment Analysis Techniques

In addition to being a dynamic and flourishing topic of research, sentiment analysis is also a subject that has the potential to be used in a wide range of sectors. For this reason, academics are always going through the process of proposing, evaluating, and comparing a wide variety of different approaches. Improvements in the performance of sentiment analysis and the development of solutions to the problems that are encountered in this industry are the goals of this project.

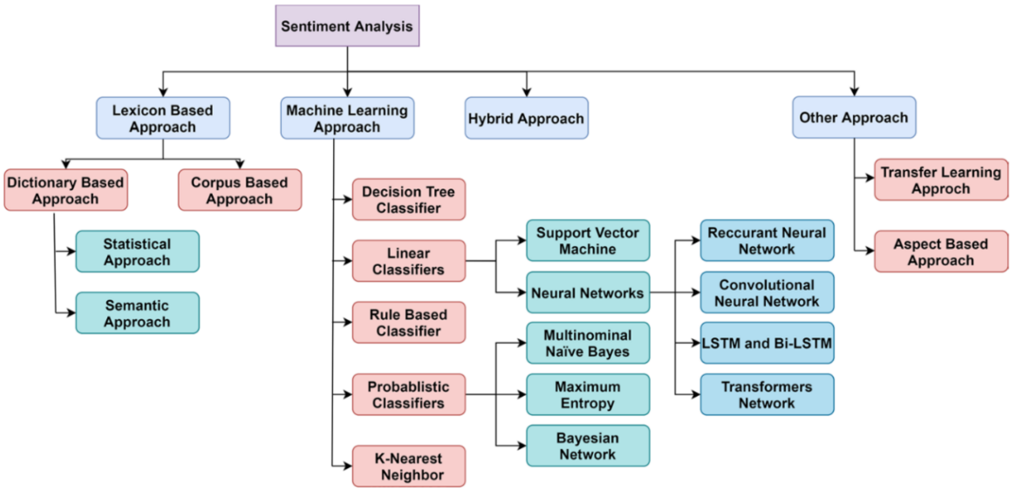

Figure 2.3 Approach of sentiment analysis.

2.5.1 Machine Learning Approach

Through the use of both training and testing datasets in combination with machine learning methods, it is possible to achieve the categorization of the polarity of sentiment, which includes categories such as negative, positive, and neutral. There are a variety of categories that may be applied to these approaches, including supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning, to name just a few. These are only some of the many possible classifications. It is necessary to employ the supervised method in circumstances when the problem of the classification is involving a specified collection of classes. On the other hand, in situations when there is a dearth of data that has been labelled, the unsupervised method could present itself as the most successful strategy. Unlabelled datasets that contain a subset of labelled samples are a good candidate for the semi-supervised technique, which may be utilized in such situations. It is the goal of reinforcement learning algorithms to equip the agent with the capacity to interact with its environment in order to maximise the accumulation of rewards. These algorithms make use of methods that incorporate trial and error in order to accomplish this goal. It is feasible for machine learning algorithms to recognise domain-specific patterns from text, which may lead to improved classification results. This is a possibility. In spite of this, these approaches often require enormous training datasets in order to obtain the best possible performance. It is of utmost importance to take note of the fact that a classifier that has been trained on a particular dataset does not exhibit the same degree of efficacy when it is applied to a different domain (Pathak et al., 2020).

2.5.1.1 Supervised Learning

It is necessary to have training papers that are labeled for supervised techniques, with the labels often indicating the classes (for example, positive, neutral, and negative). Linear, probabilistic, rule-based, and decision tree classification techniques are the four types that fall under the umbrella of supervised classification technique (Yusof et al., 2015). Following this, the remaining subsections will offer a quick description and comparison of the most common supervised classification algorithms that are used for sentiment analysis

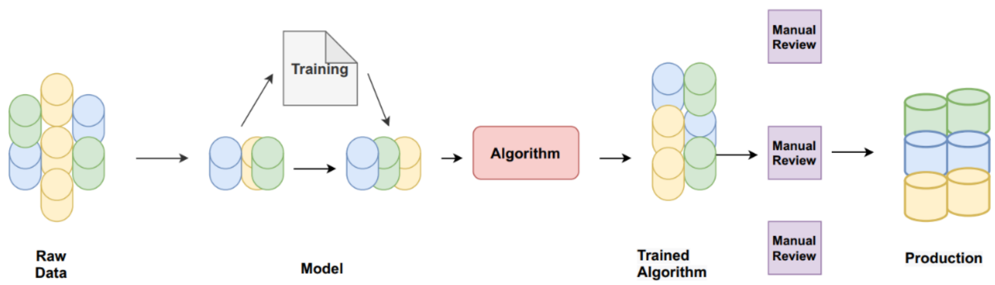

Figure 2.4 Sentiment procedure for supervised machine learning [Wankhade et al., 2022]

When it comes to the topic of text categorisation, Naïve Bayes (NB) is a fundamental classifier that is often utilised as one of the strategies that is employed the most frequently. Within the framework of the model, the Bayes Theorem is utilised, and the Bag of Words feature extraction approach is utilised in conjunction with it. Because of this, the location of a word within the text is not taken into consideration, and the presence of a certain word is not dependent on the presence of any other words. Consequently, this is the case. With the application of Bayes' theorem, Naïve Bayes assigns a document (d) to the category (c) that maximises the probability of P(c/d) in the most effective manner

Table 2.1 SV MaA Neural Networks Comparison with One Another

Classifier	**Advantages**	**Disadvantages**
SVM	When operating in high dimensional spaces, efficiently and reliably. In comparison to other machine learning algorithms, it is simpler to train and achieves a high level of accuracy. memory-efficient as a result of the advantages that kernel mapping to high-dimensional feature spaces provides. When operating in high-dimensional spaces, efficiently and reliably. In comparison to other machine learning algorithms, it is simpler to train and achieves a high level of accuracy. memory-efficient as a result of the advantages that kernel mapping to high-dimensional feature spaces provides.	In situations when the amount of features greatly exceeds the number of samples, suboptimal performance is more likely to occur. It is very necessary to choose a kernel function that is an appropriate fit. Because there is no probabilistic reason for the noncategorical, there is a limited amount of interpret ability of the data.
ANN	Capable of managing complex relationships among variables and achieving superior generalization, especially with noisy data. Efficient for high-dimensional problems with rapid execution times.	Theoretically complex and challenging to implement. Requires substantial memory and significant training time compared to other methods. May necessitate a large dataset for optimal performance.

$$P(c|d) = \frac{P(c)p(d|c)}{p(d)} \tag{2.1}$$

In this equation, the term" P(c)" refers to the prior category of the probability c," P(d|c)" represents the conditional likelihood of document (d) being classified into category c, and" P(d)" represents the prior of the document to be classified into category d. All of these terms are used interchangeably. A method known as Naïve Bayes is capable of computing the posterior probability of a class by utilising the word distribution that is present within the text. In order to achieve this goal, it is necessary to make the assumption that particular feature needs are not reliant on one another. According to one possible interpretation, the equation that was discussed before might be rewritten as follows:

$$P(c|d) = \frac{P(c)p(\omega_1|c)^{*}\ldots\ldots{}^{*}\mathrm{P}(\omega_{\mathrm{n}}|\mathrm{c})}{p(d)} \tag{2.2}$$

The Naïve Bayes method was utilized in a multitude of studies as a classification software. (Hasan et al., 2015) has developed a classifier that utilities the Naïve Bayes approach to classify opinions expressed in both English and Bangla. This classifier achieves a remarkable level of accuracy. In addition, they utilized their classifiers to conduct an analysis of a number of random evaluations and tweets, and they achieved remarkable outcomes in the majority of the cases.

Rule-based approach Taking this into consideration, the classifiers that are utilised in this technique are dependent on a certain set of criteria in order to carry out the process of sentiment categorisation. As an illustration, let's consider a rule that can be stated as LHS → RHS. Disjunctive Normal Form (DNF) is used to articulate the feature set, while the left-hand side (LHS) indicates the antecedent of the rule or a collection of conditions that apply to the feature set. At the opposite end of the spectrum, the right-hand side (RHS) of the rule is a representation of the conclusion or outcome (class label) of the rule, which is dependent upon the left-hand side (LHS) being met (Tung, 2009). The efficacy of rule-based classifiers is equivalent to that of decision trees, and they are able to categorise new examples in a very short amount of time. An additional advantage of utilising the rule-based method is that it allows for the prevention of over-fitting, which is a problem that might arise. If, on the other hand, there are an excessive number of regulations, then the interpretation of those rules becomes extremely challenging and laborious. Furthermore, it has a performance that is below average when dealing with noisy data. This is a significant limitation.

In the field of classification, rule-based classification refers to a strategy that employs IF-THEN rules with the purpose of predicting the outcomes of classes (Tung, 2009). The classifiers that are used in this method are thus dependent on a predetermined set of criteria in order to carry out the process of sentiment categorization. LHS → RHS is one possible expression for a rule. In Disjunctive Normal Form (DNF), the left-hand side (LHS) represents the antecedent of the rule or a set of conditions related to the feature set. On the other hand, the right-hand side (RHS) indicates the conclusion or result (class label) of the rule based on the fulfilment of the label on the left-hand side (LHS) (Sankar and Subramaniyaswamy, 2017).

Classifiers that are based on rules are able to quickly categorise new cases, and their efficiency is equivalent to that of decision trees. The ability of the rule-based approach to avoid overfitting is yet another advantage of using this methodology. In spite of this, its interpretation becomes difficult and complicated when there are a great number of regulations to consider. Furthermore, when it is presented with noisy data, it displays performance that is not enough.

2.5.1.2 Semi – Supervised Learning

In situations when it is difficult to collect labelled data, semi-supervised learning (SSL) techniques are used. However, in contrast to unsupervised approaches, this strategy makes use of a limited quantity of initial labelled training data in order to have an effect on the process of feature learning. As a consequence of this, it is situated somewhere in the centre of the continuum that separates supervised and unsupervised methods. Utilizing SSL techniques allows for the efficient use of large amounts of inexpensive unlabelled data in a cost-effective manner. These approaches save a considerable amount of time and effort while also producing a classifier that is capable of strong generalization with a greater quantity of data that has been labelled. (Zhu et al., 2013). (Hussain and Cambria, 2018) offered a technique to learning that is semi-supervised and was designed for the goal of analysing Big Social Data. The use of support vector machines (SVM) and random projection scaling in combination with one another is what contributes to the realization of this possibility. It seems that this semi-supervised model has the potential to considerably increase the performance of some natural language processing tasks, such as sentiment analysis, based on the data that has been collected. Research that was carried out not too long ago on SSL-based sentiment analysis may be broken down into five main categories: generative, co-training, self-training, graph-based, and multi-view learning. (Xia et al., 2015).

Generative approach In order for this method to be able to compute the parameters of each distribution, there must be at least one data point identified as belonging to each category. (Han et al., 2020). This technique is predicated on the premise that data from a variety of categories correspond to a variety of distributions. Following the training of the model for each class, a generative model will generate distributions for the inputs. Bayes' theorem will then be utilised in order to make a predictive prediction about the label (class) of a test input. After the model has been trained for each possible class, this step is carried out. (Mesnil et al., 2014) This article presents a straightforward yet reliable ensemble approach for sentiment analysis. A tree complementing technique and theoretically baseline models are both incorporated within this strategy. The usage of a generative process was really implemented in one of them. When the entire system is applied to the dataset that is comprised of IMDB movie reviews, it achieves a performance that is unparalleled and at the cutting edge of industry standards. The fact that this is the case suggests that ensemble learning may be utilised in circumstances that are either semi-supervised or in which there is no supervision that is present.

Self-training approach**.** The concept of self-training may be broken down into two distinct stages. A little amount of labelled data is used to train the classifier in the beginning stages of its training. In the succeeding step, the trained classifier is used to classify unlabelled data, therefore including the samples with the highest level of confidence into the initial training set as new labelled data (Gao et al., 2014). Iterative execution will be used to carry out the final phase, which will include the newly tagged data. After that, the model that was produced is evaluated by making use of the test data. The approach in question has been used widely within the field of sentiment analysis, as seen by the citation (Hajmohammadi et al., 2015). It was presented by (He and Zhou, 2011) that a new framework was developed by using a self-training approach. This methodology gets information from labelled features rather than labelled instances itself. As a result of the outcomes of the experiment, it may be concluded that their approach outperformed various recognized methodologies.

2.5.1.3 Reinforcement Learning

One of the approaches to machine learning is called reward learning, which is sometimes referred to as reinforcement learning (RL). In this approach, an agent is rewarded in the subsequent time step depending on the evaluation of the activity that it has completed in the time step before it. Reinforcement learning algorithms make use of methods that include trial and error in order to enable the interaction of the agent with its environment in order to maximize the cumulative rewards. This makes it possible for the algorithms to maximize the cumulative rewards. (Li et al., 2020). Among the many challenges that have been addressed via the use of reinforcement learning, robotic control is one of the issues that has been addressed. The majority of its uses, on the other hand, have been in the gaming business. (Wan and Gao, 2015). On the other hand, despite the fact that it has the capability to tackle complicated tasks, particularly with the incorporation of Neural Networks, the application of this technology to solve problems associated with sentiment analysis is still rather limited. One of the most significant benefits of this approach is that it is fairly comparable to the way in which people learn, which is a characteristic that is greatly sought after in the field of study pertaining to sentiment analysis. throughout the training phase, mistakes that were made throughout the decision-making process are corrected through the use of reinforcement learning, which makes use of previously learnt experiences in order to enhance decision-making and move closer to optimality. In the other way, the process of constructing the model for reinforcement learning might be difficult to do. Furthermore, reinforcement learning necessitates a substantial quantity of data as well as a substantial amount of processing work when it is implemented.

Convolutional neural networks (CNN) Computer vision was the principal use of this design, which is a feed-forward that is a subset of neural networks. Nonetheless, Recent research has shown that it is successful in a number of different areas, including natural language processing and recommender systems, among others. A convolutional neural network (CNN) consists of three types or layer: an input layer, an output layer, and a hidden layer made up of several convolutional, pooling, normalization, and fully connected layers. All convolutional neural networks (CNNs) have these three layers. Convolutional layers filter inputs (such as word embeddings in text sentiment classification) to extract features. The resolution of the features is decreased by pooling layers, which makes feature recognition more robust against noise and changes of a small magnitude. To provide better convergence during training, the normalization layer standardizes the output of the previous layer and works in tandem with the fully linked layers to perform the classification job. The area of sentiment analysis has seen a rise in the use of Convolutional Neural Networks, or CNNs, in the past several years. (Zhang and Wallace, 2015) presented a CNN model for sentiment analysis that has gained widespread recognition by the scientific community. The author conducted an evaluation of a CNN model that was built using pre-trained word2vec for the goal of categorizing sentiments at the sentence level. In addition to demonstrating that pre-trained word embeddings have the potential to serve as excellent features for natural language processing applications that make use of deep learning, the model outperformed other strategies.

2.5.2 Lexicon – Based Approach

Two fundamental methodologies are employed for sentiment analysis. One of these techniques is the Lexicon-Based approach, commonly known as the knowledge-based approach. A lexical resource termed an opinion lexicon is essential; it is a predetermined compilation of words that associates terms with their semantic orientation, categorizing them as negative or positive according to assigned ratings. (Hu and Liu, 2004). It is possible for a score to represent a fundamental polarity value, such as plus one, minus one, or zero, which correspond to words that are positive, negative, or neutral, respectively, or a value that shows the degree or intensity of feeling conveyed by the individual. Both of these values are feasible. The procedure that is utilized in order to discover the ultimate orientation of the composition is the computation of the semantic orientation values of the component words that make up a text. The text is disassembled into its component pieces, which can be either individual words or microphrases, and the lexicon's emotion values are assigned to each of these parts. The microphrases and individual words include the components. One method that may be utilized in order to determine the overarching feeling that is communicated by a piece of literature is the utilization of a formula or algorithm, such as the summing and averaging method

When it comes to doing sentiment analysis, the lexicon-based strategy is a very effective method that can be utilised at both the sentence and feature levels. Taking into consideration the fact that it does not require any training data, it is conceivable to categorise it as an unsupervised technique. On the other hand, the most significant issue with this strategy is that it is dependent on the domain. This is because words may have several meanings and interpretations; hence, a statement that is judged to be successful in one domain could not have the same connotation in another domain. This is because words can have multiple meanings and interpretations. For instance, when we analyse the word" small" in connection with the words" The TV screen is too small" and" This camera is very small," the word" small" in the first sentence assumes a negative meaning due to the fact that people have a tendency to favour screens that are bigger. However, the word" small" has a positive meaning in the second phrase, which implies that a tiny camera is suitable for transportation. This is because the word" small" conveys a positive connotation. On the other hand, the deployment of a lexicon adaption strategy or the construction of a sentiment lexicon that is specific to a domain are also viable answers to this problem. (Sanagar and Gupta, 2020a) proposed a strategy for changing an emotion language that is unique to a certain genre of writing. This unique technique takes use of unlabelled data in order to construct sentiment lexicons for both the source domain and the destination domain. This is in contrast to standard implementations of adaptation, which are dependent on data that has been labelled. As demonstrated by the research carried out by (Sanagar and Gupta, 2020b), transfer learning strategies have the potential to be utilised in order to acquire new lexicons that are specialised to a certain domain. A methodology for unsupervised sentiment lexicon learning that is adaptable to new domains within the same genre was proposed by the authors.

Following the acquisition of polarity seed words from corpora spanning many source domains, the information that is particular to the genre is then transferred to the target domains. The performance of the lexicon-based technique is worse in comparison to the performance of the machine learning approach when a significant training dataset is available. In the following paragraphs, we will discuss the three basic approaches to the process of constructing and annotating sentiment lexicons.

2.5.2.1 Manual Approach

Through the use of the manual technique, human intervention is required in order to annotate the vocabulary. Within the process of developing sentiment lexicons, there are two stages: the first is the compilation of a list of words that carry a certain feeling, and the second is the assignment of sentiment labels to these words. Despite the fact that this process is often laborious, costly, and time-consuming, it has the potential to produce a vocabulary that is reliable and consistent. In order to speed up this process, it is possible to develop and deploy an automated method. In order to limit the number of errors that occur, a manual approach is used as a benchmarking process. A great number of lexicons have been compiled by the use of physical labour. At the same time as (Wilson, 2005) produced the MPQA Subjectivity Lexicon, (Taboada et al., 2011) established the Semantic Orientation Calculator (SO-CAL). Both of these lexicons are dependent on collections of negators and intensifiers that have been manually selected.

Crowdsourcing and gamification are two more methods that researchers could use. The technique of bringing together a group of people to work towards a common goal via the use of online platforms is known as crowdsourcing. The term" gamification," on the other hand, describes the process of incorporating components of games into situations that are not games. Tower of Babel is a game that was designed by (Hong et al., 2013) with the intention of encouraging players to attach emotion polarity to words in order to establish a sentiment lexicon.

2.5.2.2 Dictionary – Based Approach

This method is based on the idea that words that are synonymous have the same sensation polarity, but those that are antonymous have the opposite degree of polarity.

Established dictionaries, such as WordNet9 (Miller et al., 1990) or thesauri (Mohammad et al., 2009), are used in the construction of the sentiment lexicons that are utilized in this technique. Manual compilation is used to create a collection of initial seed words that have a preset orientation. In the second step, the vocabulary is expanded by investigating synonyms and antonyms with the use of new lexical resources. Until there are no more words to be found, the newly discovered words are gradually added to the list that is already in existence (Wankhade et al., 2022). There is the possibility of doing an additional manual review in order to correct and remove errors. SentiWordNet 3.0 is a well-known lexicon that was constructed by (Baccianella et al., 2010). This innovative lexicon was created by the automatic annotation of all synsets in WordNet 3. A method for constructing a thesaurus lexicon was provided by (Wankhade et al., 2022) who suggested using three online dictionaries as the tools. The polarity lexicon was investigated by the writers from two different points of view. The first step that they took was to develop the first techniques for developing a polarity lexicon in the first aspect. With regard to the second component, the authors presented relevant information on the open-source polarity lexicon that is available to the public. The end of the article outlined the challenges that are still being faced in the field of research as well as potential paths that may be taken in order to improve polarity lexicons.

The inability of dictionary-based and all lexicon-based techniques to recognize sentiment words with domain-specific connotations is the key problem with these approaches. As a result, these methodologies are not ideal for context and domain-specific classification when it comes to identifying sentiment words. In addition, the compilation of dependency rules is a difficult and labour-intensive process. On the other hand, this method is not computationally intensive as long as there is no dataset training involved. Furthermore, it is an efficient strategy for rapidly constructing a lexicon that includes a significant number of sentiment words and their orientations.

2.5.2.3 Corpus – Based Approach

Techniques that are based on a corpus, as opposed to those that are based on a dictionary, make use of syntactic or co-occurrence patterns in order to identify new emotion words that have their preferred orientation within a large corpus. Additionally, linguistic constraints on connectives (such as AND, OR, and BUT) are used in order to recognize additional emotion terms. When a conjunction is used to link two adjectives, such as in the phrase" simple AND easy," the orientation of the adjectives is often the same. In addition to maintaining consistency in emotion, rules may be developed for these connectives; nevertheless, it is possible that these rules will not always be consistent in practice. Methods such as clustering may be used after the operation has been completed in order to generate sets of emotion words, which may include both positive and negative phrases (Liu and Liu, 2011). (Hatzivassiloglou, 1997) were the ones who originally presented this method to the public. The writers chose words that appeared in the pattern W_1 and W_2 and had the same orientation in order to widen the initial collection of often recurring adjectives with their orientation. This was done in order to broaden the scope of the collection. The researchers used a network containing words as vertices and their pairings as edges, as well as a log-linear model, in order to identify whether or not two conjoined adjectives had opposing orientations and to categories them into positive and negative terms. Corpus-based techniques are straightforward; but, in order to recognize the polarity of words and the emotions they convey in text, they need a large dataset (Agarwal et al., 2016). Numerous strategies that are based on corpora are often categorized as either statistical or semantic approaches (Vyas and Uma, 2019), as will be detailed in the subsequent subsections.

2.6 Statistical Approach

Semantic approach in contrast to the previous method, the ontology-based approach measures word similarity and gives the same emotion value to semantically comparable phrases (Araque et al., 2017). Typically, this technique searches sentiment dictionaries for synonyms, antonyms, and related terms to expand a vocabulary and analyses sentiment, as shown by (Zhang et al., 2012). The authors developed Weakness Finder, an expert system that analyses Chinese reviews to identify product weaknesses using a statistical and semantic methodology. (Dong and Dong, 2006) lexicon was utilized to determine word similarity. Experimental findings showed that the suggested expert system performed well.

2.7 Sentiment Analysis Challenges

2.7.1 The Detection of Sarcasm

Sarcasm may be defined as" the act of speaking in a manner that is intended to ridicule someone or communicate displeasure." This is one definition of sarcasm out of many. In the Macmillan English Dictionary, this term may be found. Sarcasm may also be defined as the act of conveying the opposite of what one is actually attempting to say. This is another definition of sarcasm.

2.7.2 Negation Handling

In the field of sentiment analysis, it is essential to properly manage negation phrases like not, neither, nor, and so on since these terms have the potential to reverse the polarity of a sentence. For instance, the phrase" This movie is good." is considered to be a positive statement, but the statement" The movie is not good." is considered to be a negative statement (Wankhade et al., 2022).

Regrettably, in certain methods, negation words are not included because they are included in Stop-Word lists or because they are implicitly eliminated due to their neutral emotion value in a lexicon. This does not have any effect on the final polarity. When compared to other techniques, which do not omit negation concepts, this way of thinking is different. On the other hand, carrying out this function by reversing the polarity is challenging due to the fact that negation words may be inserted in a phrase without having any impact on the sensation that is communicated by the text.

According to (Lazib et al., 2020), In order to accomplish the task of negation scope detection, it is recommended that a hybrid neural network that makes use of a syntactic route approach be utilised. The CNN model is utilised to extract pertinent syntactic characteristics between the token and the cue along the shortest syntactic path in both constituency and dependency parse trees. On the other hand, the Bi-LSTM model is utilised to gain contextual representation throughout the entirety of the sentence in both forward and backward orientations. The bidirectional LSTM and CNN are both incorporated with this technology. It is common practice to employ both of these models in tandem with one another. 90.82 percent out of a possible 100 was the F-score that their model was able to obtain.

2.7.3 Spam Detection

The detection of spam is a phase that is of utmost significance in the process of sentiment analysis. The opinions that are published on the internet have a significant impact on the decisions that customers make regarding their purchases; thus, spam and fake reviews have the ability to damage the reputations of companies and affect the impressions that customers have regarding products, services, enterprises, or other affiliations. (Cardoso et al., 2018). A tremendous obstacle is presented by the absence of obvious contrasts between the assessments, which makes it difficult to develop a spam detection system that is able to differentiate fake reviews from a large number of legitimate reviews. An approach for detecting spam that was developed by (Saumya and Singh, 2018) makes efficient use of three characteristics: the mood of the review and the comments that accompany it, content-based variables, and rating deviation. In this method, the review is classified as either spam or non-spam based on the collection of comment data.

2.7.4 Low – Resource Language

The bulk of research in the field of sentiment analysis has focused on the English language or other languages that have a sufficient number of linguistic resources, such as sentiment lexicons and labelled text corpora. This is because the English language is the medium through which sentiment analysis is conducted. For the purpose of sentiment analysis, the most common approach is to employ supervised learning algorithms. On the other hand, these approaches are highly dependent on linguistic resources, which may be rather costly to obtain for languages that are not as widely spoken [Ren et al., 2014]. When it comes to linguistic resources, languages that are considered to be low-resource languages (or under resourced languages) are those that are lacking in resources. There are a few different approaches that may be taken to solve this issue: The process of creating a language resource from the ground up by using unsupervised, semi-supervised, and transfer learning approaches.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

Using natural language processing, this chapter provides an illustration of the general framework that is used in the process of doing research for the purpose of performing sentiment analysis on evaluations of Amazon office products. Everything from the preliminary investigation of the sentiment analysis to the assessment and comparison of the models is included in the research process. In this chapter, the data that were employed and the types of models that were applied will be identified and illustrated.

3.2 Research Framework

In order to accomplished the sentiment analysis in its entirety, the research was carried out in 4 distinct phases. The completion of each phase brought about the achievement of a crucial milestone. The following processes are provided in the order that they are shown: data collection, data preparation and exploratory data analysis, VADER sentiment analysis and Reberta analysis, and model assessment and comparison correspondingly.

There is an illustration of the process of the work in figure 3.1. There will be a discussion of each step in the sub section that follows.

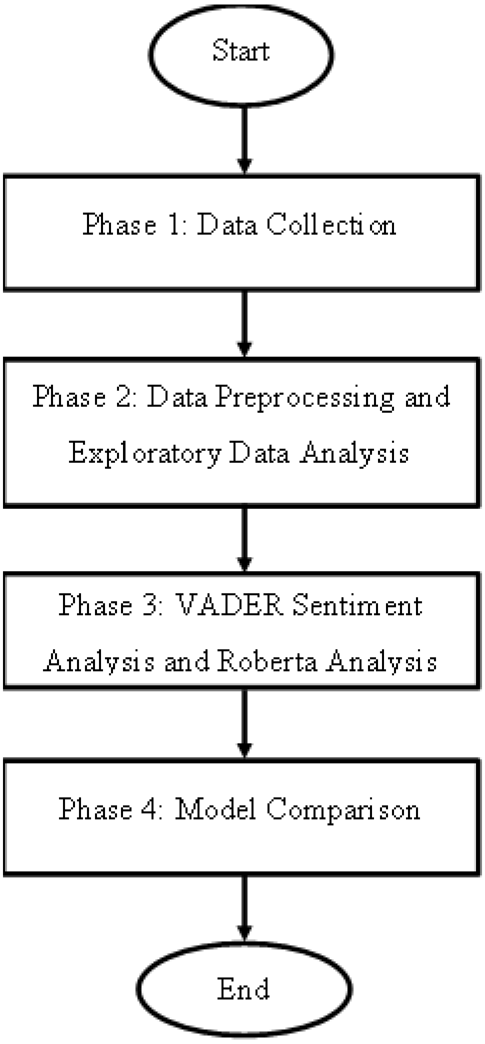

Figure 3.1 Overall research methodology

3.2.1 Data Collection

A dataset consisting of reviews of office products was gathered from the Amazon Reviews Repository collection [Amazon Reviews'23]. There are a total of 200,000 rows of data that represent each unique review, and there are eight columns that include the following information: parent asin, user id, helpful vote, asin, review, timestamp, verified buy, and rating. It offered a valuable perspective on the total pleasure of purchasers.

	parent_asin	user_id	helpful_vote	asin	text	timestamp	images	verified_purchase	title	rating
0	B01MZ3SD2X	AFKZENTNBQ7A7V7UXW5JJI6UGRYQ	0	B01AHHL4X2	Lovely ink. Writes well. The right amount of w...	1677939345945	[]	True	Pretty & I love it!	5.0
1	B08L6H23JZ	AFKZENTNBQ7A7V7UXW5JJI6UGRYQ	0	B08L6H23JZ	Overall I'm pretty happy with this purchase bc...	1677939160682	[]	True	2 excellent 1 extremely dry (blue)	4.0
2	B07JDZ5J46	AFKZENTNBQ7A7V7UXW5JJI6UGRYQ	2	B07JDZ5J46	[[VIDEOID:63276c19932aa4f3687042b8b9f8613c]] U...	1660188831933	[]	True	I don't get the reviews. Mine are garbage.	1.0
3	B07BR2PBJN	AFKZENTNBQ7A7V7UXW5JJI6UGRYQ	0	B004MNX7EW	It's a beautiful color, but even though it had...	1659806066713	[{'small_image_url': 'https://m.media-amazon.c...	True	Ordering Ink online: never a good idea I guess.	4.0
4	B097SFY5ZS	AFKZENTNBQ7A7V7UXW5JJI6UGRYQ	0	B019YLRFFS	Idk if I just got a bad batch which is possibl...	1659799390978	[]	True	Mine are iffy at best.	3.0

Figure 3.2 Review Dataset

3.2.2 Data Preprocessing and Exploratory Data Analysis

Due to the fact that that particular user-provided product review data is absent from the dataset that we are using, it is not possible to conduct an assessment using this sample. As a consequence of this, it is imperative that they be removed. Additional, in the event that the ratings have values that are missing. In light of this, it is able to either substitute NaN with the average of the other samples or get rid of those samples altogether. On account of the large number of samples that we have, it is possible to remove them.

When doing any kind of data analysis that is associated with text, the first step is to clean the text data. This is done in order to establish some straightforward methods that can be used to clean and prepare text data for modelling and machine learning.

(a) lowercase
(b) remove whitespaces
(c) replace digit with spaces
(d) replace punctuations with spaces
(e) remove extra spaces and tabs

Additionally, First, the text should be tokenized into words or sub-word units, and then stemming should be used in order to standardize word forms. Next, after cleaning.

3.2.3 VADER Sentiment Analysis and Roberta Analysis

To begin, make use of VADER, which stands for Valence Aware Dictionary and sentiment Reasoner. This tool offers a speedy dictionary-based method to determine the polarity of sentiment and provide compound scores that range from -1 (the most negative) to +1 (the most positive). This technique is especially useful for analyzing shorter text samples, and it provides an easy threshold-based classification (for example, positive, negative, and neutral). However, in order to capture language and context with a greater degree of detail, I additionally fine-tune a Roberta model. Roberta, which is an improved transformer-based framework, not only has a deeper understanding of the links between words, but she also has a more precise ability to manage small alterations in mood. The combination of these two approaches provides me with a time-efficient, rule-based sentiment analysis alternative known as VADER, as well as a strong context-aware model known as Roberta, which has the potential to greatly increase classification results.

3.2.4 Model Evaluation and Comparison

After the implementation of natural language processing, the accuracy, precision, and recall of the model may be evaluated. The F1-score is a metric that compares the model's capacity to recognize all positive data to the harmonic mean of precision and recall. This is done to determine how effectively the model is able to categorize reviews as either positive, neutral, or negative. And last, choose the model that is the most correct.

3.3 Chapter Summary

The methodology of the research is broken out in great detail in this chapter, beginning with the gathering of data and continuing with the evaluation of the categorization model and comparison. This procedure guarantees that the procedure of doing sentiment analysis of the office product reviews is carried out in a methodical and data-driven manner.

CHAPTER 4

INITIAL RESULTS

4.1 Overview

This chapter discusses the preliminary results and sentiment analysis of the Amazon office products. This chapter starts with the identification of the data set which is the exploratory data analysis (EDA), followed by VADER (Valence Aware Dictionary and sEntiment Reasoner) and Reberta (A Robustly Optimized BERT Pretraining Approach) analysis sentiment analysis, the creating and implementing the model using machine learning techniques.

4.2 Exploratory Data Analysis

EDA, which stands for exploratory data analysis, is one of the strategies that may be used to learn every information about the dataset. The patterns, trends, and relationships that are present within the data are included here. Before beginning to construct the machine learning method, it is quite helpful to have a solid grasp of the data structure already in place.

The text column describes the buyer reviewer of the product on Amazon platform in the office products category. Then the reviews will be analysed to obtain the results of sentiment analysis of the office products whether it is positive, negative or neutral.

One example of the raw dataset that was loaded from Python is shown in Figures 4.1 and 4.2, which can be found as follows. In all, there are 200000 rows of data, and there are 10 columns.

parent_asin	user_id	helpful_vote	asin	text	timestamp	images	verified_purchase	title	rating
B095CPWNTQ	AFITLXUBYKIELXW4EEA7IT5KEQQQ	0	B007D930YO	Was easy to setup and it work good with google...	1525408368407	[]	True	Was easy to setup and it work good with google...	5.0
B00BUV7C9A	AFITLXUBYKIELXW4EEA7IT5KEQQQ	0	B00BUV7C9A	Works as good with GV and works as good as Oom...	1441344761000	[]	True	Five Stars	5.0
B00006IEI4	AEB2U6KK3TFESGJY2PAHYW3M2QAQ	0	B00006IEI4	Works great!	1477875005000	[]	True	Five Stars	5.0
1604189274	AEB2U6KK3TFESGJY2PAHYW3M2QAQ	0	1604189274	Great quality, has many different uses.	1477874895000	[]	True	Five Stars	5.0
B09B1PNJ9Q	AEB2U6KK3TFESGJY2PAHYW3M2QAQ	0	B003F189HM	Love putting my students birthdays on these!	1477874814000	[]	True	Five Stars	5.0

Figure 4.1 Dataset

timestamp	images	verified_purchase	title	rating
1677939345945	[]	True	Pretty & I love it!	5.0
1677939160682	[]	True	2 excellent 1 extremely dry (blue)	4.0
1660188831933	[]	True	I don't get the reviews. Mine are garbage.	1.0
1659806066713	[{'small_image_url': 'https://m.media-amazon.c...	True	Ordering Ink online: never a good idea I guess.	4.0
1659799390978	[]	True	Mine are iffy at best.	3.0

Figure 4.2 Dataset

In figure 4.3 shows the dataset information of each column also the type of the data that used. It can be seen that all columns are non-null, consisting of 6 objects, 3 int64, 1 bool, and 1 float64.

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200000 entries, 0 to 199999
Data columns (total 10 columns):
 #   Column             Non-Null Count   Dtype
---  ------             --------------   -----
 0   parent_asin        200000 non-null  object
 1   user_id            200000 non-null  object
 2   helpful_vote       200000 non-null  int64
 3   asin               200000 non-null  object
 4   text               199975 non-null  object
 5   timestamp          200000 non-null  int64
 6   images             200000 non-null  object
 7   verified_purchase  200000 non-null  bool
 8   title              199964 non-null  object
 9   rating             200000 non-null  float64
dtypes: bool(1), float64(1), int64(2), object(6)
memory usage: 13.9+ MB
```

```
df.columns
```

```
Index(['parent_asin', 'user_id', 'helpful_vote', 'asin', 'text', 'timestamp',
       'images', 'verified_purchase', 'title', 'rating'],
      dtype='object')
```

Figure 4.3 Data Information

In figure 4.4 shows the dataset description which is about the basic statistical analysis such as mean, standard deviation, minimum and maximum values.

```
df.describe()
```

	helpful_vote	timestamp	rating
count	200000.000000	2.000000e+05	200000.000000
mean	1.108975	1.545500e+12	4.412790
std	10.618726	8.653161e+10	1.111684
min	0.000000	9.587741e+11	1.000000
25%	0.000000	1.482169e+12	4.000000
50%	0.000000	1.558833e+12	5.000000
75%	0.000000	1.614727e+12	5.000000
max	1561.000000	1.679245e+12	5.000000

Figure 4.4 Dataset Description

A word cloud including reviews with positive sentiments was shown in the figure below. Through the use of word cloud analysis, it was determined that the words "easy," "perfect," "love," "great," "nice," "beautiful," "good", "pen", "better" were the ones that were used the most often in the evaluation of the products. It was clear that the office product was helping them to meet with their needs since the term "good quality" was used. The term "works great" was a representation of the products that satisfy them. During this time, the words "price" and "happy" were used to symbolize their satisfaction of the price of the products which is affordable.

Figure 4.5 World Cloud of Positive Sentiment

A word cloud including reviews with negative sentiments was shown in the figure below. Through the use of word cloud analysis, it was determined that the words "small," "little," "bad," "waste," and "disappointed" were the ones that were used the most often in the reviews of the products. The words "time," "hard," and "small " demonstrate that a significant number of individuals will have a negative attitude towards a number of the office supplies that have been purchased.

Figure 4.6 World Cloud of Negative Sentiment

Figure below illustrate the word cloud for neutral sentiment reviews. Word cloud analysis illustrated that "use," "color," "wish," "product," were the most frequent used words in the review. These words can represent that the products don’t meet their meet as their expectation.

Figure 4.7 World Cloud of Neutral Sentiment

A representation of the rating distribution for the office product through amazon seen in Figure 4.8. The rating ranges from one to five, and as can be seen, the majority of the ratings are five, which indicates that the majority of sellers are offering items of a high quality at prices that are within reasonable ranges. Indicating most of the customers are satisfied with their purchase.

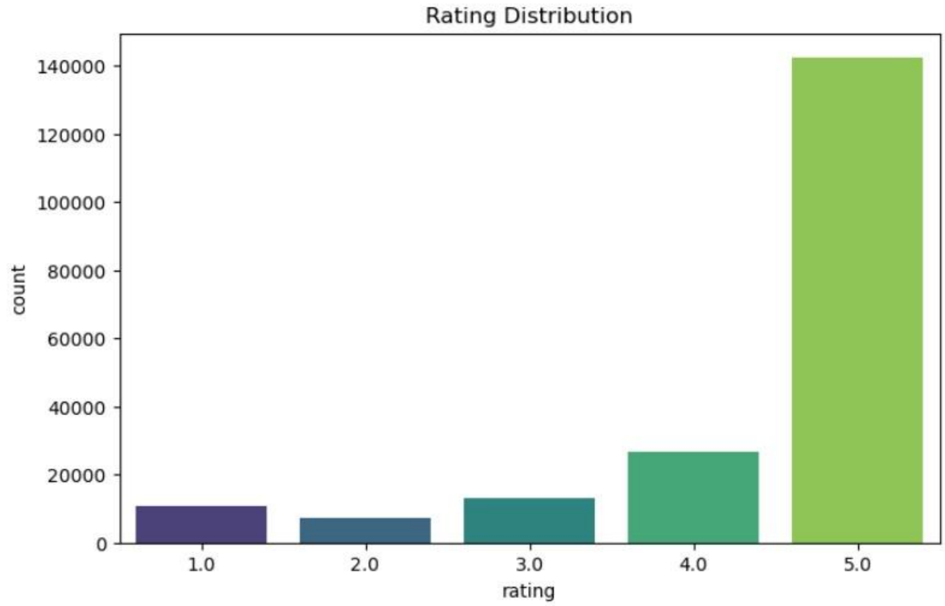

Figure 4.8 Rating Distribution

Figure 4.9 illustrate the verified purchase distribution which indicate the majority of the customers are verified their purchase. Whereas smaller portion are not verifying their purchase.

Figure 4.9 Verifies Purchase Distribution

A list of the top ten titles of reviews submitted by consumers is shown in the figure below. "Good Product" accounts for the smallest share, while "Five Star" is the term that receives the most amount of portion.

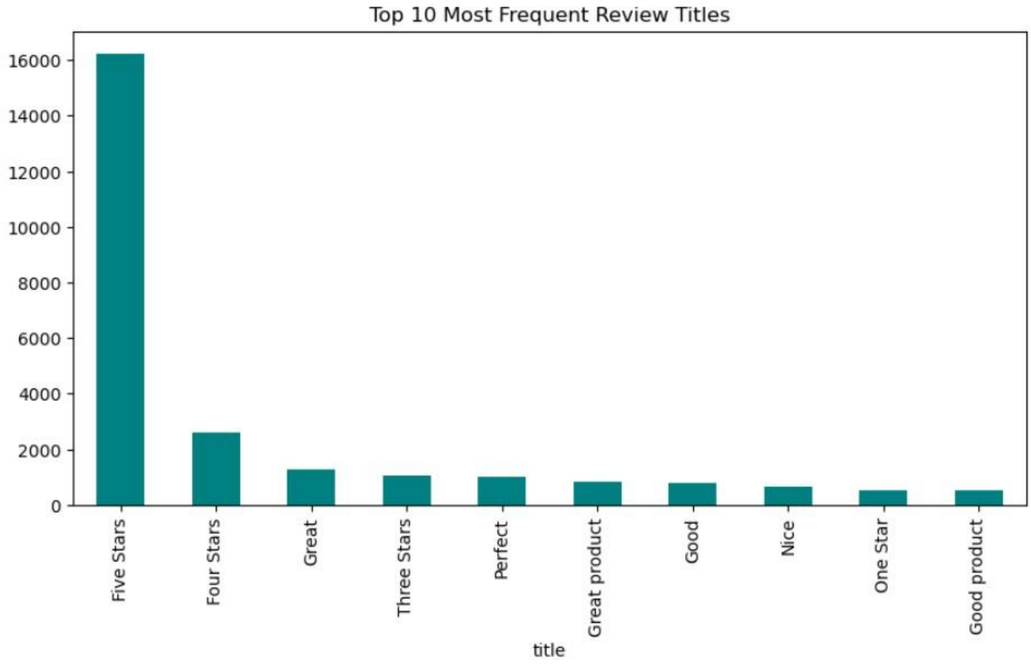

Figure 4.10 Top 10 Most Frequent Review Titles

A helpful association between rating and helpful vote was indicated by the scatter plot as illustrated below. The growing number of customers who considered the evaluations to be helpful, as well as the rising ratings of the ability of the products meet their needs. As a result, helpful vote that have received higher ratings are more likely to the customer shows that the product is helpful.

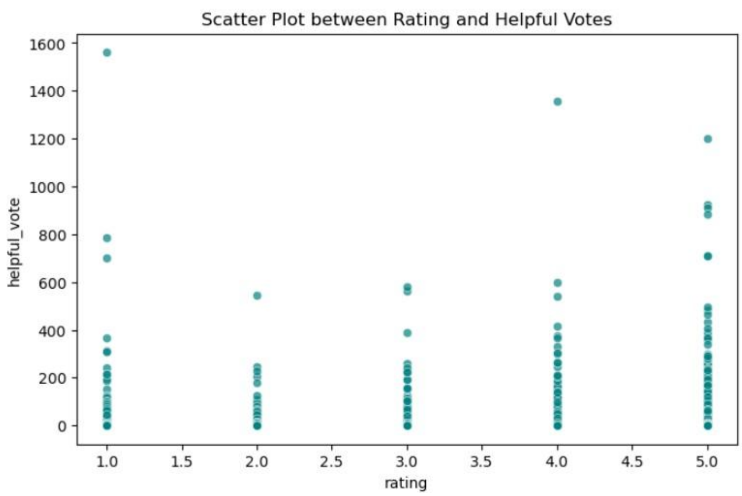

Figure 4.11 Scatter Plot between Rating and Helpful_vote

4.3 Data Cleaning

During the process of sentiment analysis, data cleaning is an essential step, particularly for the purpose of ensuring that the data utilized is clean, pertinent, and capable of being processed effectively by the model. The following are the processes that were taken to clean the data that was collected from the Amazon dataset of the office products category. Figure 4.12 illustrate the steps of the data cleaning.

Removing the unnecessarily columns help to minimize the columns in the dataset to work with in easily. Moreover, rearranging the columns in a specific way to make easy to reach. Furthermore, drop any rows that having missing values, convert to lower case which the majority of the letters in the text have been changed to lowercase. Due to the fact that both capital and lowercase letters are handled in the same manner, the analysis becomes more consistent. For instance, the terms "LIKE" and "like" are often used interchangeably. lastly, remove digits, punctuate, and extra space to make the text more effective and neater to do the sentiment analysis.

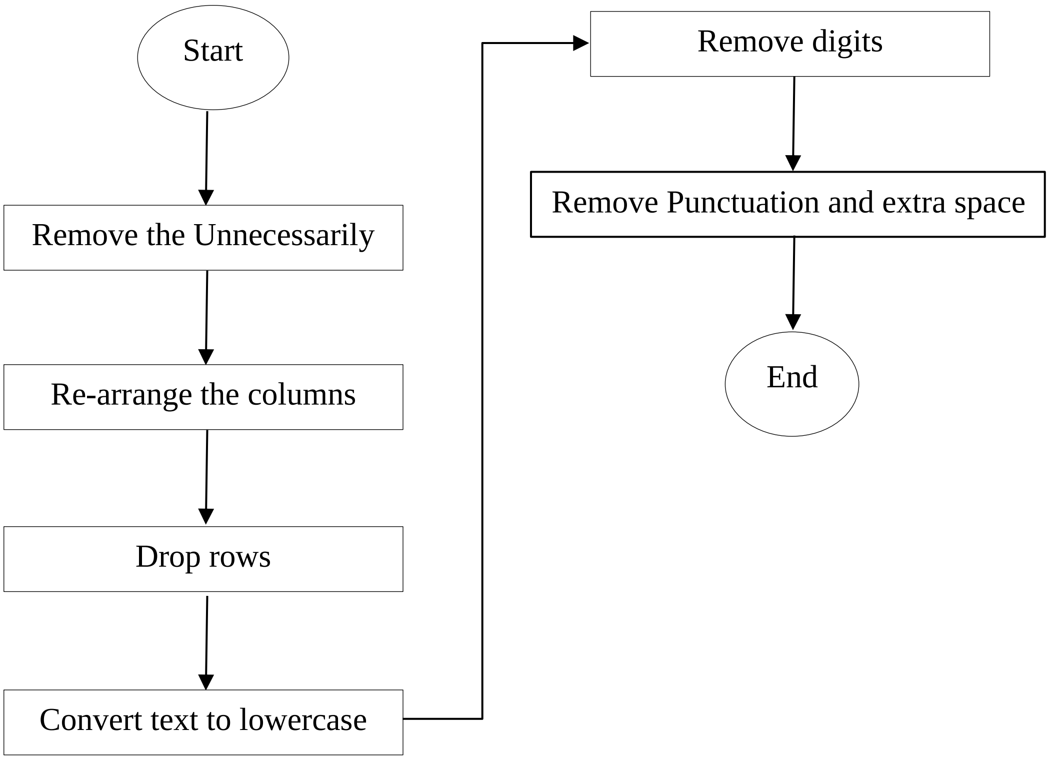

Figure 4.12 flowchart of data cleaning

4.4 Data Preparation

In order to guarantee that the data is clean and formatted before any further processing takes place, the preparation of the data is an essential step. Therefore, the reviews by the customer in the dataset should apply to them snowball stemmed to return the word to it’s original like ‘lovely’ become ‘love’ to make the process faster later on when applying the sentiment analysis

Therefore, tokenize the text and slice it for faster processing, this technique is known as tokenisation, and it involves separating a text into its component words, also known as tokens as illustrated in the figure below.

```
example = df_cleaned ['Text'] [49]
print (example)
```

```
This helps you find the remotes.
```

```
tokens= nltk.word_tokenize (example)
tokens[:10]
```

```
['This', 'helps', 'you', 'find', 'the', 'remotes', '.']
```

Figure 4.13 Example of Tekenisation

The next phase is Part-of-speech (POS) tagging, which is used to give a part-of-speech tag to each tokenized word in the review. POS tagging plays a significant role in numerous NLP tasks by offering linguistic insights and enabling the analysis and interpretation of textual data.

```
tagged= nltk.pos_tag(tokens)
tagged[:10]
```

```
[('This', 'DT'),
 ('helps', 'VBZ'),
 ('you', 'PRP'),
 ('find', 'VBP'),
 ('the', 'DT'),
 ('remotes', 'NNS'),
```

Figure 4.14 POS Tagging

4.5 Sentiment Analysis

Within the context of this section on sentiment analysis, we will locate each word that appears in reviews from Amazon dataset and classify it according to whether it is positive, negative, or neutral according to Valence Aware Dictionary and sEntiment Reasoner and Reberta analysis. Take a look at the following examples below of sentences that are positive, negative, and neutral.

Table 4.1 Some Examples of Sentiment Analysis sentences

Review	**Sentiment**
They work great. I love the colours	Positive
These were okay some worked and some didn't No worries don't have this printer anymore	Neutral
Not impressed. Print is very poor quality. Looks faded.	Negative

4.6 Model Development

In this research using two models in the sentiment analysis which are Valence Aware Dictionary and sEntiment Reasoner and Robustly Optimized BERT Pretraining Approach using Python. Some libraries were used including the NLTK, vaderSentiment, Transformers, Pandas, Scikit-learn, and seaborn. Those are helping to approach the sentiment analysis in effective way.

For the rule-based sentiment analysis, the VADER (Valence Aware Dictionary and sEntiment Reasoner) model was implemented using the NLTK library. VADER operates by assigning sentiment scores to text based on a pre-defined lexicon and set of rules. It produces four scores: positive, negative, neutral, and a compound score that combines these sentiments into a single metric .Sentiment classification was performed by applying thresholds to the compound score, categorizing reviews as positive, neutral, or negative. Due to its simplicity and efficiency, VADER was particularly useful for quickly analysing large volumes of reviews.

The phrases "neg," "neu," and "pos" are used in VADER data frame as shown below is referring to various characteristics of the sentiment that is communicated in a piece of text. These elements are influenced by the degree to which the text contains positive, negative, and neutral feelings, and the strength of those sentiments.

	Id	neg	neu	pos	compound	Product_ID	Helpful_Vote	Rating	Time	verified_purchase	Summary	Text
0	1	0.000	0.677	0.323	0.9300	B01MZ3SD2X	0	5.0	1677939345945	True	Pretty & I love it!	Lovely ink. Writes well. The right amount of w...
1	2	0.051	0.771	0.178	0.9481	B08L6H23JZ	0	4.0	1677939160682	True	2 excellent 1 extremely dry (blue)	Overall I'm pretty happy with this purchase bc...
2	3	0.070	0.815	0.115	0.9498	B07JDZ5J46	2	1.0	1660188831933	True	I don't get the reviews. Mine are garbage.	[[VIDEOID:63276c19932aa4f3687042b8b9f8613c]] U...
3	4	0.072	0.755	0.173	0.9941	B07BR2PBJN	0	4.0	1659806066713	True	Ordering Ink online: never a good idea I guess.	It's a beautiful color, but even though it had...
4	5	0.142	0.776	0.082	-0.9306	B097SFY5ZS	0	3.0	1659799390978	True	Mine are iffy at best.	Idk if I just got a bad batch which is possibl...

Figure 4.15 Data Frame of Vader Model

The neg score, also known as the negative score, is a statistic that represents the percentage of negative emotion that is present in the text. The amount to which the language displays negative emotions, such as despair, or frustration, is represented by this metric. The value neg can range from 0 to 1, with higher values signifying a stronger degree of negativity than lower values on the scale.

The pos score, also known as the positive score, is measurement of the percentage of positive emotion that is present in the text. The amount to which the text communicates pleasant feelings, such as happiness, joy or contentment, is represented by this metric. Additionally, the value pos can vary from 0 to 1, with larger values suggesting a more robust positive.

The neu score, also known as the neutral score, is a type of score that reflects the percentage of the text that has neutral sentiment. The extent to which the text is free of strong emotional polarity or lacks a strong emotional polarity is represented by this. A higher neu score indicates that the text comprises a greater amount of neutral language and a lower amount of information that is emotive. As is the case with neg and pos, the value neu might be anywhere between 0 and 1.

In the context of a piece of writing, the compound score is a single value that indicates the overall sentiment polarity of the text. In order to offer a thorough evaluation of the text's sentiment, it takes into account not only the positive and negative sentiment ratings, but also the intensities of those values. Between -1 and 1, it is a range.

This suggests that the majority of the reviews are positive and have a strong sentiment polarity in the positive direction, whilst the number of negative reviews is quite minimal and has a low sentiment polarity.

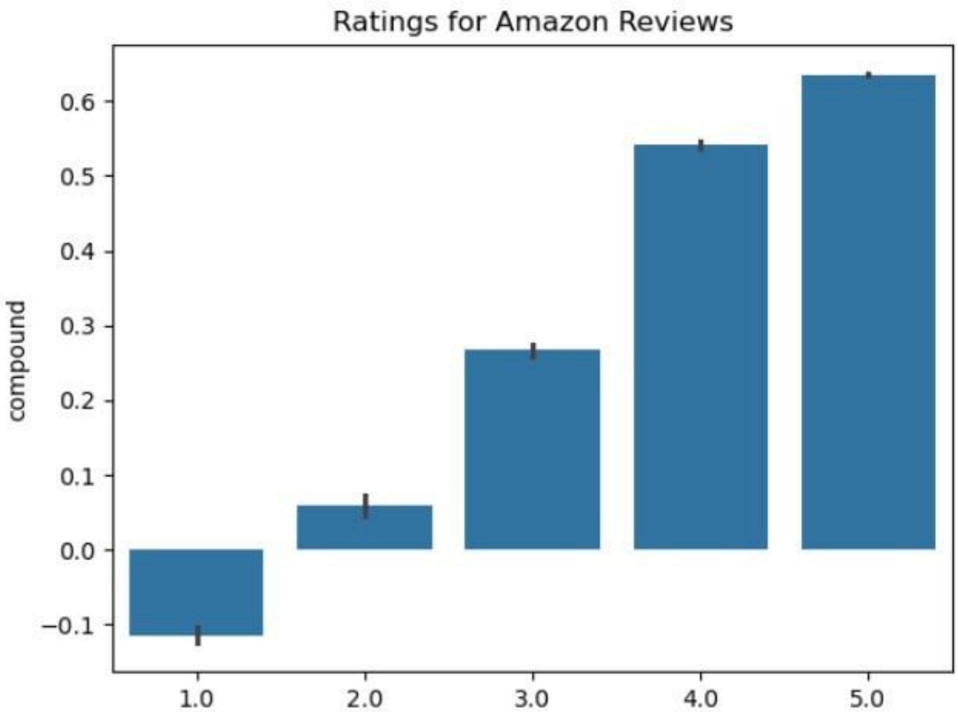

Figure 4.16 Visualization of VADER Sentiment

In contrast, the RoBERTa (Robustly Optimized BERT Pretraining Approach) model provided a more sophisticated approach to sentiment analysis. RoBERTa is a transformer-based deep learning model capable of understanding contextual

The phrases "neg," "neu," and "pos" are used in RoBERTa data frame as shown below is referring to various characteristics of the sentiment that is communicated in a piece of text. These elements are influenced by the degree to which the text contains positive, negative, and neutral feelings, respectively, and the strength of those sentiments.

roberta_neg	roberta_neu	roberta_pos	Product_ID	Helpful_Vote	Rating	Time	verified_purchase	Summary	Text
0.001184	0.016531	0.982284	B01MZ3SD2X	0	5.0	1677939345945	True	Pretty & I love it!	Lovely ink. Writes well. The right amount wet/...
0.066594	0.202709	0.730696	B08L6H23JZ	0	4.0	1677939160682	True	2 excellent 1 extremely dry (blue)	Overall I'm pretty happy purchase bc ink good ...
0.907260	0.081818	0.010923	B07JDZ5J46	2	1.0	1660188831933	True	I don't get reviews. Mine garbage.	[[VIDEOID:63276c19932aa4f3687042b8b9f8613c]] U...
0.156052	0.423542	0.420406	B07BR2PBJN	0	4.0	1659806066713	True	Ordering Ink online: never good idea I guess.	It's beautiful color, even though packed extre...
0.745801	0.219413	0.034786	B097SFY5ZS	0	3.0	1659799390978	True	Mine iffy best.	Idk I got bad batch possible I suppose bc let'...

Figure 4.17 Data frame of Roberta Model

4.7 Summary

In conclusion there were two models were developed in this work which are VADER and Roberta to do the sentiment analysis of office product reviews in Amazon. The VADER sentiment is provided to give some insight of the behavior of the customer as positive, negative, and neutral. The most portion of VADER analysis is positive which reflect that the customer are satisfied. On the other hands, Roberta analysis was developed by the data frame

CHAPTER 5

DISCUSSION AND FUTURE WORK

5.1 Introduction

This project aims to identify pattern of the customer in office product through amazon to enhance the understanding of the behaviour of the customer. To ensure the result obtained from the analysis is accurate and reliable, a few steps have been taken to improve the data quality. Preprocessing steps such as EDA and data preprocessing were carried out to retrieve the input data for further analysis. This preprocessing ensures that the dataset is consistent and complete. Then, preparing the models which are VADER and Roberta to do the sentiment analysis of the reviews to group it into positive, neutral, and negative.

5.2 Future Work

In this project, the research framework only achieved the halfway. Research planning and initial study that conducted during the first step provide with the comprehensive understanding of the sentiment analysis with different types of machine learning. Meanwhile, data preparation that is carried out during second step allows a cleaned dataset to be collected by cleaning process and normalizing text data. Among the total of 200,000 office products reviews available in the dataset were successfully processed in third step which is to get the analysis into positive, neutral, and negative, to derive the behaviour of the customers.

Future works in this project are:

(a) Get the sentiment analysis of both VADER and Roberta.
(b) Determine the accuracy of both models to get which one is more accurate of the sentiment analysis
(c) Visualization of the insights from office product reviews through dashboard

REFERENCES

Agarwal et al., 2016. Agarwal, B., Mittal, N., Agarwal, B., and Mittal, N. (2016). Semantic orientation-based approach for sentiment analysis. *Prominent feature extraction for sentiment analysis*, pages 77–88.

Al Amrani et al., 2018. Al Amrani, Y., Lazaar, M., and El Kadiri, K. E. (2018). Random forest and support vector machine based hybrid approach to sentiment analysis. *Procedia Computer Science*, 127:511–520.

Al Moubayed et al., 2020. Al Moubayed, N., McGough, S., and Hasan, B. A. S. (2020). Beyond the topics: how deep learning can improve the discriminability of probabilistic topic modelling. *PeerJ Computer Science*, 6:e252.

Alqaryouti et al., 2020. Alqaryouti, O., Siyam, N., Monem, A. A., and Shaalan, K. (2020). Aspect-based sentiment analysis using smart government review data. *Applied Computing and Informatics*, 20(1/2):142–161.

Annett and Kondrak, 2008. Annett, M. and Kondrak, G. (2008). A comparison of sentiment analysis techniques: Polarizing movie blogs. In *Advances in Artificial Intelligence: 21st Conference of the Canadian Society for Computational Studies of Intelligence, Canadian AI 2008 Windsor, Canada, May 28-30, 2008 Proceedings 21*, pages 25–35. Springer.

Araque et al., 2017. Araque, O., Corcuera-Platas, I., Sánchez-Rada, J. F., and Iglesias, C. A. (2017). Enhancing deep learning sentiment analysis with ensemble techniques in social applications. *Expert Systems with Applications*, 77:236–246.

Baccianella et al., 2010. Baccianella, S., Esuli, A., Sebastiani, F., et al. (2010). Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining. In *Lrec*, volume 10, pages 2200–2204. Valletta.

Bastı et al., 2015. Bastı, E., Kuzey, C., and Delen, D. (2015). Analyzing initial public offerings' short-term performance using decision trees and svms. *Decision Support Systems*, 73:15–27.

Batrinca and Treleaven, 2015. Batrinca, B. and Treleaven, P. C. (2015). Social media analytics: a survey of techniques, tools and platforms. *Ai & Society*, 30:89–116.

Behdenna et al., 2018. Behdenna, S., Barigou, F., and Belalem, G. (2018). Document level sentiment analysis: a survey. *EAI endorsed transactions on context-aware systems and applications*, 4(13):e2–e2.

Bhatia et al., 2015. Bhatia, P., Ji, Y., and Eisenstein, J. (2015). Better document-level sentiment analysis from rst discourse parsing. *arXiv preprint arXiv:1509.01599*.

Bird et al., 2019. Bird, J. J., Ekárt, A., Buckingham, C. D., and Faria, D. R. (2019). High resolution sentiment analysis by ensemble classification. In *Intelligent Computing: Proceedings of the 2019 Computing Conference, Volume 1*, pages 593–606. Springer.

Birjali et al., 2021. Birjali, M., Kasri, M., and Beni-Hssane, A. (2021). A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226:107134.

Biyani et al., 2013. Biyani, P., Caragea, C., Mitra, P., Zhou, C., Yen, J., Greer, G. E., and Portier, K. (2013). Co-training over domain-independent and domain-dependent features for sentiment analysis of an online cancer support community. In *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pages 413–417.

Blum and Mitchell, 1998. Blum, A. and Mitchell, T. (1998). Combining labeled and unlabeled data with co-training. In *Proceedings of the eleventh annual conference on Computational learning theory*, pages 92–100.

Broekens et al., 2015. Broekens, J., Jacobs, E., and Jonker, C. M. (2015). A reinforcement learning model of joy, distress, hope and fear. *Connection Science*, 27(3):215–233.

Cambria et al., 2017. Cambria, E., Das, D., Bandyopadhyay, S., Feraco, A., et al. (2017). *A practical guide to sentiment analysis*, volume 5. Springer.

Cardoso et al., 2018. Cardoso, E. F., Silva, R. M., and Almeida, T. A. (2018). Towards automatic filtering of fake reviews. *Neurocomputing*, 309:106–116.

Chen et al., 2019. Chen, C., Zhuo, R., and Ren, J. (2019). Gated recurrent neural network with sentimental relations for sentiment classification. *Information Sciences*, 502:268–278.

Chen et al., 2017. Chen, T., Xu, R., He, Y., and Wang, X. (2017). Improving sentiment analysis via sentence type classification using bilstm-crf and cnn. *Expert Systems with Applications*, 72:221–230.

Collomb et al., 2014. Collomb, A., Costea, C., Joyeux, D., Hasan, O., and Brunie, L. (2014). A study and comparison of sentiment analysis methods for reputation evaluation. *Rapport de recherche RR-LIRIS-2014-002*.

Cui et al., 2005. Cui, X., Potok, T. E., and Palathingal, P. (2005). Document clustering using particle swarm optimization. In *Proceedings 2005 IEEE Swarm Intelligence Symposium, 2005. SIS 2005.*, pages 185–191. IEEE.

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

De Saa and Ranathunga, 2020. De Saa, E. and Ranathunga, L. (2020). Self-reflective and introspective feature model for hate content detection in sinhala youtube videos. In *2020 From Innovation to Impact (FITI)*, volume 1, pages 1–6. IEEE.

Devi et al., 2019. Devi, D. N., Venkata Rajini Kanth, T., Mounika, K., and Sowjanya Swathi, N. (2019). Assay: Hybrid approach for sentiment analysis. In *Information and Communication Technology for Intelligent Systems: Proceedings of ICTIS 2018, Volume 1*, pages 309–318. Springer.

Do et al., 2019. Do, H. H., Prasad, P. W., Maag, A., and Alsadoon, A. (2019). Deep learning for aspect-based sentiment analysis: a comparative review. *Expert systems with applications*, 118:272–299.

Dong and Dong, 2006. Dong, Z. and Dong, Q. (2006). *Hownet and the computation of meaning (with Cd-rom)*. World Scientific.

Elshakankery and Ahmed, 2019. Elshakankery, K. and Ahmed, M. F. (2019). Hilatsa: A hybrid incremental learning approach for arabic tweets sentiment analysis. *Egyptian Informatics Journal*, 20(3):163–171.

Fernández-Gavilanes et al., 2016. Fernández-Gavilanes, M., Álvarez-López, T., Juncal-Martínez, J., Costa-Montenegro, E., and González-Castaño, F. J. (2016). Unsupervised method for sentiment analysis in online texts. *Expert Systems with Applications*, 58:57–75.

Ferrari and Esuli, 2019. Ferrari, A. and Esuli, A. (2019). An nlp approach for cross-domain ambiguity detection in requirements engineering. *Automated Software Engineering*, 26(3):559–598.

Ficamos et al., 2017. Ficamos, P., Liu, Y., and Chen, W. (2017). A naive bayes and maximum entropy approach to sentiment analysis: Capturing domain-specific data in weibo. In *2017 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pages 336–339. IEEE.

Fisch et al., 2013. Fisch, D., Kalkowski, E., and Sick, B. (2013). Knowledge fusion for probabilistic generative classifiers with data mining applications. *IEEE Transactions on Knowledge and Data Engineering*, 26(3):652–666.

Flek, 2020. Flek, L. (2020). Returning the n to nlp: Towards contextually personalized classification models. In *Proceedings of the 58th annual meeting of the association for computational linguistics*, pages 7828–7838.

Fredriksen-Goldsen and Kim, 2017. Fredriksen-Goldsen, K. I. and Kim, H.-J. (2017). The science of conducting research with lgbt older adults-an introduction to aging with pride: National health, aging, and sexuality/gender study (nhas).

Gao et al., 2014. Gao, W., Li, S., Xue, Y., Wang, M., and Zhou, G. (2014). Semi-supervised sentiment classification with self-training on feature subspaces. In *Chinese Lexical Semantics: 15th Workshop, CLSW 2014, Macao, China, June 9–12, 2014, Revised Selected Papers 15*, pages 231–239. Springer.

Gope et al., 2022. Gope, J. C., Tabassum, T., Mabrur, M. M., Yu, K., and Arifuzzaman, M. (2022). Sentiment analysis of amazon product reviews using machine learning and deep learning models. In *2022 International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE)*, pages 1–6. IEEE.

Gupta and Joshi, 2019. Gupta, I. and Joshi, N. (2019). Enhanced twitter sentiment analysis using hybrid approach and by accounting local contextual semantic. *Journal of intelligent systems*, 29(1):1611–1625.

Gutiérrez et al., 2019. Gutiérrez, L., Bekios-Calfa, J., and Keith, B. (2019). A review on bayesian networks for sentiment analysis. In *Trends and Applications in Software Engineering: Proceedings of the 7th International Conference on Software Process Improvement (CIMPS 2018) 7*, pages 111–120. Springer.

Hajmohammadi et al., 2015a. Hajmohammadi, M. S., Ibrahim, R., and Selamat, A. (2015a). Graph-based semi-supervised learning for cross-lingual sentiment classification. In *Intelligent Information and Database Systems: 7th Asian Conference, ACIIDS 2015, Bali, Indonesia, March 23-25, 2015, Proceedings, Part I 7*, pages 97–106. Springer.

Hajmohammadi et al., 2015b. Hajmohammadi, M. S., Ibrahim, R., Selamat, A., and Fujita, H. (2015b). Combination of active learning and self-training for cross-lingual sentiment classification with density analysis of unlabelled samples. *Information sciences*, 317:67–77.

Han et al., 2020. Han, Y., Liu, Y., and Jin, Z. (2020). Sentiment analysis via semi-supervised learning: a model based on dynamic threshold and multi-classifiers. *Neural Computing and Applications*, 32:5117–5129.

Hasan et al., 2015. Hasan, K. A., Sabuj, M. S., and Afrin, Z. (2015). Opinion mining using naive bayes. In *2015 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, pages 511–514. IEEE.

Hatzivassiloglou, 1997. Hatzivassiloglou, V. (1997). Predicting the semantic orientation of adjectives. In *Proceedings of the 8th conference on European chapter of the Association for Computational Linguistics*.

He and Zhou, 2011. He, Y. and Zhou, D. (2011). Self-training from labeled features for sentiment analysis. *Information Processing & Management*, 47(4):606–616.

Hong et al., 2013. Hong, Y., Kwak, H., Baek, Y., and Moon, S. (2013). Tower of babel: A crowdsourcing game building sentiment lexicons for resource-scarce languages. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 549–556.

Hu and Liu, 2004. Hu, M. and Liu, B. (2004). Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177.

Hussain and Cambria, 2018. Hussain, A. and Cambria, E. (2018). Semi-supervised learning for big social data analysis. *Neurocomputing*, 275:1662–1673.

Indurkhya and Damerau, 2010. Indurkhya, N. and Damerau, F. J. (2010). *Handbook of natural language processing*. Chapman and Hall/CRC.

Jain et al., 2020. Jain, D., Kumar, A., and Garg, G. (2020). Sarcasm detection in mash-up language using soft-attention based bi-directional lstm and feature-rich cnn. *Applied Soft Computing*, 91:106198.

Jain et al., 2019. Jain, P. K., Pamula, R., Ansari, S., Sharma, D., and Maddala, L. (2019). Airline recommendation prediction using customer generated feedback data. In *2019*

4th International Conference on Information Systems and Computer Networks (ISCON), pages 376–379. IEEE.

Jalilvand and Salim, 2012. Jalilvand, A. and Salim, N. (2012). Sentiment classification using graph based word sense disambigution. In *Advanced Machine Learning Technologies and Applications: First International Conference, AMLTA 2012, Cairo, Egypt, December 8-10, 2012. Proceedings 1*, pages 351–358. Springer.

Joachims, 1998. Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features. In *European conference on machine learning*, pages 137–142. Springer.

Joulin et al., 2016. Joulin, A., Grave, E., Bojanowski, P., and Mikolov, T. (2016). Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.

Jurek et al., 2015. Jurek, A., Mulvenna, M. D., and Bi, Y. (2015). Improved lexicon-based sentiment analysis for social media analytics. *Security Informatics*, 4:1–13.

Khalid et al., 2020. Khalid, M., Ashraf, I., Mehmood, A., Ullah, S., Ahmad, M., and Choi, G. S. (2020). Gbsvm: sentiment classification from unstructured reviews using ensemble classifier. *Applied Sciences*, 10(8):2788.

Korkontzelos et al., 2016. Korkontzelos, I., Nikfarjam, A., Shardlow, M., Sarker, A., Ananiadou, S., and Gonzalez, G. H. (2016). Analysis of the effect of sentiment analysis on extracting adverse drug reactions from tweets and forum posts. *Journal of biomedical informatics*, 62:148–158.

Kumar and Garg, 2020. Kumar, A. and Garg, G. (2020). Systematic literature review on context-based sentiment analysis in social multimedia. *Multimedia tools and Applications*, 79(21):15349–15380.

Lazarova and Koychev, 2015. Lazarova, G. and Koychev, I. (2015). Semi-supervised multi-view sentiment analysis. In *Computational Collective Intelligence: 7th International Conference, ICCCI 2015, Madrid, Spain, September 21-23, 2015, Proceedings, Part I*, pages 181–190. Springer.

Lazib et al., 2020. Lazib, L., Qin, B., Zhao, Y., Zhang, W., and Liu, T. (2020). A syntactic path-based hybrid neural network for negation scope detection. *Frontiers of computer science*, 14:84–94.

Li et al., 2014. Li, C., Xu, B., Wu, G., He, S., Tian, G., and Hao, H. (2014). Recursive deep learning for sentiment analysis over social data. In *2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT)*, volume 2, pages 180–185. IEEE.

Li and Liu, 2014. Li, G. and Liu, F. (2014). Sentiment analysis based on clustering: a framework in improving accuracy and recognizing neutral opinions. *Applied intelligence*, 40:441–452.

Li et al., 2020a. Li, W., Qi, F., Tang, M., and Yu, Z. (2020a). Bidirectional lstm with self-attention mechanism and multi-channel features for sentiment classification. *Neurocomputing*, 387:63–77.

Li et al., 2020b. Li, Y., Fang, Y., and Akhtar, Z. (2020b). Accelerating deep reinforcement learning model for game strategy. *Neurocomputing*, 408:157–168.

Liu, 2022. Liu, B. (2022). *Sentiment analysis and opinion mining*. Springer Nature.

Liu and Liu, 2011. Liu, B. and Liu, B. (2011). Opinion mining and sentiment analysis. *Web data mining: exploring hyperlinks, contents, and usage data*, pages 459–526.

Liu et al., 2018. Liu, W., Zhang, L., Tao, D., and Cheng, J. (2018). Reinforcement online learning for emotion prediction by using physiological signals. *Pattern Recognition Letters*, 107:123–130.

Ma et al., 2017. Ma, B., Yuan, H., and Wu, Y. (2017). Exploring performance of clustering methods on document sentiment analysis. *Journal of Information Science*, 43(1):54–74.

Maas et al., 2011. Maas, A., Daly, R. E., Pham, P. T., Huang, D., Ng, A. Y., and Potts, C. (2011). Learning word vectors for sentiment analysis. In *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies*, pages 142–150.

Medhat et al., 2014. Medhat, W., Hassan, A., and Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, 5(4):1093–1113.

Mesnil et al., 2014. Mesnil, G., Mikolov, T., Ranzato, M., and Bengio, Y. (2014). Ensemble of generative and discriminative techniques for sentiment analysis of movie reviews. *arXiv preprint arXiv:1412.5335*.

Miller et al., 1990. Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., and Miller, K. J. (1990). Introduction to wordnet: An on-line lexical database. *International journal of lexicography*, 3(4):235–244.

Mohammad et al., 2009. Mohammad, S., Dunne, C., and Dorr, B. (2009). Generating high-coverage semantic orientation lexicons from overtly marked words and a thesaurus. In *Proceedings of the 2009 conference on empirical methods in natural language processing*, pages 599–608.

Mohammad and Turney, 2013. Mohammad, S. M. and Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational intelligence*, 29(3):436–465.

Moraes et al., 2013. Moraes, R., Valiati, J. F., and Neto, W. P. G. (2013). Document-level sentiment classification: An empirical comparison between svm and ann. *Expert Systems with Applications*, 40(2):621–633.

Ngoc et al., 2019. Ngoc, P. V., Ngoc, C. V. T., Ngoc, T. V. T., and Duy, D. N. (2019). A c4. 5 algorithm for english emotional classification. *Evolving Systems*, 10(3):425–451.

Nisbet et al., 2009. Nisbet, R., Elder, J., and Miner, G. D. (2009). *Handbook of statistical analysis and data mining applications*. Academic press.

Pang et al., 2002. Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up? sentiment classification using machine learning techniques. *arXiv preprint cs/0205070*.

Pathak et al., 2020. Pathak, A. R., Agarwal, B., Pandey, M., and Rautaray, S. (2020). Application of deep learning approaches for sentiment analysis. *Deep learning-based approaches for sentiment analysis*, pages 1–31.

Pinto et al., 2016. Pinto, A., Gonçalo Oliveira, H., and Oliveira Alves, A. (2016). Comparing the performance of different nlp toolkits in formal and social media text. In *5th Symposium on Languages, Applications and Technologies (SLATE'16)(2016)*. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

Poria et al., 2017. Poria, S., Cambria, E., Bajpai, R., and Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. *Information fusion*, 37:98–125.

Rajat et al., 2021. Rajat, R., Jaroli, P., Kumar, N., and Kaushal, R. K. (2021). A sentiment analysis of amazon review data using machine learning model. In *2021 6th International*

Conference on Innovative Technology in Intelligent System and Industrial Applications (CITISIA), pages 1–6. IEEE.

Rana and Singh, 2016. Rana, S. and Singh, A. (2016). Comparative analysis of sentiment orientation using svm and naive bayes techniques. In *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)*, pages 106–111. IEEE.

Rathor et al., 2018. Rathor, A. S., Agarwal, A., and Dimri, P. (2018). Comparative study of machine learning approaches for amazon reviews. *Procedia computer science*, 132:1552–1561.

Rehman et al., 2019. Rehman, A. U., Malik, A. K., Raza, B., and Ali, W. (2019). A hybrid cnn-lstm model for improving accuracy of movie reviews sentiment analysis. *Multimedia Tools and Applications*, 78:26597–26613.

Ren et al., 2020. Ren, L., Xu, B., Lin, H., Liu, X., and Yang, L. (2020). Sarcasm detection with sentiment semantics enhanced multi-level memory network. *Neurocomputing*, 401:320–326.

Ren et al., 2014. Ren, Y., Kaji, N., Yoshinaga, N., and Kitsuregawa, M. (2014). Sentiment classification in under-resourced languages using graph-based semi-supervised learning methods. *IEICE TRANSACTIONS on Information and Systems*, 97(4):790–797.

Riaz et al., 2019. Riaz, S., Fatima, M., Kamran, M., and Nisar, M. W. (2019). Opinion mining on large scale data using sentiment analysis and k-means clustering. *Cluster Computing*, 22:7149–7164.

Ripa et al., 2021. Ripa, S. P., Islam, F., and Arifuzzaman, M. (2021). The emergence threat of phishing attack and the detection techniques using machine learning models. In *2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI)*, pages 1–6. IEEE.

Rojas-Barahona, 2016. Rojas-Barahona, L. M. (2016). Deep learning for sentiment analysis. *Language and Linguistics Compass*, 10(12):701–719.

Sagha et al., 2017. Sagha, H., Cummins, N., and Schuller, B. (2017). Stacked denoising autoencoders for sentiment analysis: a review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 7(5):e1212.

Saleena et al., 2018. Saleena, N. et al. (2018). An ensemble classification system for twitter sentiment analysis. *Procedia computer science*, 132:937–946.

Sanagar and Gupta, 2016. Sanagar, S. and Gupta, D. (2016). Roadmap for polarity lexicon learning and resources: A survey. *Intelligent Systems Technologies and Applications 2016*, pages 647–663.

Sanagar and Gupta, 2020a. Sanagar, S. and Gupta, D. (2020a). Automated genre-based multi-domain sentiment lexicon adaptation using unlabeled data. *Journal of Intelligent & Fuzzy Systems*, 38(5):6223–6234.

Sanagar and Gupta, 2020b. Sanagar, S. and Gupta, D. (2020b). Unsupervised genre-based multidomain sentiment lexicon learning using corpus-generated polarity seed words. *IEEE Access*, 8:118050–118071.

Sankar and Subramaniyaswamy, 2017. Sankar, H. and Subramaniyaswamy, V. (2017). Investigating sentiment analysis using machine learning approach. In *2017 International conference on intelligent sustainable systems (ICISS)*, pages 87–92. IEEE.

Saumya and Singh, 2018. Saumya, S. and Singh, J. P. (2018). Detection of spam reviews: a sentiment analysis approach. *Csi Transactions on ICT*, 6(2):137–148.

Saunders, 2021. Saunders, D. (2021). *Domain adaptation for neural machine translation*. PhD thesis.

Schmidhuber, 2015. Schmidhuber, J. (2015). Deep learning in neural networks: An overview.

Shanthi and Sangeetha, 2015. Shanthi, E. and Sangeetha, D. (2015). Analyzing data through data fusion using classification techniques. In *Computational Intelligence in Data Mining-Volume 2: Proceedings of the International Conference on CIDM, 20-21 December 2014*, pages 165–173. Springer.

Shin et al., 2016. Shin, B., Lee, T., and Choi, J. D. (2016). Lexicon integrated cnn models with attention for sentiment analysis. *arXiv preprint arXiv:1610.06272*.

Sil, . Sil, J. Intrusion detection.

Silva et al., 2016. Silva, N. F. F. D., Coletta, L. F., and Hruschka, E. R. (2016). A survey and comparative study of tweet sentiment analysis via semi-supervised learning. *ACM Computing Surveys (CSUR)*, 49(1):1–26.

Sohangir et al., 2018. Sohangir, S., Wang, D., Pomeranets, A., and Khoshgoftaar, T. M. (2018). Big data: Deep learning for financial sentiment analysis. *Journal of Big Data*, 5(1):1–25.

Su et al., 2012. Su, Y., Li, S., Ju, S., Zhou, G., and Li, X. (2012). Multi-view learning for semi-supervised sentiment classification. In *2012 International Conference on Asian Language Processing*, pages 13–16. IEEE.

Suresh and Gladston Raj, 2017. Suresh, H. and Gladston Raj, S. (2017). A fuzzy based hybrid hierarchical clustering model for twitter sentiment analysis. In *Computational Intelligence, Communications, and Business Analytics: First International Conference, CICBA 2017, Kolkata, India, March 24–25, 2017, Revised Selected Papers, Part II*, pages 384–397. Springer.

Taboada et al., 2011. Taboada, M., Brooke, J., Tofiloski, M., Voll, K., and Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2):267–307.

Thet et al., 2010. Thet, T. T., Na, J.-C., and Khoo, C. S. (2010). Aspect-based sentiment analysis of movie reviews on discussion boards. *Journal of information science*, 36(6):823–848.

Tsagkalidou et al., 2011. Tsagkalidou, K., Koutsonikola, V., Vakali, A., and Kafetsios, K. (2011). Emotional aware clustering on micro-blogging sources. In *Affective Computing and Intelligent Interaction: 4th International Conference, ACII 2011, Memphis, TN, USA, October 9–12, 2011, Proceedings, Part I 4*, pages 387–396. Springer.

Tubishat et al., 2018. Tubishat, M., Idris, N., and Abushariah, M. A. (2018). Implicit aspect extraction in sentiment analysis: Review, taxonomy, oppportunities, and open challenges. *Information Processing & Management*, 54(4):545–563.

Tung, 2009. Tung, A. K. (2009). Rule-based classification.

Vassilev, 2019. Vassilev, A. (2019). Bowtie-a deep learning feedforward neural network for sentiment analysis. In *Machine Learning, Optimization, and Data Science: 5th International Conference, LOD 2019, Siena, Italy, September 10–13, 2019, Proceedings 5*, pages 360–371. Springer.

Vinodhini and Chandrasekaran, 2016. Vinodhini, G. and Chandrasekaran, R. (2016). A comparative performance evaluation of neural network based approach for sentiment classification of online reviews. *Journal of King Saud University-Computer and Information Sciences*, 28(1):2–12.

Vyas and Uma, 2019. Vyas, V. and Uma, V. (2019). Approaches to sentiment analysis on product reviews. In *Sentiment Analysis and Knowledge Discovery in Contemporary Business*, pages 15–30. IGI global.

Wan and Gao, 2015. Wan, Y. and Gao, Q. (2015). An ensemble sentiment classification system of twitter data for airline services analysis. In *2015 IEEE international conference on data mining workshop (ICDMW)*, pages 1318–1325. IEEE.

Wankhade et al., 2022. Wankhade, M., Rao, A. C. S., and Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7):5731–5780.

Wilson, 2005. Wilson, T. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of HLT/EMNLP*.

Xia et al., 2015. Xia, R., Wang, C., Dai, X., and Li, T. (2015). Co-training for semi-supervised sentiment classification based on dual-view bags-of-words representation. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1054–1063.

Yang and Chen, 2017. Yang, P. and Chen, Y. (2017). A survey on sentiment analysis by using machine learning methods. In *2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, pages 117–121. IEEE.

Yusof et al., 2015. Yusof, N. N., Mohamed, A., and Abdul-Rahman, S. (2015). Reviewing classification approaches in sentiment analysis. In *Soft Computing in Data Science: First International Conference, SCDS 2015, Putrajaya, Malaysia, September 2-3, 2015, Proceedings 1*, pages 43–53. Springer.

Zhang et al., 2012. Zhang, W., Xu, H., and Wan, W. (2012). Weakness finder: Find product weakness from chinese reviews by using aspects based sentiment analysis. *Expert Systems with Applications*, 39(11):10283–10291.

Zhang and Wallace, 2015. Zhang, Y. and Wallace, B. (2015). A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification. *arXiv preprint arXiv:1510.03820*.

Zhao et al., 2017. Zhao, Z., Rao, G., and Feng, Z. (2017). Dfds: a domain-independent framework for document-level sentiment analysis based on rst. In *Web and Big Data:*

First International Joint Conference, APWeb-WAIM 2017, Beijing, China, July 7–9, 2017, Proceedings, Part I 1, pages 297–310. Springer.

Zhou et al., 2014. Zhou, H., Chen, L., and Huang, D. (2014). Cross-lingual sentiment classification based on denoising autoencoder. In *Natural Language Processing and Chinese Computing: Third CCF Conference, NLPCC 2014, Shenzhen, China, December 5-9, 2014. Proceedings 3*, pages 181–192. Springer.

Zhu et al., 2013. Zhu, S., Xu, B., Zheng, D., and Zhao, T. (2013). Chinese microblog sentiment analysis based on semi-supervised learning. In *Semantic web and web science*, pages 325–331. Springer.