Aspect Based Opinion Mining on Restaurant Reviews

by

Sazid Hasan Tonmoy
23241037
Faiyaj Bin Ahmed
23141084
Madhurjya Sarkar
18101574
Mehejabin Binta Bashar
18101568
Rafi Ahmed
16101065

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
School of Data and Sciences
Brac University
September 2023

Declaration

It is hereby declared that

1. The thesis submitted is our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Sazid Hasan Tonmoy

Sazid Hasan Tonmoy
23241037

Faiyaj

Faiyaj Bin Ahmed
23141084

Madhurjya Sarkar

Madhurjya Sarkar
18101574

Mehejabin

Mehejabin Binta Bashar
18101568

Rafi Ahmed
16101065

Approval

The thesis titled “Aspect Based Opinion Mining on Restaurant Reviews” submitted by

1. Sazid Hasan Tonmoy(23241037)
2. Faiyaj Bin Ahmed(23141084)
3. Madhurjya Sarkar(18101574)
4. Mehejabin Binta Bashar (18101568)
5. Rafi Ahmed(16101065)

Of Summer, 2023 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science and Engineering on October 8, 2023.

Examining Committee:

Supervisor:
(Member)

Farig Yousuf Sadeque
Associate Professor
Department of Computer Science and Engineering
Brac University

Co-Supervisor:
(Member)

Moin Mostakim
Senior Lecturer
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

____

Dr. Md. Golam Rabiul Alam
Professor
Department Computer Science and Engineering
Brac University

Head of Department:
(Chair)

____

Dr. Sadia Hamid Kazi
Chairperson and Associate Professor
Department Computer Science and Engineering
Brac University

Ethics Statement (Optional)

This is optional, if you don't have an ethics statement then omit this page

Abstract

The way businesses are operating have changed due to the explosion of the internet. Social media has an increasing number of reviews as people are keen to express their opinions based on their experiences. Online reviews have become a precious asset in various disciplines such as intelligent marketing and decision-making.The number of reviews for a well-liked product might reach thousands. This makes it challenging for a prospective buyer to go through them and make up their minds. In order to overcome this challenge, a machine-learning system is needed. Aspect based Opinion mining can be used to extract the aspects from the reviews, then we can analyze the nature of the reviews and recommend them to all the customers. We plan to classify reviews about a target entity as positive, negative and neutral so that readers of the reviews do not have to go through all the reviews but instead can focus on functional items and applicable suggestions. This thesis is specifically focused on reviews in the domain of restaurants. This study extends our knowledge of online reviews by taking into account users' wants and anticipating their future behavior. Several distinct evaluative linguistic nuances shed light on internet reviews. Using an assortment of models on generated benchmark datasets, we will also empirically show the efficacy of our strategy and show that the new techniques (or modified versions) are superior to, or at least on par with, state-of-the-art methods.

Keywords: Aspect-Based Opinion Mining(ABOM); Opinion Mining; Aspect Extraction; Opinion Extraction; Random Forest (RF); Support Vector Machine(SVM); Multinomial Naive Bayes (MNB); Deep Learning; Bidirectional Long Short-Term Memory (BiLSTM); Double Stack LSTM; Customer Feedback

Dedication (Optional)

A dedication is the expression of friendly connection or thanks by the author towards another person. It can occupy one or multiple lines depending on its importance. You can remove this page if you want.

Acknowledgement

Firstly, all praise to the Great Allah for whom our thesis have been completed without any major interruption.
Secondly, to our supervisor and co-supervisor for their kind support and advice in our work. They helped us whenever we needed help.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

In the realm of restaurant decision-making, online reviews have become an integral part of the consumer's journey. A significant number of people now make it a habit to consult restaurant reviews before visiting a new establishment. However, with the surge in the number of customer reviews comes the challenge of varying quality. Restaurants may receive hundreds to thousands of reviews in a short time, making it arduous for potential patrons to sift through them and make informed choices.

To address this issue, the emerging field of opinion Mining comes to the rescue. By scrutinizing people's sentiments, emotions, and expressions related to various aspects of a restaurant, such as food quality, hygiene, staff behavior, and cost, Opinion Mining allows us to gauge the positive and negative feedback received by the establishment. This examination can be performed with varying degrees of detail.

We calculate the total sentiment of the examination at the document level. On the other hand, Sentence-level opinion delves into individual sentences to uncover the sentiments expressed therein. Going further, we encounter the Opinion level, which examines sentiments at the word level, acknowledging that polarity can vary depending on the context.

Recognizing the various facets or properties of a restaurant's entity is an essential part of aspect-based Opinion Mining. These aspects may encompass semantic functions or specific features of the establishment. For instance, in a review that praises the restaurant's spacious size and attractive interior design but criticizes the high food prices as ”extravagant,” we observe conflicting opinions across different aspects.

Through Aspect-Based Opinion Mining, we gain valuable insights into the nuances of restaurant reviews, allowing customers to make more informed dining choices based on the aspects that matter most to them.

1.1 Research Problem

With the increase of restaurants all over the world, it is very hard to maintain communication and relationships with customers. The increasingly complex challenge associated with a growing customer base is the substantial volume of data generated in the form of natural language. Extracting all the necessary information from this data is very difficult as there exist different meanings to a sentence. To make this data useful there are major challenges to solve in aspect-based opinion mining.

Aspect extraction comes first. This task entails identifying the traits and characteristics of a target item from the review. The key objective of this work is to ascertain which particular features of a reviewed item (product) are considered important in customer evaluations. Whenever extracting an aspect it is very important to know which entity this belongs to. Aspects can be categorized into two parts, with the first one being the Explicit aspect and the second one being the Implicit aspect[17]. When the aspect is in the sentence for instance, when we say "The Beef Burger tastes awesome" here "taste" is the aspect which is used in the sentence explicitly. The sentence "The restaurant is small" talks about the size of the restaurant but it is not mentioned directly. In aspect extraction, there are some things to work with such as searching the type of the aspect for opinion mining and selecting good aspects for classification.

The Aspect of Opinion Classification is the second. It is integral to categorize the opinion words that were extracted as well as those associated with the multiple traits, into one of the three polarity scales: positive, negative, or neutral. The goal is to determine the precise numerical opinion ratings for aspects or discern whether sentiments and sentiments towards identified aspects are positive, negative, or neutral[24]. This opinion classification can be expressed through symbols like stars, thumbs-up, or thumbs-down. For instance, in the sentence "Drone's camera quality is amazing," the aspect opinion classification task would label the word "amazing" as having a positive opinion orientation toward the "camera quality" aspect. Different algorithms like n-gram, naive Bayes, vector semantics, and logistic regression are used to classify the extracted aspects.

Thirdly, Finding implicit aspects is a very challenging task as people express their opinions differently and the habits of language differ from person to person. Many reviews refer to multiple aspects of different products. For example, "The restaurant is very much smaller compared to another one", and 'The restaurant is very much smaller' tells us about the restaurant's 'size' aspect. If we identify a single implicit feature, it becomes relatively straightforward to determine the sentiment associated with that particular feature but when there is more than one implicit feature present then detecting the opinion of the feature becomes more challenging[25]. When dealing with multiple implicit features, these features may exhibit varying levels of polarity compared to the overall sentiment of the sentence. For instance, in the statement "Pictures taken can get blurred because of the lack of image stabilizer but overall a great option for a given budget," two distinct implicit aspects, namely "camera quality" and "price," are mentioned. These aspects might have different polarities than the overall sentiment expressed in the sentence. Consider this example: "The food quality of the restaurant is very good but it is not that tidy".In this review, the restaurant's food quality and hygiene are brought up. Since people have positive and negative thoughts about many things, it is wise to resolve the numerous sorts of elements and take cognizance of their polarity autonomously. Finding several facets of a feature is thus a challenging endeavor.

The fourth is the Cross-Domain adaptation. The majority of opinion mining systems heavily rely on certain domains. The same perspective phrase might convey various polarities in other contexts. Prior to mining opinions from the reviews, subject understanding is imperative. It may prove daunting for academics to generate domain-independent methodologies and algorithms[15].

Next is modeling the customer's review with a view of finding semantics aspects and opinions and projecting the overall rating review. Usually, the customer gives a review with an overall rating in the form of stars. So by utilizing the overall rating, it is possible to lead the process of the sentiment aspect of the review.

1.2 Research Objectives

This research paper's key objective is to conduct an in-depth comparison of machine learning and deep learning techniques, which include Random Forest, Support Vector Machine (SVM), Multinomial Naive Bayes (Multinomial NB), Bidirectional Long Short-Term Memory (BiLSTM), and Double Stack LSTM. This analysis is carried out for the purpose of Aspect-Based Opinion Mining on restaurant reviews. More precisely, our objective is to:

1. Investigate the effectiveness of these diverse techniques in capturing and categorizing sentiment and opinions expressed in restaurant reviews at the aspect level.

2. Assess the performance of each method with respect to accuracy, precision, recall, F1-score, and computational efficiency for aspect-based sentiment classification.

3. Explore the potential benefits of deep learning models, including BiLSTM and Double Stack LSTM, in handling the nuanced and context-dependent nature of aspect-based Opinion Mining.

4. Offer insights into the strengths, limitations, and practical suitability of each technique for aspect-based opinion mining in restaurant reviews, advancing the field through a comprehensive comparative study applicable not only in the restaurant industry but also across various domains.

Chapter 2

Literature Review

For decision making opinions can play a major role in choosing from multiple choices involving valuable resources. Until recently, friends and specialized magazines or websites were the main sources of information. Our ability to easily produce and exchange ideas with everyone linked through discussion boards, blogs, social media platforms, and content-sharing services has evolved thanks to the internet, which has given us an abundance of possibilities and new tools. According to Statista [22], in 2020, social media users reached over 3.6 billion people worldwide, a number projected to increase to almost 4.41 billion in 2025. However, a new study by Kepios claims that by April 2022, there will be approximately 4.65 billion online social networking users globally, which is 58.7% of the whole of humanity's populace[30]. Due to the pandemic, social media engagement rates have surged over the previous 12 months as well, with 326 million new individuals joining during this period last year. This translates to a yearly increase of 7.5 percent, or in excess of 10 new users per second on average[32]. As a result, the number of shared opinions on various topics over the internet is exponentially increasing. Due to the unstructured nature of this information, it's not machine-processable. The scientific community is becoming increasingly proactive in gathering public opinion on many issues because of the possible difficulties that may arise from exploring uncharted areas[31]. The mining of opinions and sentiments is one of the new areas that have emerged as a consequence of this. Aspect Based Opinion Mining(ABOM) aims to extract aspects of products and classify the corresponding polarities of the user in the review[20]. Previously, several approaches have been used to study ABOM based on text reviews. Vector extraction is now utilized in opinion extraction and sentiment evaluation to acquire the most prominent and significant linguistic properties. Frequency and presence are the two most common features in Vector classification [21]. Additionally, n-grams, which are often bigrams and trigrams, are seen as having useful qualities. While linguistic evaluation employs part of speech (also known as POS) information, such as adjectives, nouns, adverbs, and verbs, as a fundamental kind of Word-Sense Disambiguation (WSD), other approaches additionally rely on the disparity within words. The aforementioned techniques are firmly constrained by topic and domain.[3].

2.1 Related Works

This section attempts to critically evaluate earlier significant research in the area of opinion mining within the structure for aspects-based sentiment analysis [32]. We examined the various techniques and tactics used to achieve the intended outcome. Mining data, computational languages, as well as the processing of natural languages (NLP) are all included in opinion mining. Opinion mining doesn't necessitate a complete understanding of the text, unlike its predecessor, regular syntactical NLP. While Syntactical NLP places a strong emphasis on summarization and automatic categorization, it is mainly focused on semantics inference and emotional data connected to natural language [32]. At several levels, such as the documenting, term, single entity, and aspect levels, opinion mining can be applied. Whether a whole opinion paper reflects a favorable or negative mood is dependent on the document level. As opposed to sentence-level analysis, which looks at each sentence to see if it's neutral, negative, or reflects an opinion. Entity and aspect level analysis is more accurate since it focuses on the perspective itself rather than the structures of language (documents, portions, sentences, clauses, or phrases) [22]. It is predicated on the notion that a point of view is composed of a subject matter, a sentiment (whether favorable or unfavorable), and both. The three primary tasks in ABM 3 are aspect identification, ABM word recognition, and ABM word orientation detection. Consider a review of a restaurant that begins, "While the food at this place is excellent, the hospitality is not that good." In this context, the adjectives "not that good" and "excellent" are used to describe the aspects of service and food.

The scientists involved in this work [22] classified articles by overall opinion rather than themes and found that ordinary machine learning algorithms outperform baselines produced by humans. They did this using reviews of movies as their data. They employed naive Bayes models, highest-entropy categorizing, and support vector algorithms as their three machine-learning techniques. These techniques perform well for topic-based categorization but poorly for sentiment classification. Three of the subtasks include aspects words the extraction process, aspects of categorization recognition, and aspect sentiment prediction. Alghunaim et al. [22] have investigated the effectiveness of vector representations over different text data. The vector space approach used in this work performed well in comparison to the baselines, having an F1 score of 79.91. The authors of [4] asserted that they have demonstrated the effectiveness of using aspect-based sentiment analysis by employing a hierarchy-based bidirectional LSTM to characterize the links between phrases in a review. Without using features that were manually created or outside resources, the proposed hierarchical model outperforms two non-hierarchical baselines, producing outcomes comparable to those of the most advanced datasets and even surpassing the results of a few multilingual individual multi-domain datasets. In this study, [10] a method termed Sentiment utility Logistic Model (SULM) has been suggested for identifying the most beneficial components of prospective user encounters.

Our method performed incredibly well when put to the test with real reviews from three applications from the real world. Additionally, it was able to forecast the 6-level unknown ratings of the reviews, which is comparable to the most sophisticated HFT model. Furthermore, it assumed the group of elements that a user would bring up in a hypothetical future evaluation of a certain item on the point of the most advanced LRPPM. In order to provide better services for users, service providers can profit from the favorable user experience qualities provided by SULM. The research study [11] provided the first deep learning method for extracting opinion targets from texts that have strong opinions expressed in them. Seven layers make up the deep CNN architecture that has been proposed: The final result is combined from two convolution layers, a max-pooling layer after each, a fully linked layer, and an output layer with a single neuron for each word. Word embedding features are included in the input layers for each word in the phrase. A set of already created heuristic patterns of language were linked with the neural network classifier. Performance using this model was far better than using cutting-edge methods.

The absence of fine-grained labeled data makes supervised learning methods ineffective for Aspects Based Sentiment Analysis notwithstanding their effectiveness in the field. In this study, a brand-new domain adaptation paradigm dubbed Cross-domain reviews generation (CDRG) was put out as a solution to this problem. CDRG may provide fine-grained annotation-based target-domain reviews based on the source-domain labeled review. Through careful testing, CDRG has been shown to be superior to modern domain adaptation techniques [8].

Dealing with annotated datasets with a variety of demographics is crucial, as sociocultural variables and demography are key determinants of a user's attitude and preference, according to researcher Sazzad [9]. This study initially creates a dataset with local demography as its primary emphasis. They suggested a hybrid methodology and outperformed the most effective lexicon-based and machine-learning (ML) oriented classifiers without depending on any labeled data. A second analysis was conducted to determine the influence of demographics over linguistic characteristics utilizing two diverse local and global datasets. This study's findings indicated that user demographics are very important for understanding how reviews are written.

The DomBERT extension of the general-purpose language model, BERT, is described in this study [27]. DomBERT will incorporate each appropriate boundary corpora and in-domain corpora. The goal was to lower the resource requirements of traditional domain language models by combining domain-oriented language models plus general-purpose language models. The experiment's results were reassuring, further demonstrating DomBERT's value to ABSA.

One of the most difficult ABSA tasks, particularly for supervised models, has always been the lack of fine-grained labeled data. Researchers mainly utilized based features domain adaption or instance-based domain adaption to mitigate this restriction and dependence on labeled data. But each of these approaches has advantages and disadvantages. Unified Domain Adaptation (UDA), an end-to-end framework that combines feature-based approaches adaption with instance-based adaption for the objectives of cross-domain extraction of aspects and cross-domain End2End ABSA, was proposed by researchers to address this constraint. The experiment was carried out and demonstrated to be a major improvement over the current state-of-the-art approaches for both tasks, achieving exceptional results on four benchmarks [23].

Chapter 3

Methodology

To accomplish this, we leveraged a substantial dataset obtained from Kaggle, consisting of approximately 2.7 million reviews sourced from various countries. Each sentence in this dataset was meticulously annotated with labels such as positive, negative, neutral, or N/A (indicating instances where reviewers did not mention specific aspects), ensuring that even sentences with multiple sentiments were appropriately classified.

Given the inherent challenges of processing unstructured online reviews, we recognized the paramount importance of comprehensive Exploratory Data Analysis (EDA). The preprocessing steps for this research encompassed the following:

1. Eliminating null values
2. Removing duplicate entries
3. Discarding irrelevant features
4. Converting all review text to lowercase
5. Expanding contractions
6. Eliminating digits and words containing digits
7. Removing punctuation and extra spaces
8. Removing common stopwords
9. Eliminating frequently occurring words and rare words

Following the data pre-processing stage, we proceeded to the Aspect-Related Opinion Identification phase, wherein we aimed to identify aspects related to opinion words. This was achieved through topic modeling, specifically employing Latent Dirichlet Allocation (LDA) to uncover keywords within the dataset. Subsequently, we manually selected the four most pertinent and significant keywords to serve as aspects.

For the Subjectivity and Objectivity Classification stages, we leveraged a diverse range of traditional text classification methods, including Multinomial Naive Bayes (MNB), Support Vector Machines (SVM), Random Forest, and Long Short-Term Memory (LSTM) networks. These methods were employed to differentiate between subjective and objective sentences based on aspects and opinion words.

We performed a detailed performance assessment, including metrics like F1 score,

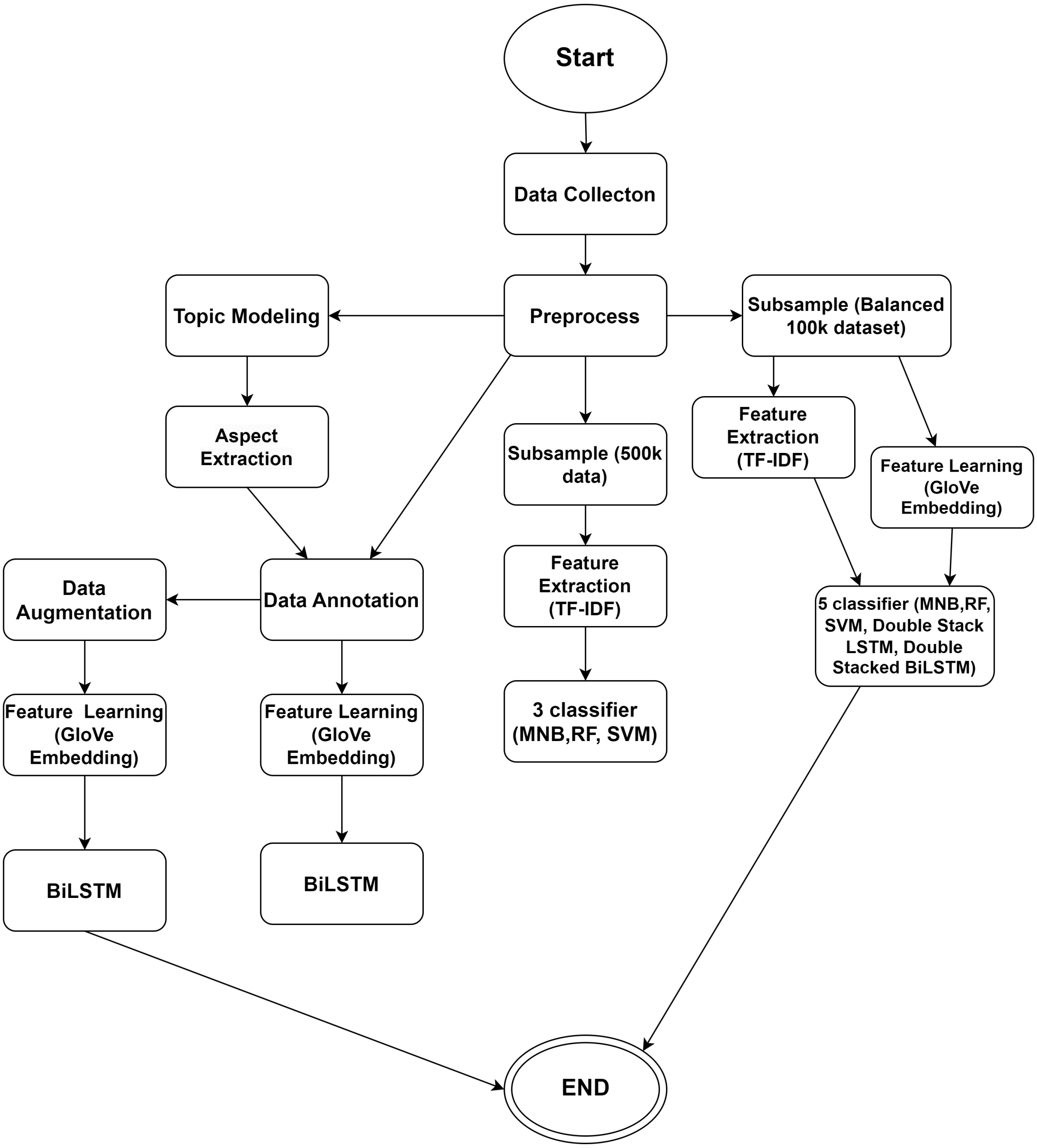

Figure 3.1: Research Methodology

	Unnamed: 0	parse_count	restaurant_name	rating_review	sample	review_id	title_review	review_preview	review_full	date	city	url_restaurant	author_id
612629	196818	196819	The_Queens_Head	4.0	Positive	review_700475093	Authentic British pub in Piccadilly	This pub has two floors, ground floor is a pub...	This pub has two floors, ground floor is a pub...	August 18, 2019	London_England	https://www.tripadvisor.com/Restaurant_Review-...	UID_145480
964835	550230	550231	ROSSODISERA_il_marchigiano	5.0	Positive	review_450783906	A little piece of Italy in the heart of Covent...	We stumbled across this gem of a restaurant wh...	We stumbled across this gem of a restaurant wh...	January 9, 2017	London_England	https://www.tripadvisor.com/Restaurant_Review-...	UID_327828
2207341	441815	441816	Red_Lobster	3	Negative	review_355823130	Good, but not great	While looking for a place to eat during out NY...	While looking for a place to eat during out NY...	March 15, 2016	New_York_City_New_York	https://www.tripadvisor.com/Restaurant_Review-...	UID_8890
238532	239160	239161	El_pebrot_i_el_petit_cargol	5	Positive	review_519641933	Awesome food and service at affordable prices!	We tried the rabbit stew with snails: it was a...	We tried the rabbit stew with snails: it was a...	August 30, 2017	Barcelona_Catalonia	https://www.tripadvisor.com/Restaurant_Review-...	UID_156430
223636	224179	224180	Art_i_Tapes	5	Positive	review_536122152	We loved it	Very nice food and staff. The Sangria was real...	Very nice food and staff. The Sangria was real...	October 26, 2017	Barcelona_Catalonia	https://www.tripadvisor.com/Restaurant_Review-...	UID_39913
...	...	...	...	...	...	...	...	...	...	...	...	...	...
2501157	220211	220212	Domaine_de_Lintillac	3	Negative	review_179115415	Good but could have been better!	Our daughter had tried this restaurant before ...	Our daughter had tried this restaurant before ...	September 29, 2013	Paris_Ile_de_France	https://www.tripadvisor.com/Restaurant_Review-...	UID_22053
690666	275077	275078	108_Brasserie	4.0	Positive	review_211093089	Really Good	An all round good experience for business lunc...	An all round good experience for business lunc...	June 19, 2014	London_England	https://www.tripadvisor.com/Restaurant_Review-...	UID_9803
2049184	281376	281377	Amy_Ruth_s_Home_Style_Southern_Cuisine	5	Positive	review_422939844	Amazing Soul Food	This is Soul Food like "Mom" makes. The food i...	This is Soul Food like "Mom" makes. The food i...	September 27, 2016	New_York_City_New_York	https://www.tripadvisor.com/Restaurant_Review-...	UID_164112
2426410	144967	144968	L_Arpege	5	Positive	review_75386861	Triumph of vegetables	Alain Passard is a king of vegetables. Perfect...	Alain Passard is a king of vegetables. Perfect...	August 15, 2010	Paris_Ile_de_France	https://www.tripadvisor.com/Restaurant_Review-...	UID_101802
315395	317146	317147	Los_Caracoles	4	Positive	review_230172205	Anniversary meal	We were recommended this hotel so booked for o...	We were recommended this hotel so booked for o...	September 21, 2014	Barcelona_Catalonia	https://www.tripadvisor.com/Restaurant_Review-...	UID_190381

100 rows × 13 columns

Figure 3.2: Dataset before preprocessing

	parse_count	restaurant_name	rating_review	review_id	title_review	review_full	date	review_full_wo_freqwords	review_full_wo_rarewords	review_full_lemmatized	review_len	word_count	polarity
612629	196819	The_Queens_Head	4.0	review_700475093	Authentic British pub in Piccadilly	This pub has two floors, ground floor is a pub...	August 18, 2019	pub two floors ground floor pub small restaura...	pub two floors ground floor pub small restaura...	pub two floor ground floor pub small restauran...	306	47	-0.167308
964835	550231	ROSSODISERA_il_marchigiano	5.0	review_450783906	A little piece of Italy in the heart of Covent...	We stumbled across this gem of a restaurant wh...	January 9, 2017	stumbled across gem restaurant whilst weekend ...	stumbled across gem restaurant whilst weekend ...	stumbled across gem restaurant whilst weekend ...	434	65	0.273333
2207341	441816	Red_Lobster	3	review_355823130	Good, but not great	While looking for a place to eat during out NY...	March 15, 2016	looking place eat ny trip found ourdself red l...	looking place eat ny trip found ourdself red l...	looking place eat ny trip found ourdself red l...	405	64	0.097222
238532	239161	El_pebrot_i_el_petit_cargol	5	review_519641933	Awesome food and service at affordable prices!	We tried the rabbit stew with snails: it was a...	August 30, 2017	tried rabbit stew snails absolutely delicious ...	tried rabbit stew snails absolutely delicious ...	tried rabbit stew snail absolutely delicious €...	122	17	0.712000
223636	224180	Art_i_Tapes	5	review_536122152	We loved it	Very nice food and staff. The Sangria was real...	October 26, 2017	nice food staff sangria really impressive enjo...	nice food staff sangria really impressive enjo...	nice food staff sangria really impressive enjo...	113	14	0.700000

Figure 3.3: Dataset after preprocessing

accuracy, precision, and recall, to evaluate the effectiveness of our proposed model. This allowed us to assess how well our system performed in comparison to other well-known, cutting-edge techniques.

In addition to our data preprocessing efforts and aspect-related opinion identification, we also delved into the algorithms utilized in our model, which encompassed Random Forest, SVM, Multinomial Naive Bayes, BiLSTM (Bidirectional LSTM), and Double Stack LSTM. These algorithms collectively contributed to our Aspect-Based Opinion Mining on restaurant reviews, showcasing the comprehensive nature of our research.

3.1 Data Collection

Finding the labeled data for the text is very hard. There are multiple ways of extracting the data and one of the ways is by web scraping. But we have used a dataset from Kaggle where there are reviews from customers who have reviewed specific restaurants [7]. In our dataset, there are around 2.7 million reviews from different countries. Finding fine-grained labeled datasets for large text corpus is extremely hard. The dataset we used was taken from Kaggle under the name of “Six TripAdvisor Datasets for NLP Tasks” related to the paper “Explain and Conquer: Personalised Text-based Reviews to Achieve Transparency”[29]. It has 2.7M rows with 13 columns. All the columns are object types. This dataset contains restaurant reviews from six cities across the world, Barcelona, London, Madrid, New Delhi, New York, and Paris. An excerpt from the dataset before the cleaning process, as illustrated in Figure - 3.2.

3.2 Data Pre-Processing

3.2.1 Dropping irrelevant columns

In our dataset, we initially had around 14 columns, but we discovered that the majority of them were not relevant to our specific work. Therefore, during the pre-processing phase, we decided to eliminate those irrelevant columns from the dataset. This not only streamlined the data for our analysis but also had the added benefit of reducing the memory footprint, making it more manageable for further processing.

3.2.2 Dropping the duplicate rows

Eliminating duplicate rows from our dataset was a crucial task for us because such redundant entries can significantly impact the accuracy of our prediction results. Therefore, we made it a priority to remove these duplicate rows to ensure a clean and reliable output for our analysis.

3.2.3 Dropping the missing or null values

In a large dataset, encountering missing or null values is a common occurrence that can introduce bias and impact the final results. Consequently, it is essential to address these issues appropriately, and one commonly employed method is to remove the rows containing missing or null values from the dataset. This ensures that the data is handled properly, leading to more accurate and unbiased analyses.

3.2.4 Lowercasing the texts

Machines can interpret single words in various forms, which can lead to different meanings. For instance, the word 'cat' written as 'Cat', 'cat', or 'CAT' may be perceived differently by a machine. To ensure proper understanding, it is crucial to convert all text to lowercase, enabling the machine to interpret words consistently and accurately.

3.2.5 Expand Contractions

Contractions, which involve shortening words or phrases by omitting letters or syllables and replacing them with apostrophes, serve as a way to make written and spoken language more concise. Expanding contractions become particularly valuable in formal writing or when you desire greater clarity in your communication. By expanding contractions, you standardize the text, converting it into its complete word form. This standardization process simplifies the consistent application of text analysis techniques.

Certain contractions, such as "it's," possess multiple meanings that hinge on the surrounding context. The act of expanding these contractions into "it is" or "it has" can facilitate a more accurate comprehension of the context by natural language processing models.

In the realm of Opinion Mining, the meanings of words play a pivotal role. The expansion of contractions guarantees that Opinion Mining tools correctly interpret the text, as the sentiment may fluctuate when contractions are employed.

3.2.6 Remove digits and words containing digits

We decided to remove digits and words containing digits from the dataset because they do not hold any sentimental value. This step was taken to enhance the overall cleanliness and readability of the dataset.

3.2.7 Remove Punctuation and extra spaces

To improve the comprehensibility of the reviews, we eliminate punctuation and extra spaces from the dataset since they do not carry any sentimental value. This process helps to make the text more understandable and facilitates better analysis.

3.2.8 Stop Words Removal

Stop word removal: We focused on enhancing the dataset's quality by eliminating stopwords, which are commonly used but lack substantial meaning. This step aimed to streamline the dataset, emphasizing valuable content and reducing memory usage during review classification. The removal of stopwords is a well-established technique in applications like search algorithms and text categorization, contributing to more effective analysis[33].

3.2.9 Removal of rare and frequent words

In line with our previous preprocessing step, we extended our efforts to eliminate both frequent and rare words from the dataset. These steps are crucial in enhancing the precision and relevance of the information we gather.

This preprocessing step involves eliminating highly frequent words from the text. Such words, often referred to as stopwords, do not significantly contribute to understanding the text and can burden the NLP model with an excessive number of features. By removing these common terms, we streamline the dataset, reducing its size and ensuring that the NLP model focuses on more meaningful content. Frequent words tend to cluster together, and their high frequency can increase data size without adding substantial value.

Conversely, rare words are removed to address a different concern. These terms occur infrequently and their associations with other words are often dominated by noise. In cases where a document contains a single instance of a rare word, it may be highly relevant. However, the surrounding words might be more influenced by chance than actual context. By removing rare words, we mitigate this noise, allowing the NLP model to concentrate on more meaningful linguistic patterns and associations within the text.

3.2.10 Lemmatization

Another technique to the Text normalization process that is utilized to get the words ready for the procedure is this one. This aids in capturing the essence of words and lowering the dimension of text data[34]. There are some important factors that affect the lemmatization process. To start with, The canonical form of a word, which embodies its essential meaning, is the base or root form acquired through lemmatization. As an illustration, the lemma of the phrases ”running,” ”runs,” as well as ”ran” is ”run.” Furthermore, lemmatization determines the suitable lemma by taking into account the grammatical perspectives of phrases, including tense, gender, number, and part of speech tags, which is referred to as morphological analysis. Lemmatization is separate from stemming from, another text normalization method, in addition. Lemmatization takes into consideration the context as well as the significance of words, as opposed to stemming, which merely strips words of their suffixes and prefixes to produce their root form. Lemmatization is a useful NLP method for standardizing words and lowering the level of dimensionality on text data overall. It is frequently used to increase the precision and effectiveness of machine learning models in a variety of text analysis applications.

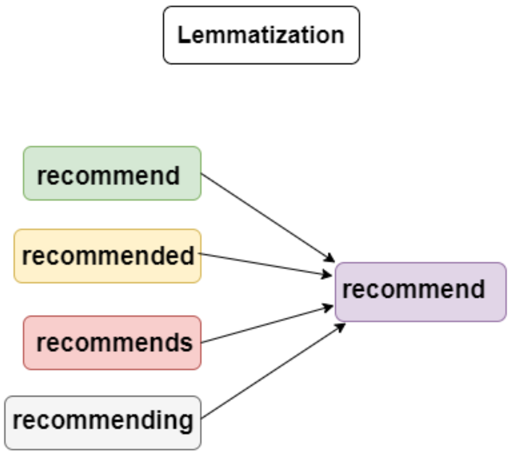

Figure 3.4: Process of Lemmatization

3.2.11 Stemming

This is one of the Text Normalization procedures used to get the text ready for more work. The major goal of this technique is to normalize texts by deleting prefixes and suffixes, resulting in the treatment of multiple word variations as one and the same word[35]. There are a few important variables that affect stemming. First of all, the basic meaning of the word stem is represented by getting by stemming. As an illustration, the root of the terms ”running,” ”runs,” as well as ”ran” is just ”run.” Second, to find and eliminate frequent suffixes and prefixes from words, stemming algorithms employ linguistic guidelines and heuristics.t. Through this process, stemming aims to cluster similar words together and reduce them to a common structure. Additionally, to complete the stemming process, stemming algorithms depend on established rules and patterns. These criteria are made to deal with word variations that are linguistically distinctive and morphological alterations. Stemming is a typical method applied in NLP to handle variances and standardize words. Despite the fact that it could fall short of more sophisticated methods like lemmatization in terms of capturing the complete semantic meaning of words, it can be a valuable preparatory step for a variety of text analysis applications.

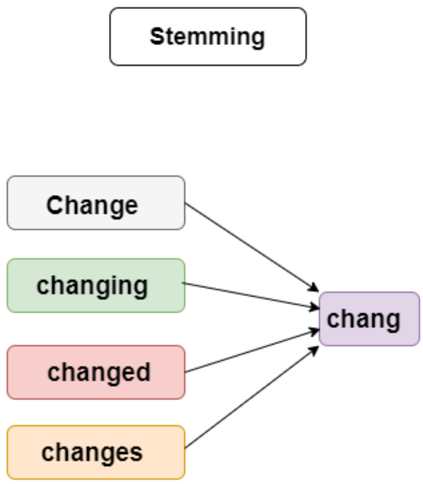

Figure 3.5: Process of Stemming

3.2.12 Conversion of Emojis and emotes to Words

Emojis and emoticons play a significant role in conveying important data within a dataset. Emojis, in particular, excel at representing emotions, making them an excellent tool for expression. They often carry more information than plain text alone. However, it's essential to be aware that emojis can also be used sarcastically. For instance, if someone sarcastically comments while giving a very low rating to a restaurant, they use the emoji to express their negative criticism in a humorous way. Without the emoji, the entire text could be interpreted differently, leading to misunderstandings. To ensure the dataset retains essential information and facilitates a better understanding of emotions, we have translated emojis into words. This approach aims to help everyone accurately comprehend sentiments and make appropriate judgments based on the data provided.

3.2.13 Removal of URLs

Because individual URL removal tends to have such a poor success rate, it is crucial. URLs, also known as uniform resource locators, are used to identify references to specific online locations in the text. These, however, don't provide any ancillary data. Therefore, leaving out URLs plays a big part.

3.2.14 Chat word Conversion

To ensure better comprehension for the machine, we transformed chat words, which are commonly used to represent words in a shortened form, into their full form. This conversion was necessary to avoid missing any crucial information and enable the machine to understand the text more effectively.

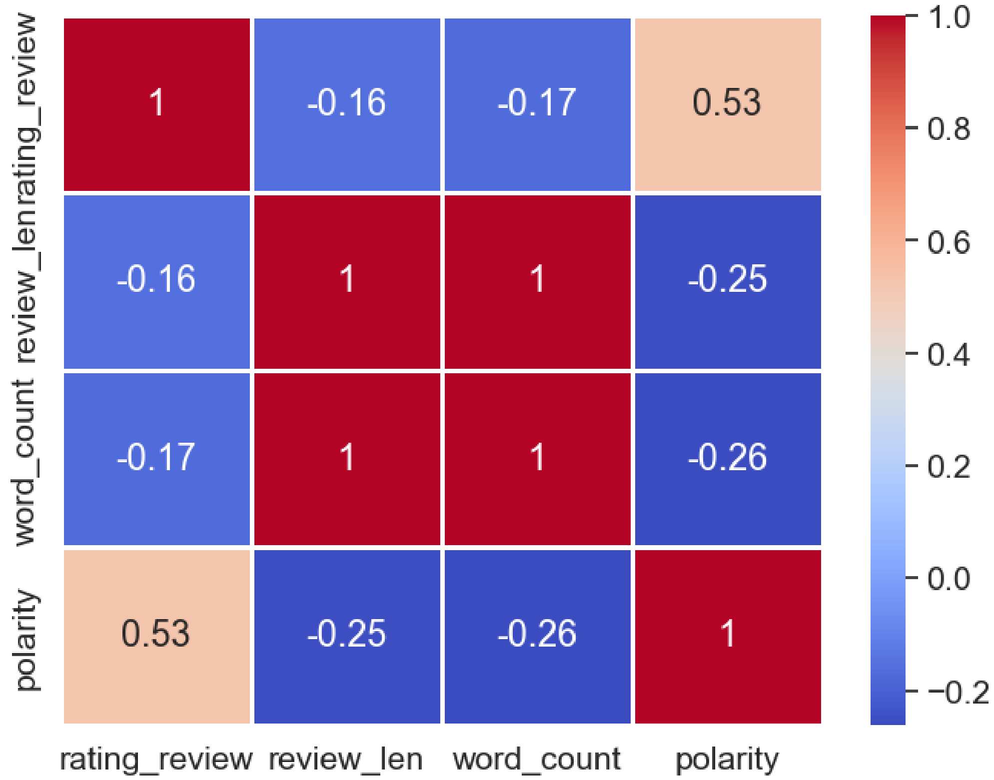

Figure 3.6: rating_review, review_len, word_count, polarity

3.2.15 Spelling Correction

In order to enhance the precision of prediction outcomes, we took the necessary step of rectifying all spelling mistakes in the dataset. By doing so, we ensured that the machine could correctly grasp the intended meaning of words, enhancing the overall performance of the model.

3.2.16 Part of Speech Tagging

Part-of-speech tagging, a widely used technique in Natural Language Processing, involves classifying words in a text corpus based on their specific grammatical roles. This includes categorizing words as nouns, verbs, adjectives, and various other grammatical categories. POS tagging plays a crucial role in numerous NLP applications, such as information extraction, entity recognition, and machine translation. It allows algorithms to grasp the sentence's grammatical structure and disambiguate words with multiple meanings, thereby improving the accuracy of text classification and the identification of named entities

3.3 Topic Modeling

Topic modeling is a powerful unsupervised learning technique designed to discover and allocate topics within a given corpus of text data. In today's data-driven world, where vast amounts of information are expressed through textual content, effectively

Figure 3.7: Visualization of text data using wordcloud

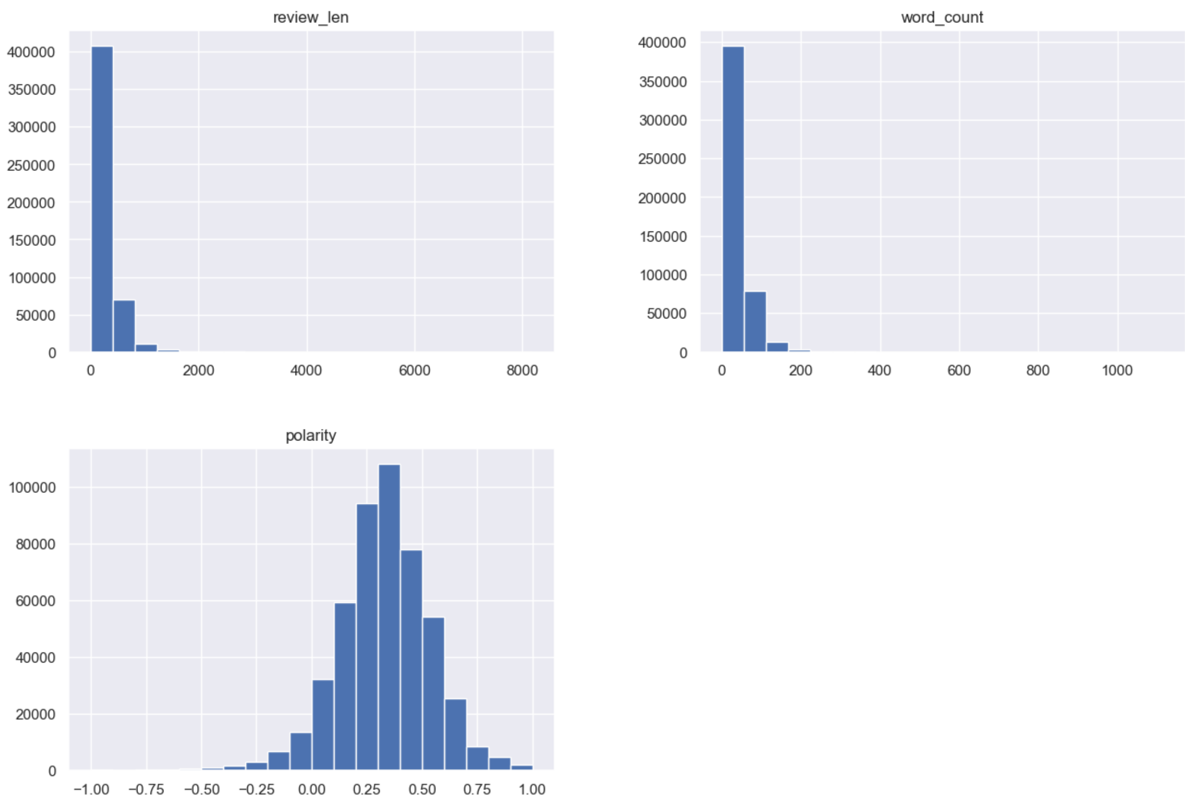

Figure 3.8: Histogram of review_len, word_count and polarity

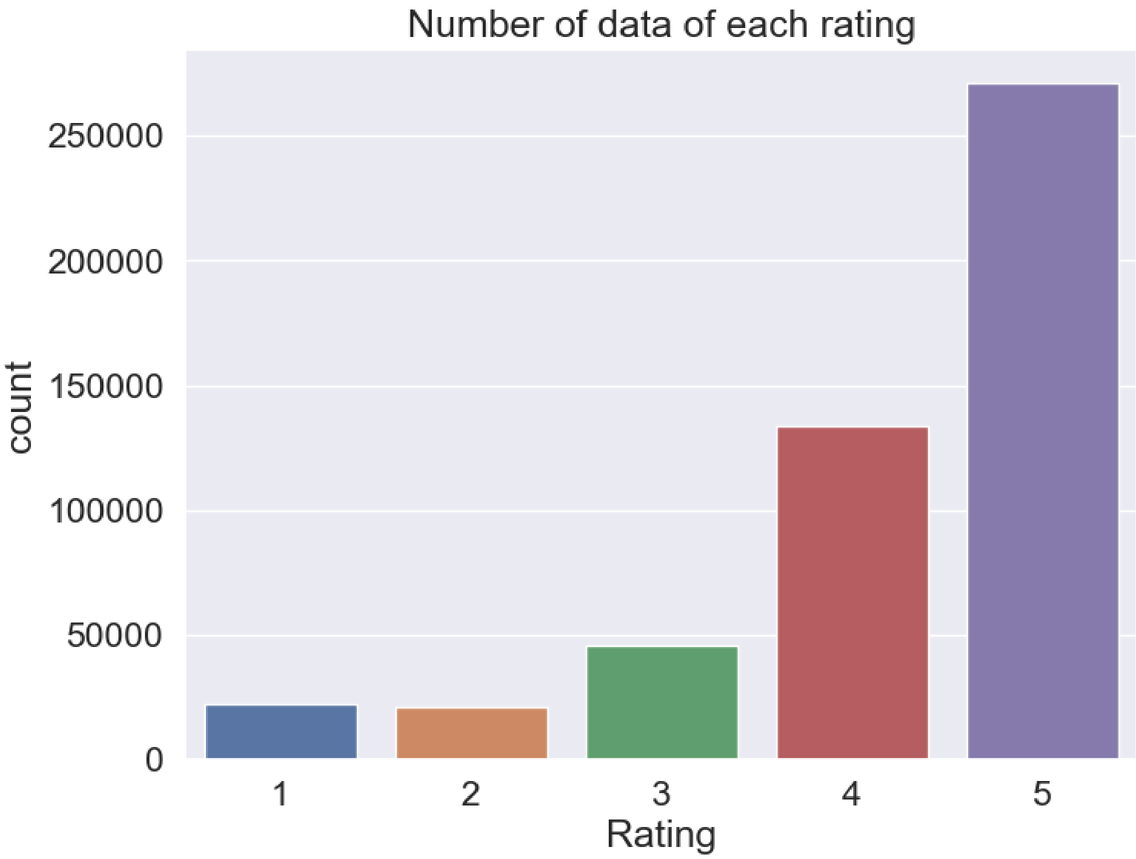

Figure 3.9: Number of data each rating

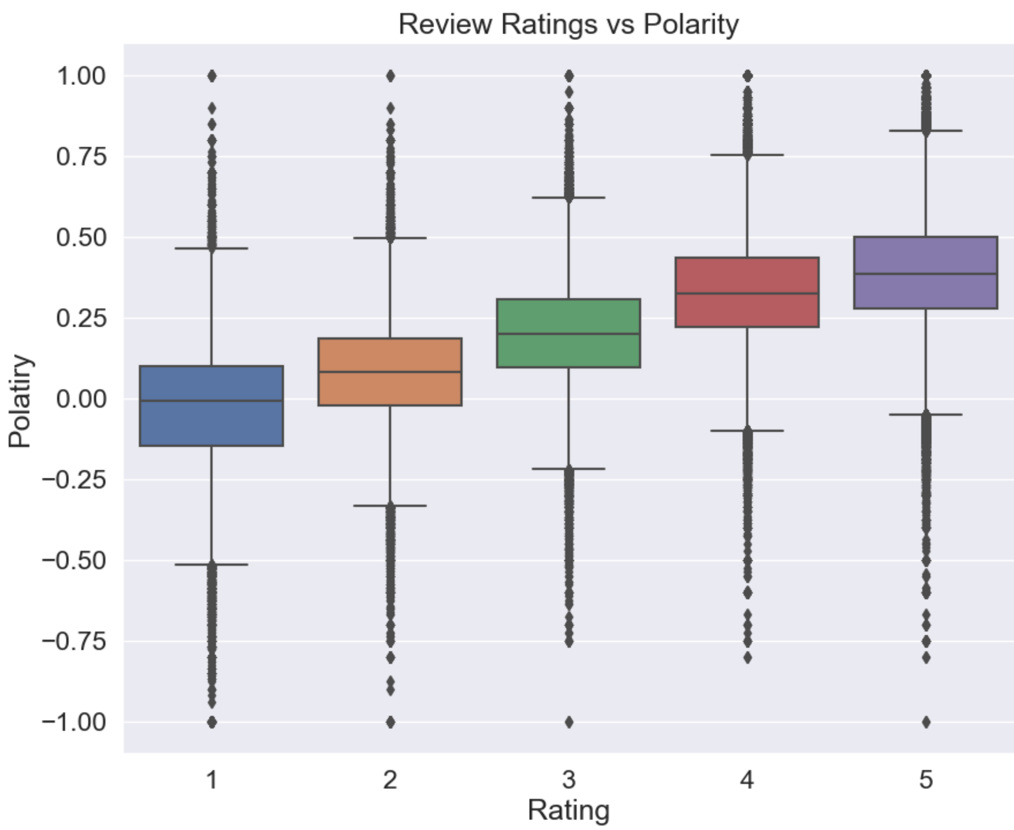

Figure 3.10: Box plot of review rating of vs. polarity.

Figure 3.11: Point plot of product rating vs review length

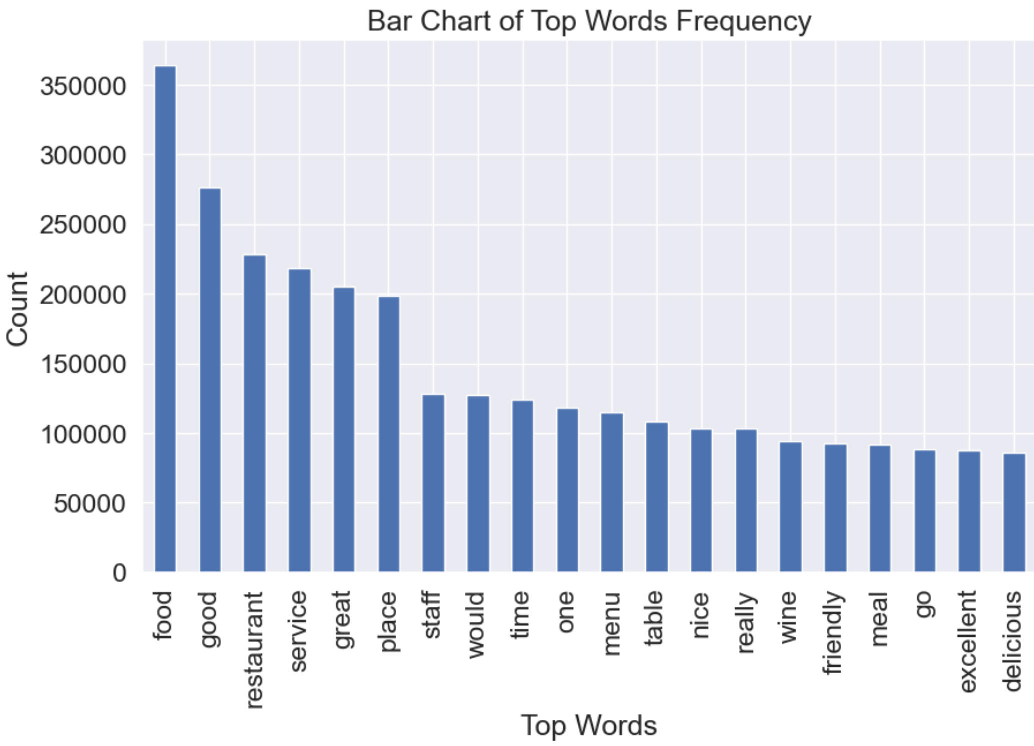

Figure 3.12: Bar chart of top world frequency

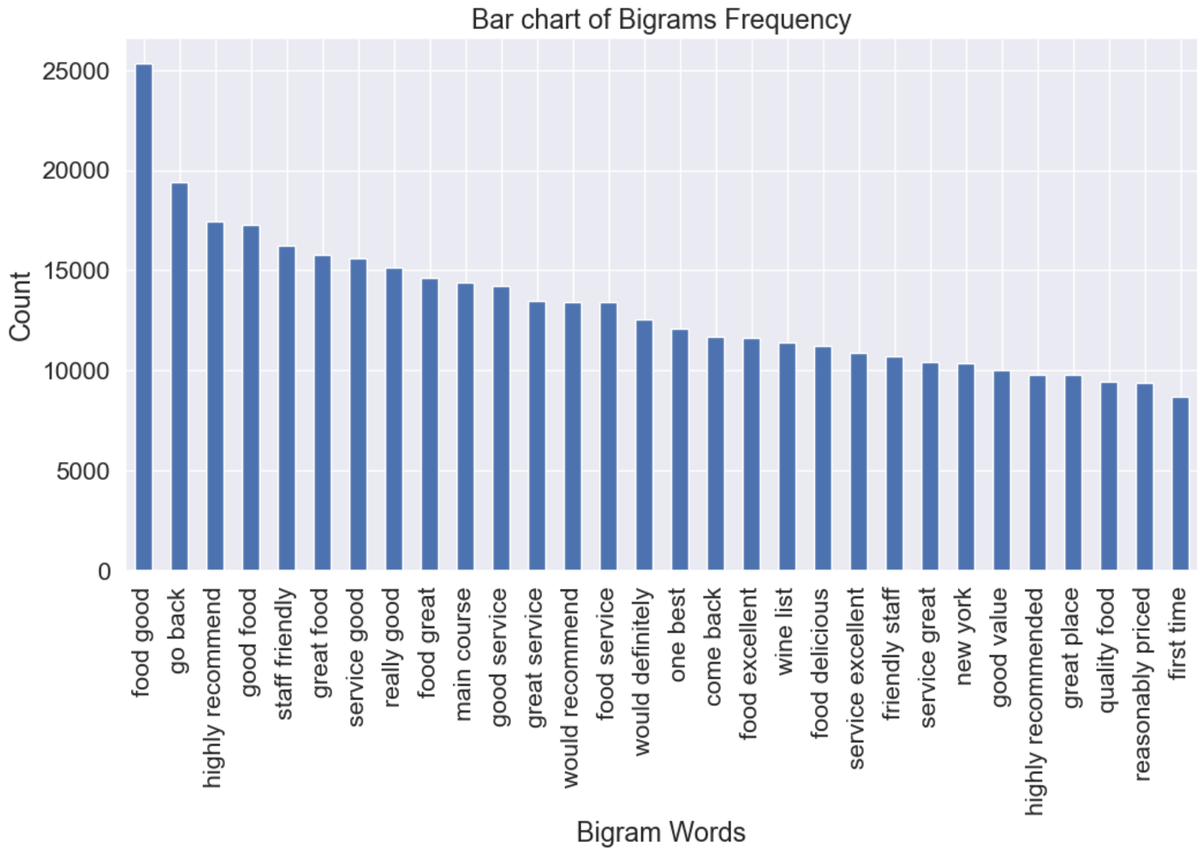

Figure 3.13: Bar chart of bigram frequency

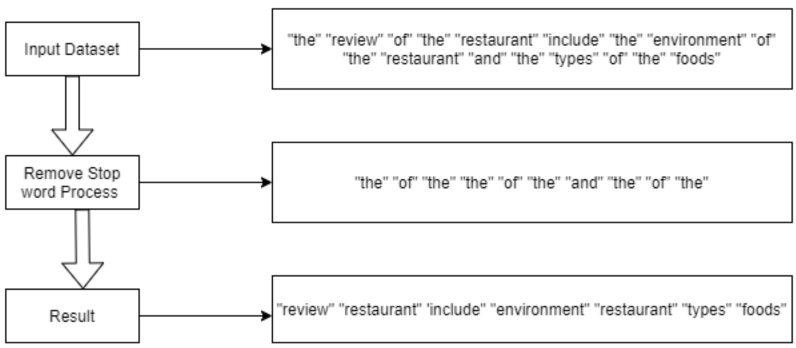

Figure 3.14: Stop words removal process

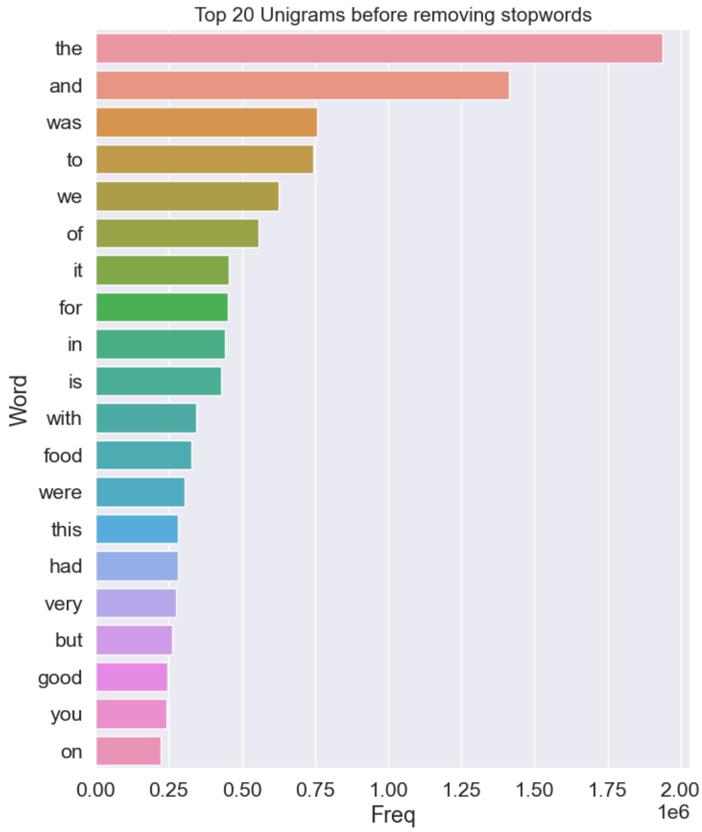

Figure 3.15: Top 20 unigrams before removing stopwords

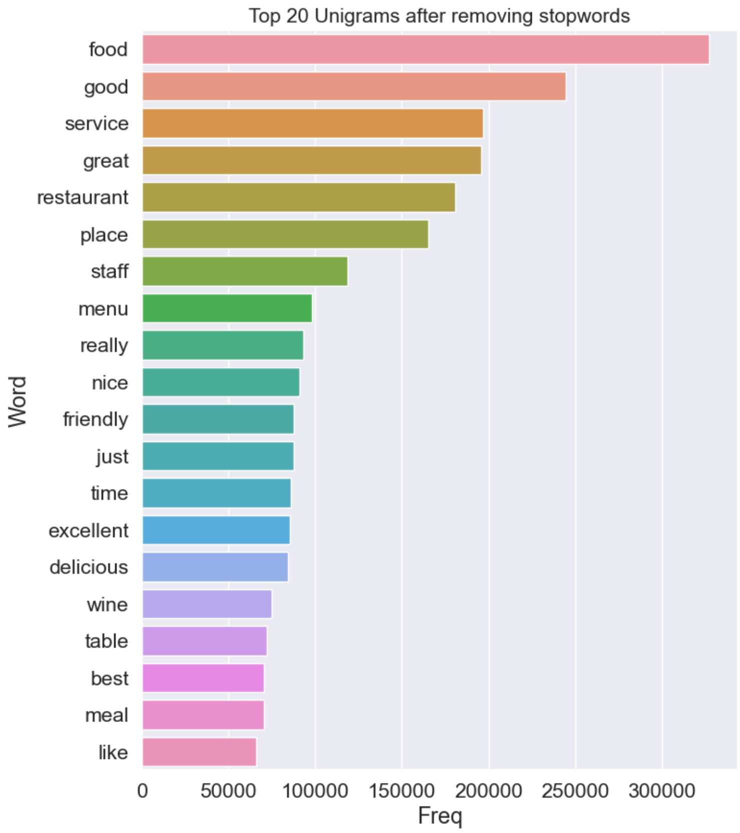

Figure 3.16: Top 20 unigrams after removing stopwords

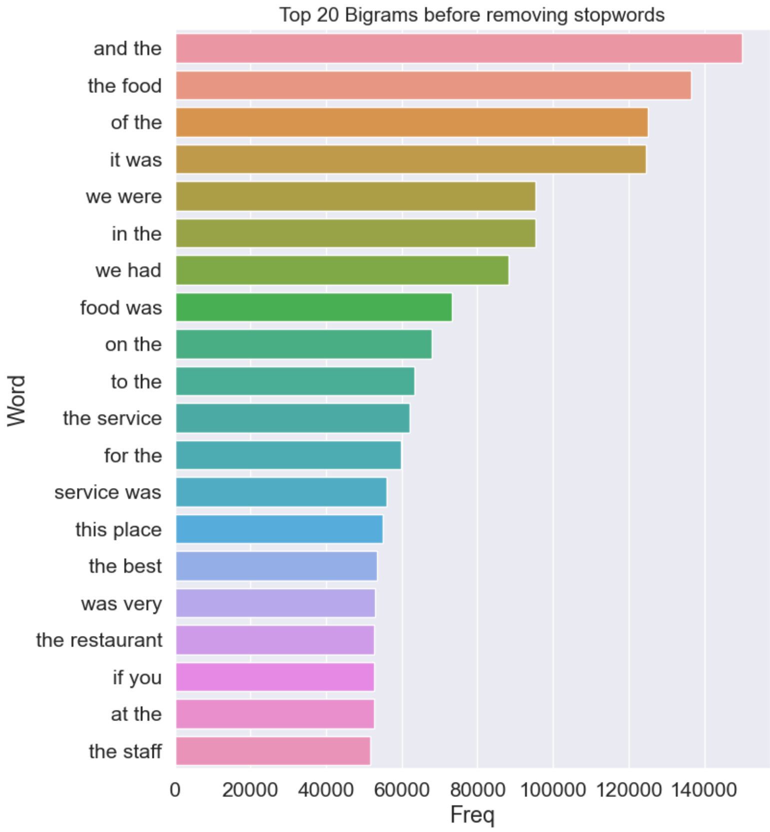

Figure 3.17: Top 20 bigrams before removing stopwords

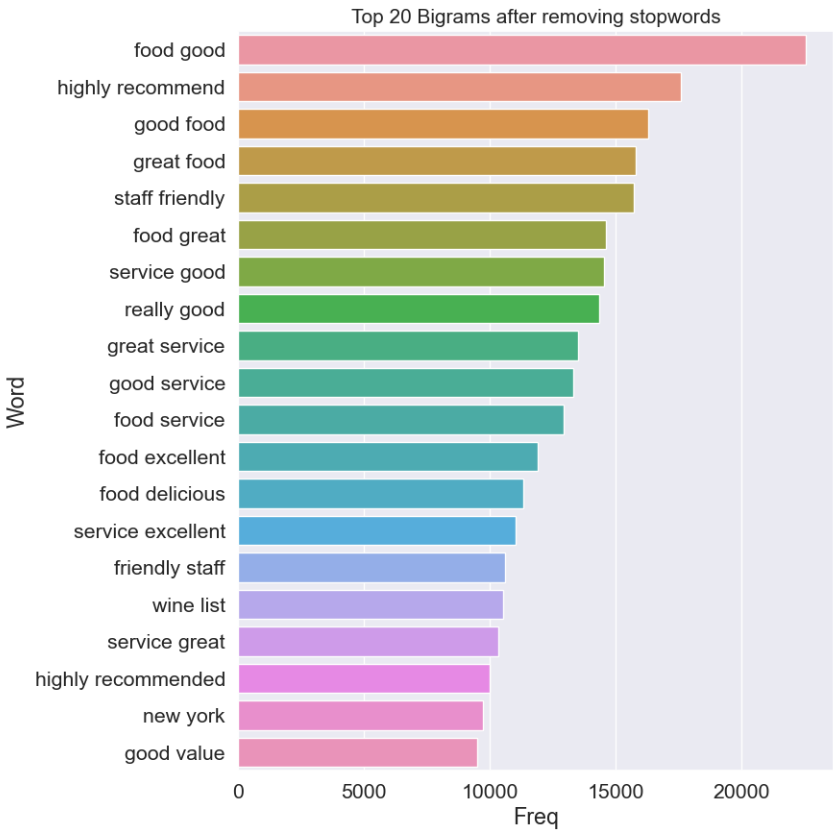

Figure 3.18: Top 20 bigrams after removing stopwords

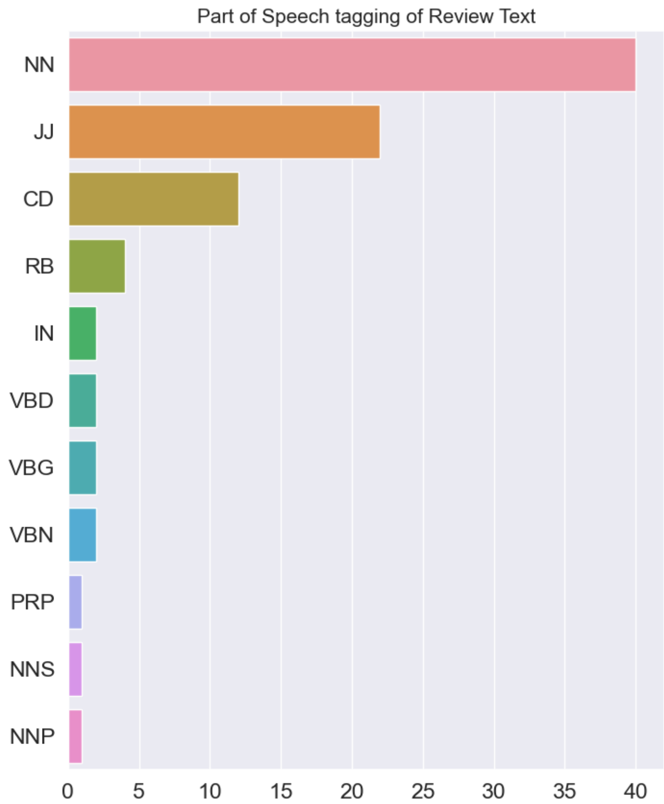

Figure 3.19: Part of speech tagging of Review Text

categorizing documents becomes crucial. This is where topic modeling steps in as a valuable tool[19].

Consider the scenario of a restaurant where customers leave reviews covering various aspects of their experience. Identifying all the specific features mentioned in these reviews manually can be an overwhelming task. However, by employing topic modeling, we can automatically categorize and extract key topics or themes from these reviews.

The process of topic modeling involves extracting essential keywords and patterns from each document within the corpus. By analyzing the words present in the text, the algorithm uncovers latent topics that are prevalent across the entire dataset[12]. These topics can represent various aspects, attributes, or themes discussed in the reviews, such as food quality, service, ambiance, pricing, and more.

Through topic modeling, we gain a structured and organized representation of the underlying content within the restaurant reviews. This categorization of features not only aids in understanding customer sentiments better but also facilitates decision-making processes for restaurant owners, managers, and potential patrons alike. Ultimately, topic modeling serves as a valuable tool to unlock valuable insights from the vast sea of textual data available in today's digital landscape.

3.3.1 Latent Dirichlet Allocation

A common method for topic modeling is Latent Dirichlet Allocation (LDA), which aims to reveal latent themes or subjects within a group of documents. The term

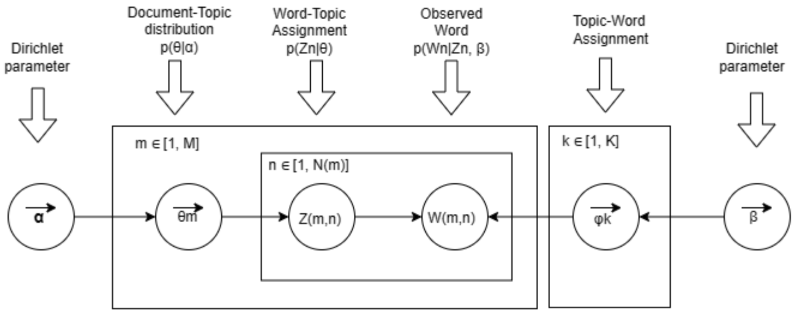

Figure 3.20: LDA Architecture

"Latent" alludes to the fact that these subjects are obscure and unexplored. The term "Dirichlet" denotes that the model assumes that the themes in the documents are distributed according to a Dirichlet distribution. The term "Allocation" refers to the process of assigning topics to the documents[26].

LDA is a powerful model of generative probabilistic reasoning specifically designed for analyzing discrete data, such as text corpora[2]. It forms a hierarchical Bayesian model consisting of three layers. Each item in the data, represented by words, is modeled as a composition of a finite number of topics. These topics, in turn, are modeled as an infinite composition of different topic probabilities, effectively capturing the context of the documents[1].

Each word in LDA is regarded as a distinct data point, expressed as a piece of vocabulary with an index of 1 to V. The model implies that documents are created via a statistical process that involves a variety of subjects and a variety of possible terms for each topic.

There are three sorts of hyper-parameters needed for LDA to function. The first one, alpha (α), regulates how many themes are included in the texts. The second hyperparameter, beta (β), controls how many words are allocated to each topic in the document. The number of topics to be extracted is represented by the third hyper-parameter, k[14].

In the absence of labeled data, it becomes necessary to draw certain inferences while selecting attributes for analysis[28]. Based on the aforementioned evaluation employing the Latent Dirichlet Allocation (LDA) technique, we have discerned several salient aspects that have been prominently cited in the majority of comments. These critical aspects identified through the LDA model are as follows:

1. Food (pertaining to the restaurant's culinary offerings, quality, taste, variety, and presentation of dishes)
2. Place (referring to the restaurant's ambiance or setting)
3. Service (pertaining to the quality of service provided)
4. Price (concerning the cost or pricing of the offerings)

These identified aspects hold significant importance in shaping the overall sentiment expressed in the comments and provide valuable insights for further analysis and understanding of customer perceptions within the context of the studied restaurant reviews.

We acquire the proportions of these 20 topics for each review in order to create a 20-vector feature representation that can be used to leverage the findings of LDA for supervised classification. This feature vector is intended to be used to ascertain whether a sentence exhibits a positive or negative mood.

3.3.2 Data Augmentation

In order to address the challenges of limited data and enhance the performance of our Opinion Mining models, particularly for aspect-based Opinion Mining, we employed data augmentation techniques using the NLPaug library. Data augmentation is a critical step in improving model generalization and robustness, especially when dealing with imbalanced datasets and underrepresented aspects[36].

NLPaug Library

We leveraged the NLPaug library, which offers a wide range of data augmentation methods specifically designed for natural language processing tasks[37]. These techniques aim to diversify the training data by generating new examples while preserving the original meaning and context of the text.

Aspect-Based Augmentation

Our data augmentation process was tailored to each of the four aspects under analysis: Food, Place, Service, and Price. For each aspect, we applied augmentation techniques individually to ensure that the generated data aligned with the specific aspect's context and sentiment[38].

Augmentation Techniques

1. Synonym Augmentation (Positive Sentiment): To capture the richness of positive sentiment expressions, we employed synonym augmentation. This technique involved replacing words in the original reviews with their synonyms, effectively expanding the vocabulary and allowing for diverse ways to convey positive sentiments related to each aspect[39].

2. Antonym Augmentation (Negative Sentiment): To accentuate the diversity of negative sentiment expressions, antonym augmentation was employed. In this process, we replaced words with their antonyms, enabling the model to grasp varying ways customers convey negative sentiments about the aspects.

3. Neutral Sentiment Preservation: To maintain the integrity of neutral sentiments, we refrained from applying augmentation techniques, keeping the text unaltered in cases where the original sentiment expression was neutral.

4. Review Length Standardization: To ensure uniformity in the augmented dataset, we enforced a consistent review length of 50 words per review. This standardization helped the model process and learn from the augmented data more effectively.

5. Augmentation Hyper-parameters: For the synonym and antonym augmentation techniques, we set the augmentation probability (aug_p) at 0.5. This balanced approach allowed for a substantial augmentation while still preserving some original sentiment expressions.

3.4 Classifications

3.4.1 Multinomial Naive Bayes

A prominent probabilistic classifier for text classification applications, such as Opinion Mining, is Multinomial Naive Bayes (MNB). Based on the Bayes theorem, it makes the assumption that given the class label, the features—in this example, the words—are conditionally independent. MNB is computationally effective and frequently performs well on text categorization problems despite its simplicity[40]. Let's break down the components of MNB for sentiment classification:

1. The Problem

In Opinion Mining, we seek to identify the sentiment polarity (positive, negative, or neutral) of a certain text document, such as a customer review or an online community post[6].

2. Representing the Text

We must translate the text data into a numerical format before we can apply MNB. The Bag-of-Words (BoW) format or the Term FrequencyInverse Document Frequency (TF-IDF) representation are frequently used for this[5].

- Term Frequency-Inverse Document Frequency (TF-IDF):
A numerical representation called TF-IDF takes into account the significance of each word across the board. The inverse document frequency (IDF), a measure of how uncommon the word is over the entire data set, and the word's frequency in the document (TF) are used to determine the weight of each word in each document[41].

3. The Naive Bayes Classifier

The Bayes theorem[42], which determines the conditional probability of a class label given the observed features, is the foundation of the Naive Bayes classifier. Let's denote: C: Class label (e.g., positive, negative, neutral)
X: The features (in our case, TF-IDF representation of the text document)

Bayes' theorem states:

$$P(\frac{C}{X}) = \frac{P(\frac{X}{C}) \cdot P(C)}{P(X)} \tag{3.1}$$

where:
$P(\frac{C}{X})$: The posterior probability of class C given features X (the probability that the document belongs to class C given its features).
$P(\frac{X}{C})$: The likelihood of observing features X given class C (the probability of seeing the TF-IDF representation X given the class label C).
$P(C)$: The prior probability of class C (the probability of a document belonging to class C, independent of its features).
$P(X)$: The evidence probability (the probability of observing the features X).

4. The Multinomial Naive Bayes Model

In MNB, we assume that the features (words) are multinomially distributed within each class. This means that for each class C, the likelihood $P(\frac{X}{C})$ follows a multinomial distribution[43]. The likelihood$P(\frac{X}{C})$ can be estimated using the TF-IDF representation. For TF-IDF, it is as follows:

$$P(\frac{X}{C}) = \prod (P(\frac{w_i}{C}))^{\text{TF-IDF}(w_i)} \tag{3.2}$$

where:
w_i: The i-th word in the vocabulary V.
$P(\frac{w_i}{C})$: The probability of observing word w_i given class C.
$TF - IDF(w_i)$: The frequency of word w_i in the document X.

This formula takes into account the TF-IDF scores of words in the document when estimating the likelihood of the document belonging to class C in a Multinomial Naive Bayes classifier using TF-IDF representation.

5. Laplace Smoothing

One issue with MNB is that if a word in the test document has not been seen in the training data for a particular class, its probability will be zero. To avoid this problem, we apply Laplace smoothing (additive smoothing), which adds a small constant (usually denoted by alpha) to each word count. This ensures that no word has a zero probability[13].

The smoothed probability of observing word w_i given class C is given by:

$$P\left(\frac{w_i}{C}\right) = \frac{\text{TF-IDF}(w_i) + \alpha}{N + \alpha \cdot |V|} \tag{3.3}$$

where:
α: The smoothing parameter.

N: Sum of TF-IDF scores of all words in document X.
$TF - IDF(w_i)$: This is the TF-IDF score of word wi in the document.
The denominator is the sum of TF-IDF scores of all words in the document X.
$|V|$: This is the size of the vocabulary.

6. Opinion Mining using MNB

Once we have calculated the likelihood $P(\frac{X}{C})$ for each class C, we can use Bayes' theorem to compute the posterior probability $P(\frac{C}{X})$ for each class. The class with the highest posterior probability is chosen as the predicted sentiment label for the input text document[16].

$$P(\frac{C}{X}) = \frac{P(\frac{X}{C}) \cdot P(C)}{P(X)} \tag{3.4}$$

We calculate $P(\frac{C}{X})$ for all classes and select the class with the highest probability as the predicted sentiment label.

In summary, MNB for Opinion Mining involves representing text documents using TF-IDF, estimating the likelihood $P(\frac{X}{C})$ using multinomial distribution assumptions, and applying Bayes' theorem to classify the sentiment polarity of the text documents[44]. The use of Laplace smoothing helps handle words not seen in the training data and improves the classifier's robustness.

3.4.2 Random Forest

Opinion Mining is one classification problem that uses the ensemble learning technique known as Random Forest. During the training stage, numerous decision trees are constructed, and their predictions are then combined to arrive at a final classification choice. Random Forest is a reliable and well-liked option for text classification problems because of its propensity to handle high-dimensional input and minimize overfitting[45].

1. The Problem

In Opinion Mining, as previously noted, we seek to ascertain the sentiment polarity (positive, negative, or neutral) of a particular text document, such as a customer review or social media post[46].

2. Data Representation

Similar to the previous explanations, the text data needs to be represented in a numerical format using the Term Frequency-Inverse Document Frequency (TF-IDF) representation[41].

3. Random Forest

An assortment of decision trees makes up a random forest. A random subset of the training data and a random subset of the characteristics are used to train each

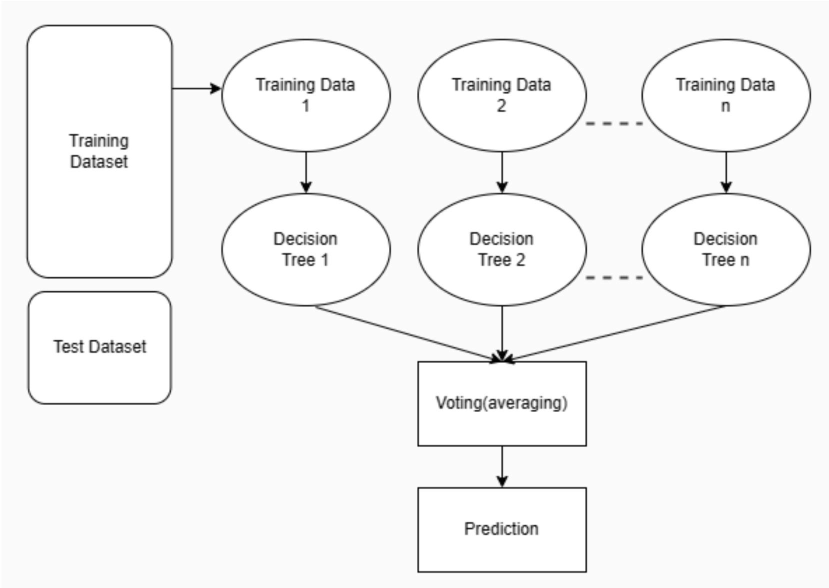

Figure 3.21: Random forest algorithm.

decision tree. By building different and less linked decision trees, Random Forest's main goal is to reduce overfitting[47].

3.1 Decision Trees

A decision tree is a hierarchical structure in which each leaf node represents a class label (in our example, sentiment polarity) and each interior node reflects a judgment based on a particular attribute[48].

3.2 Bootstrap Aggregating (Bagging)

The method used by Random Forest is known as Bootstrap Aggregating or Bagging. By sampling the training data with replacement samples (bootstrap samples), bagging entails constructing several datasets. Then, a different decision tree is trained for each dataset[49].

3.3 Feature Randomness

Only a random subset of features is taken into account at each node of each decision tree as it is being built. The trees are further adorned by this feature unpredictability, which also lowers the likelihood of overfitting[50].

4. Training Random Forest

Let's denote: X: The training data is represented using TF-IDF. Y: The corresponding sentiment labels (positive, negative, or neutral).

The steps to train a Random Forest are as follows:

Step 1: Initialize the number of decision trees $n_{\text{estimators}}$, the maximum depth of each tree (max_{depth}), and the number of features to consider at each node (max_{features}).

Step 2: For each decision tree (t = 1 to $n_{\text{estimators}}$): Create a bootstrap sample X_t and Y_t by randomly selecting data points with replacements from X and Y.
Train a decision tree T_t on the bootstrap sample $(X_t,\ Y_t)$, limiting its depth to max_{depth} and considering only max_{features} features at each node[51].

Step 3: Predictions from the Random Forest
To make predictions for a new text document, the Random Forest combines the predictions of all its decision trees. The final prediction is obtained through a majority vote among the decision trees for classification tasks like Opinion Mining[50].

5. Opinion Mining

The Random Forest can forecast the sentiment polarity of new text documents once it has been trained on the preprocessed and feature-engineered data. The Random Forest's ensemble of decision trees aids in lowering overfitting and enhancing model generalization[50].

In conclusion, Random Forest is an ensemble learning technique that generates several decision trees using feature randomness and bootstrap aggregation in order to produce a reliable and precise sentiment categorization model. Random Forest can handle high-dimensional text input and produce high classification performance for Opinion Mining applications by utilizing a variety of decision trees[49].

3.4.3 Support Vector Machine

One widely used and effective method in supervised machine learning for tasks like Opinion Mining is the support vector machine (SVM). It operates in a high-dimensional feature space to find the best hyperplane that can separate data points belonging to different classes most effectively. SVM relies on two key elements: the margin and the kernel function[52].

1. The Problem

In opinion categorization, we seek to identify the opinion polarity (positive, negative, or neutral) of a certain text document, such as a consumer review or a social network post[53].

2. Data Representation

Similar to the above rationales, SVM must be used using a quantitative representation of the linguistic sample. We shall employ the Term Frequency-Inverse Document Frequency (TF-IDF) format[54].

3. SVM with RBF Kernel

The non-linear SVM variation with an RBF (Radial Basis Function) kernel enables more intricate boundary conditions for decisions. Using a kernel function, it transforms the data points into a higher-dimensional space, where it then looks for the most optimal hyperplane[55].

The RBF kernel function is defined as:

$$K(x_i, x_j) = \exp(-\gamma \cdot ||x_i - x_j||^2) \tag{3.5}$$

where:
γ : a hyperparameter that regulates the Gaussian kernel's width[56].
$||x_i - x_j||$: The Euclidean distance between data points x_i and x_j.

The decision function of SVM with RBF kernel is given by:

$$f(x) = \sum \alpha_i \cdot y_i \cdot K(x_i, x) + b \tag{3.6}$$

where:
α_i : The Lagrange multipliers (obtained during training)[57].
y_i : The class label of the $i - th$ document (positive, negative, or neutral).

4. Training SVM

SVM seeks to identify the best hyperplane throughout training which optimizes the distance across each point of information of various categories[58]. The method in question necessitates minimizing the subsequent objective function in order to solve a convex optimization problem:

$$\text{minimize} = \frac{1}{2}||w||^2 + C \sum \max\left(0, 1 - y_i \cdot (w \cdot x_i + b)\right)^2 \tag{3.7}$$

where:
$||w||$: The norm of the weight vector w.
C : The regularization parameter that controls the trade-off between maximizing the margin and minimizing the classification error.
$\sum$: The summation of all training data points[59].

5. Opinion Mining with SVM

We train the SVM model using the preprocessed and feature-engineered data (TF-IDF representation) together with their matching opinion labels in order to conduct opinion Mining using SVM[44]. By determining the decision function for each text document and determining the class label based on the sign of the decision Value, the model may predict the opinion polarity of fresh text documents once it has been trained[60].

In conclusion, SVM employing RBF kernel projects the data into a higher-dimensional space, allowing for non-linear decision boundaries. Model effectiveness may differ depending on the jobs' requirements for opinion categorization, the difficulty of the opinion mining task, and the kernel parameter selection (For instance, the RBF kernel's gamma)[61].

3.4.4 Deep Learning Algorithm: RNN

Assessing the sentiment polarization toward particular aspects or characteristics stated in a text document is frequently necessary for aspect-based opinion mining. To accomplish this and improve the efficiency of opinion mining, sophisticated Recurrent Neural Network (RNN) architectures like Double-layered LSTM and Bidirectional LSTM (BiLSTM) are combined with pre-trained word embeddings like GloVe[62].

1. GloVe Word Embeddings

GloVe, a widely utilized pre-trained word embedding technique, encodes the nuanced connections among words through an analysis of their co-occurrence patterns within a vast textual dataset. It accomplishes this by transforming words into compact vectors within a continuous-dimensional space[63].

2. Double-layered (Stacked) LSTM

A Stacked LSTM, also referred to as a Double-layered LSTM, employs a configuration where two LSTM layers are placed one on top of the other. In this setup, each subsequent LSTM layer uses the output of the preceding layer as its input. This stacking of LSTM layers enhances the model's capability to discern intricate patterns and extended connections within the input data, thereby improving its ability to capture intricate information[62].

The hidden state (h_t) of a Double-layered LSTM at time step t is computed as follows[62]:

$$h_t^1 = \mathrm{LSTM}_1(x_t, h_{t-1}^1) \tag{3.8}$$

$$h_t^2 = \mathrm{LSTM}_2(h_t^1, h_{t-1}^2) \tag{3.9}$$

where:
x_t : The input vector (word embedding)[64] at time step t.
h_t^1 : The hidden state of the first LSTM layer at time step t.
h_t^2 : The hidden state of the second LSTM layer at time step t.
h_{t-1}^1 : The hidden state of the first LSTM layer from the previous time step.
h_{t-1}^2 : The hidden state of the second LSTM layer from the previous time step.
LSTM_1 : The first LSTM layer.
LSTM_2 : The second LSTM layer.

The final aspect-based Opinion Mining can be performed using the output of the second LSTM layer (h_t^2) or by applying additional layers (e.g., fully connected layers) on top of it for sentiment prediction[65].

3. Bidirectional LSTM (BiLSTM)

It is crucial to comprehend the context around a particular aspect while conducting an aspect-based opinion study. As a type of recurrent neural network (RNN) that can process input sequences in both forward and backward directions, the Bidirectional LSTM (also known as BiLSTM) stands out. It can successfully acquire insights from both the past and the future because of its dual-directional strategy, which results in a comprehensive understanding of the input sequence[18].

The hidden state (h_t) of a BiLSTM at time step t is computed as follows[66]:

$$h_t^f = \text{LSTM}_f(x_t, h_{t-1}^f) \tag{3.10}$$

$$h_t^b = \text{LSTM}_b(x_t, h_{t-1}^b) \tag{3.11}$$

where:
h_t^f : The hidden state of the forward LSTM (processing input from the beginning to the end of the sequence) at time step t.
h_t^b : The hidden state of the backward LSTM (processing input from the end to the beginning of the sequence) at time step t.
LSTM_f : The forward LSTM layer.
LSTM_b : The backward LSTM layer.

The final hidden state of the BiLSTM (h_t) is typically obtained by concatenating the hidden states of the forward and backward LSTMs:

$$h_t = [h_t^f, h_t^b] \tag{3.12}$$

The final aspect-based Opinion Mining can be performed using the hidden state (h_T) or by applying additional layers (e.g., fully connected layers) for sentiment prediction[67].

4. GloVe and Aspect-Based Opinion Mining with Stacked LSTM and BiLSTM

Each word in the original stream is transferred to its appropriate GloVe vector for the purpose of employing pre-trained GloVe word embeddings. The Stacked LSTM or BiLSTM opinion mining model, as described in the prior sections, will then incorporate these linguistic representations[63].

By using advanced architectures like Double-layered LSTM and Bidirectional LSTM in combination with pre-trained GloVe word embeddings, aspect-based Opinion Mining models can effectively capture contextual information and long-term dependencies, leading to improved Opinion Mining performance, especially in tasks where understanding the sentiment towards specific aspects is critical.
$p(W_n|Z_n, \beta)$

Chapter 4

Result & Analysis

In this study, we conducted research into Aspect-Based Opinion Mining (ABOM) within the context of restaurant reviews. Our objective was to uncover nuanced sentiments related to distinct aspects: Food, Place, Service, and Price. To tackle the challenges associated with large and imbalanced datasets, we devised a multi-phased investigation, each phase addressing specific domain challenges. Our primary goal was to improve the performance of conventional models while leveraging the capabilities of neural networks for aspect-based opinion mining (ABOM). This report provides a thorough analysis of our discoveries and critical insights gleaned from our experimentation.

4.1 Initial Analysis on Large-Scale Dataset

We commenced our research by working with a massive dataset comprising 2.7 million restaurant reviews from diverse global locations. Employing three classical machine learning algorithms—Multinomial Naive Bayes (MNB), Support Vector Machine with Radial Basis Function Kernel (SVM RBF), and Random Forest (RF)—we initially faced computational constraints. Consequently, we downsized the dataset to 500k samples to make SVM computations feasible. Despite achieving high F1 scores(Table: 4.1), we observed a high degree of misclassification in the confusion matrices, particularly MNB exhibited difficulties in classifying target 0, while SVM-RBF and RF showed a tendency to misclassify target 0 as well as neighboring target ratings (See Figure 4.1, 4.2 & 4.3). The dataset's inherent class imbalance, with an overwhelming majority of 5-star reviews, contributed to these challenges.

Algorithm	Accuracy	Precision	Recall	F1 Score
SVM	0.91	0.91	0.91	0.91
MNB	0.90	0.90	0.90	0.90
RandomForest	0.90	0.90	0.90	0.90

Table 4.1: Performance Evaluation on 500K Data

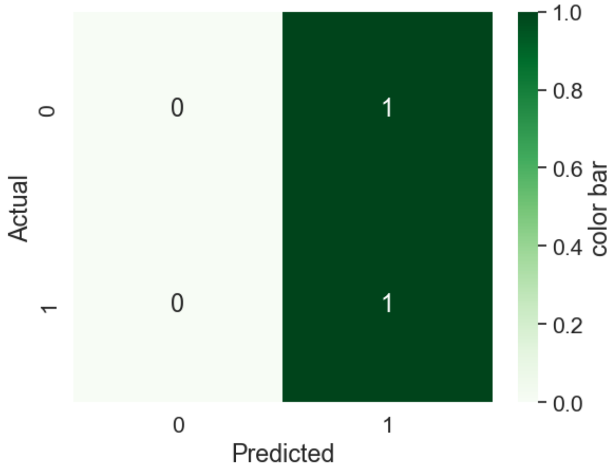

Figure 4.1: Confusion Matrix for Naive Bayes

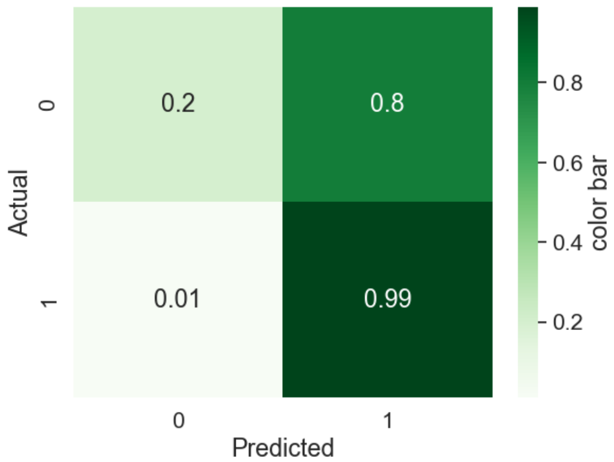

Figure 4.2: Confusion Matrix for SVM

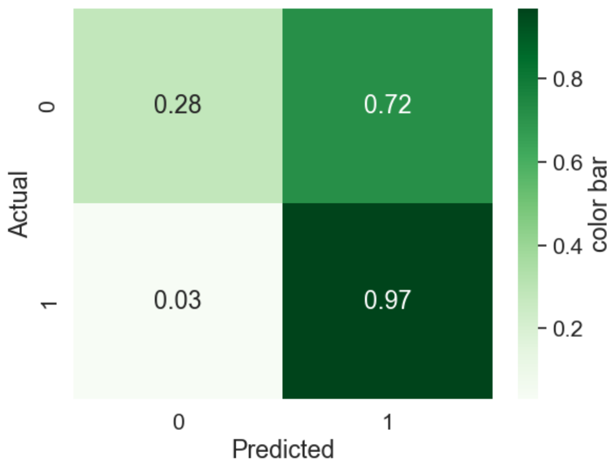

Figure 4.3: Confusion Matrix for Random Forest

4.2 Balancing the Dataset

Our initial dataset consisted of 2.7 million restaurant reviews, and it displayed a considerable class imbalance, with the majority of reviews falling into the 4 and 5-star categories. To address this imbalance and ensure fairness in our analysis, we meticulously curated a balanced dataset containing 100,000 reviews, evenly distributed among 1 to 5-star ratings. This deliberate balancing effort was undertaken to eliminate any potential biases in our models' predictions, acknowledging the stark disparity in positive reviews compared to negative and mixed ones in the original dataset.

4.3 Model Reevaluation on Balanced dataset & Introduction of Neural Networks

A series of assessments were then carried out to evaluate the effectiveness of three classic models (Naive Bayes, SVM, and Random Forest) as well as two neural networks (Double Stack LSTM and Double layered BiLSTM). Our main objective is to identify the most efficient set of settings. These variables included changes in the number of hidden layer dimensions, learning rates, and training epochs that would produce the best results. Five classifiers were used in the experiment, two of which were neural networks using GloVe word embeddings, and three of which were classic models using TF-IDF.

To ensure robust evaluation, we divided the dataset into a 60% training set, a 30% validation set, and a 10% test set. By analyzing the results based on accuracy, precision, recall, and F1 score obtained from the testing phase, we aimed to identify which models performed better and which parameter settings were most effective.

The findings of this study will offer valuable insights into the performance of different models and parameter combinations, thus aiding in selecting the most suitable approach for aspect-based opinion classification in NLP tasks.

4.3.1 Training Process on Balanced Dataset

The training, validation, and test sets were afterward generated from this balanced dataset. In our original strategy, word vectorization was used to change the training set's text data. Three classic models were trained using this format. The training data for our neural network models was then prepared. In particular, we made use of GloVe word embeddings and set their dimensionality to 300. We used the double-stack LSTM and the Double layered BiLSTM neural networks in this configuration. With a learning rate of 0.001, stochastic gradient descent was used to fine-tune these neural networks. We performed training sessions across 500 epochs with early halting as a safety measure in order to facilitate model training and reduce overfitting. Our goals included the development of a balanced dataset, the improvement of the classic model's performance, and the effective use of neural networks for opinion classifications.

4.3.2 Performance on Balanced Dataset

4.3.2.1 Classical Models

In the initial experiment, we processed the sentence tokens by converting them into word vectors, a common practice in supporting neural networks for NLP tasks. This approach is widely used in the field. By evaluating the performance of word vectors independently from POS tags and word dependencies, we aimed to compare the effectiveness of this conventional method with the proposed techniques. To begin, the sentences were tokenized, followed by the conversion of these tokens into word vectors. Subsequently, these word vectors were utilized as input for the three linear classifiers. The results of this experiment are presented in Table 4.2, allowing us to assess the performance of the classifiers based on word vectorization alone.

Models	Accuracy
Multinomial NB	0.56
Random Forest	0.56
SVM	0.57

Table 4.2: Classic Algorithm Result for 100K Data

4.3.2.2 RNN Models

For the subsequent experiment, we opted to use neural networks. Specifically, we employed the Glove word embedding, which consists of word representations in a hundred-dimensional space. Word embeddings have proven to be highly effective in handling text data. The utilization of neural networks is expected to result in higher accuracy compared to linear classifiers due to their capacity to capture more intricate patterns and relationships in the data.

To explore the performance of neural networks, we implemented two distinct architectures. The results of these experiments are presented in Table 4.3, allowing us to evaluate the effectiveness of the neural network models in opinion mining .

The Double Layered LSTM model exhibited moderate accuracy in predicting the sentiment of restaurant reviews. While not achieving high accuracy, it demonstrated an improvement compared to the 500k dataset version. However, the model still had difficulty distinguishing between adjacent star ratings, indicating room for further enhancement. See Table 4.3 and Figure 4.7. The Double Layered BiLSTM outperformed the Double Layered LSTM slightly in terms of accuracy. Similar to the other models, it showed the tendency to misclassify reviews with adjacent star ratings. Further analysis of the confusion matrices may provide insights into specific areas for improvement. See Table 4.3 and Figure 4.10.

Models	Accuracy
Double Stack LSTM	0.59
Double Layered BiLSTM	0.60

Table 4.3: Classic Algorithm Result for 100K Data

4.4 Aspect-Based Opinion Mining

Recognizing the importance of aspect-based opinion mining (ABOM), we employed Latent Dirichlet Allocation (LDA) to identify four key aspects: Food, Place, Service, and Price within our input data. We created a separate dataset of 5,000 three-star reviews as they typically exhibit a more diverse and mixed sentiment. Within this new dataset, we manually annotated the aspects and labeled them based on positive, negative, or neutral sentiment. In instances where customers did not provide any explicit opinion or sentiment on a particular aspect, we labeled them as ”N/A” to indicate the absence of their feedback regarding that specific aspect.

4.4.1 ABOM Using BiLSTM

In the context of restaurant review analysis focused on different aspects, we proceeded by segmenting our annotated dataset into three essential subsets: training, validation, and test sets. These sets were essential for the training phase of our BiLSTM model, which aimed to discern nuanced sentiments associated with each specific aspect of restaurant reviews.

For the BiLSTM model tailored to restaurant review aspect-based opinion mining (ABOM), we opted for stochastic gradient descent as our optimizer, configuring a learning rate of 0.001. Our choice for the loss function was binary cross-entropy. The BiLSTM model underwent rigorous training on this dataset, with the primary goal of capturing the intricate sentiments linked to individual aspects of restaurant reviews.

We followed a consistent approach when training the model for other aspects within the restaurant review domain. The outcome of these experiments is summarized in the Table , providing a comprehensive overview of the BiLSTM model's performance in aspect-based opinion mining (ABOM) for various aspects specific to restaurant reviews. Due to the limited size of the annotated dataset, the initial model suffered from overfitting. To mitigate this, we augmented the data using NLP techniques, resulting in improved performance across all aspects (Table 4.8).

4.4.2 Aspect-Specific Performance

Food:

The BiLSTM demonstrated strong performance in classifying Food sentiment with an F1 score of 87.36%, showcasing its ability to capture nuanced opinions about food quality. See Table 4.4

Place:

For the Place aspect, the model achieved an impressive F1 score of 92%, indicating its proficiency in discerning sentiments related to the restaurant's ambiance and location. See Table 4.5

Service:

The Service aspect also yielded a high F1 score of 92%, highlighting the model's capability to understand customer feedback regarding service quality. See Table 4.6

Price:

The Price aspect proved to be challenging, with an F1 score of 32%. The model struggled to accurately predict price-related sentiments, indicating the need for further refinement in this area. See Table 4.7.

It's important to note that while the accuracy of the neural network models may not be exceptionally high, the aspect-based sentiment analysis results are more promising. The models effectively captured nuanced sentiments associated with aspects like Food, Place, and Service, which are critical in restaurant reviews. The Price aspect, being more challenging, could be a focus for future improvements and refinements in the model architecture or data preprocessing.

Ratings	Precision	Recall	F1-score
Negative	0.88	0.85	0.87
Neutral	0.84	0.92	0.88
Positive	0.90	0.85	0.88
Macro avg	0.88	0.87	0.87
Weighted avg	0.88	0.87	0.87

Table 4.4: Result analysis for Food

Ratings	Precision	Recall	F1-score
Negative	0.91	0.93	0.92
Neutral	0.92	0.95	0.93
Positive	0.94	0.90	0.92
Macro avg	0.92	0.92	0.92
Weighted avg	0.92	0.92	0.92

Table 4.5: Result analysis for Place

Ratings	Precision	Recall	F1-score
Negative	0.93	0.90	0.92
Neutral	0.90	0.97	0.93
Positive	0.94	0.89	0.91
Macro avg	0.92	0.92	0.92
Weighted avg	0.92	0.92	0.92

Table 4.6: Result analysis for Service

Ratings	Precision	Recall	F1-score
Negative	0.44	0.06	0.10
Neutral	0.35	0.42	0.38
Positive	0.38	0.63	0.48
Macro avg	0.39	0.37	0.32
Weighted avg	0.39	0.37	0.32

Table 4.7: Result analysis for Price

Aspects	Accuracy
Food	0.87
Place	0.92
Service	0.92
Price	0.37

Table 4.8: F1-Score Comparison for Aspects

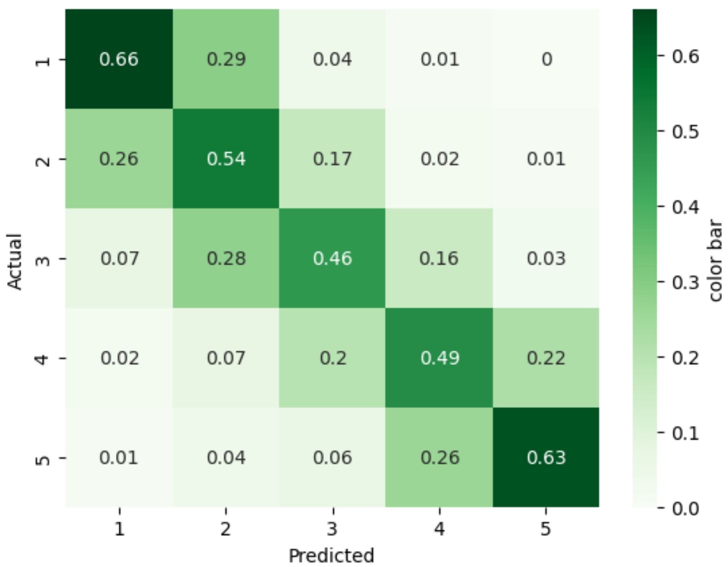

Figure 4.4: Confusion Matrix for MNB(100K Data)

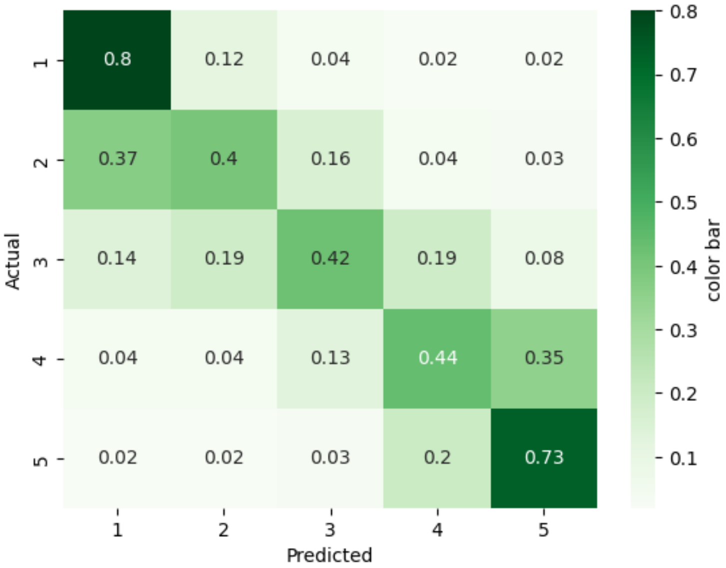

Figure 4.5: Confusion Matrix for Random Forest(100K Data)

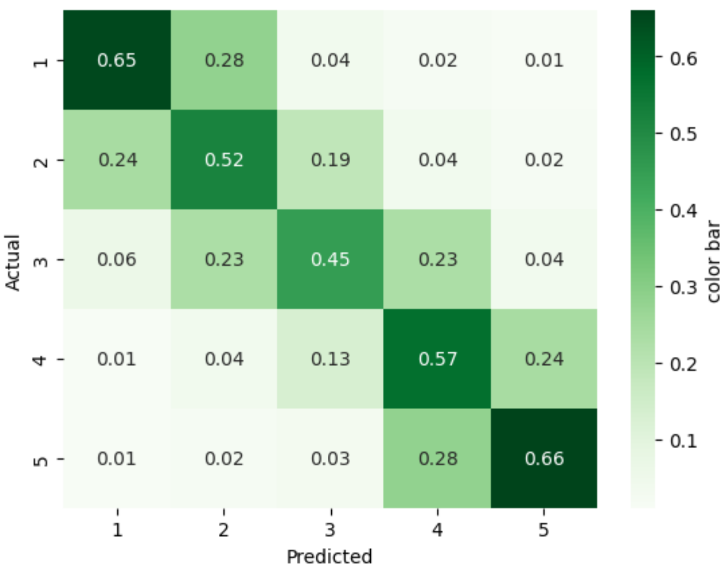

Figure 4.6: Confusion Matrix for SVM(100K Data)

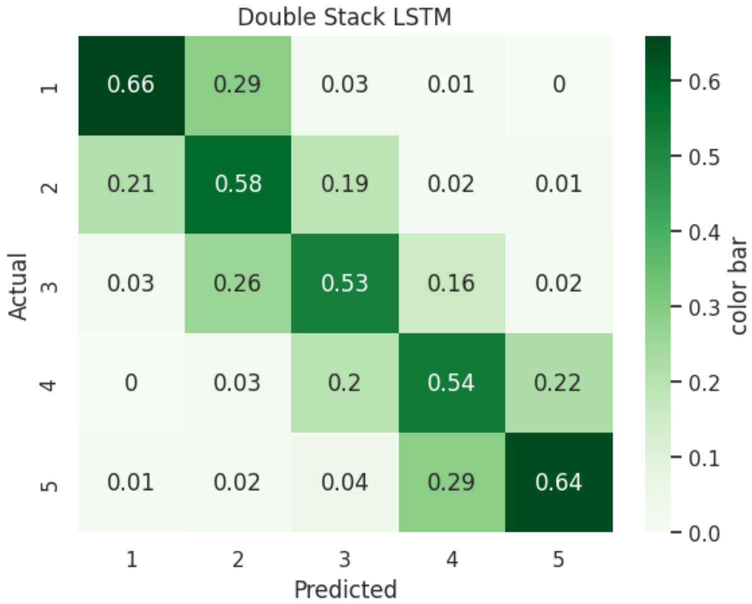

Figure 4.7: Confusion Matrix for Double Stacked LSTM(100K Data)

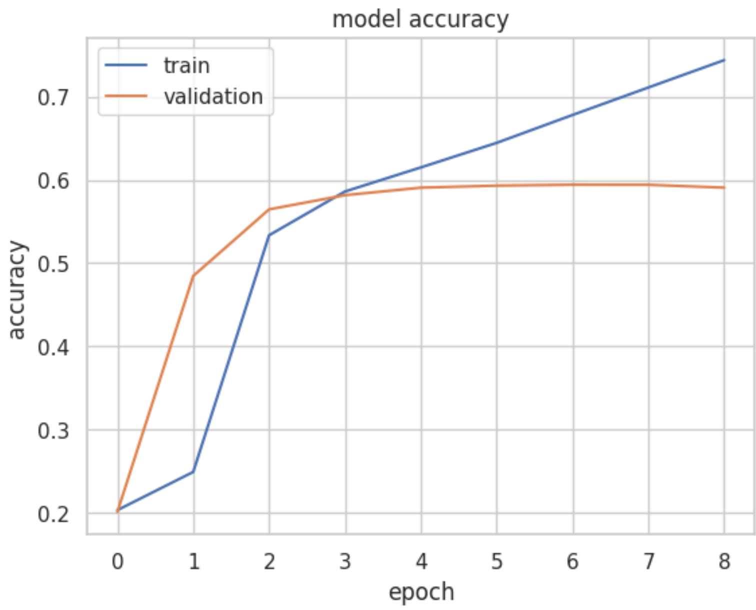

Figure 4.8: Double Stacked LSTM Accuracy for 100K Data

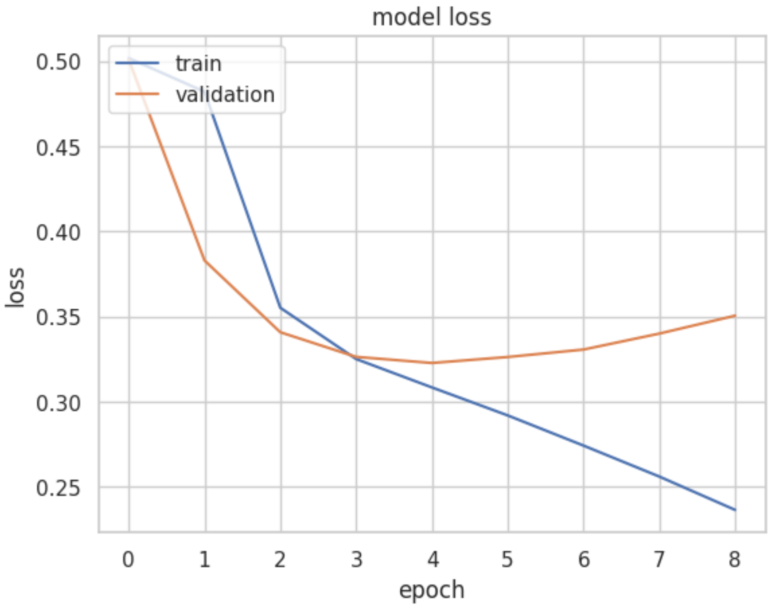

Figure 4.9: Double Stacked LSTM Loss for 100K Data

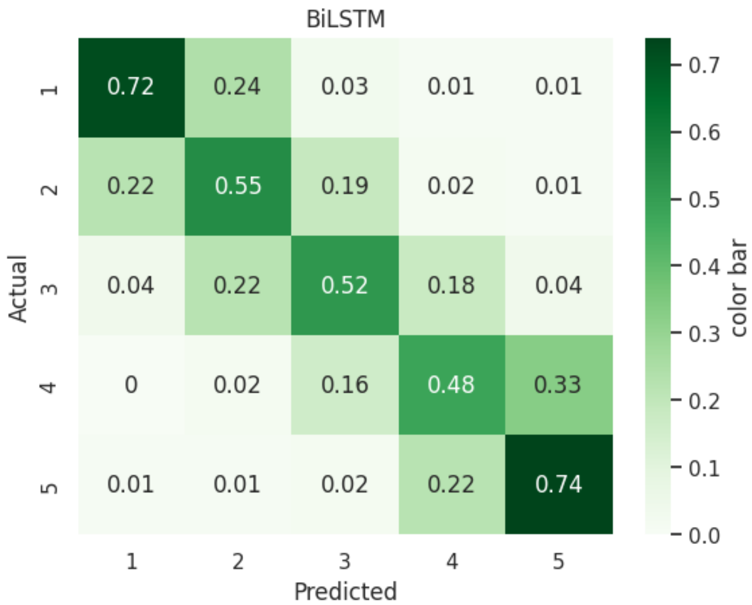

Figure 4.10: Confusion Matrix for Double Layered BiLSTM(100K Data)

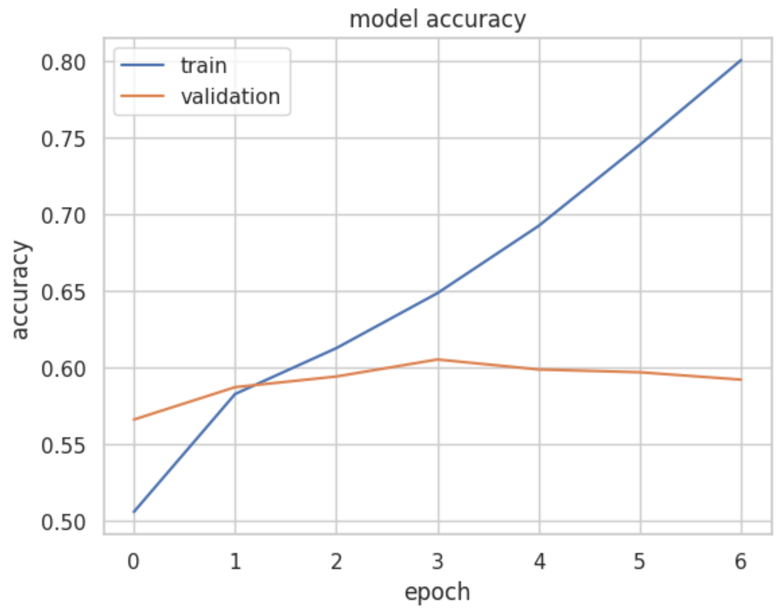

Figure 4.11: Double Layered BiLSTM Accuracy for 100k Data

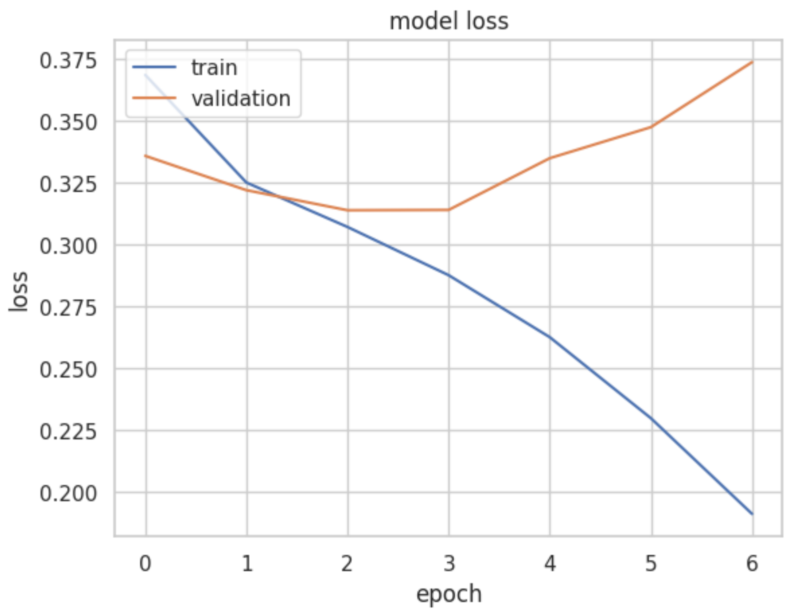

Figure 4.12: Double Layered BiLSTM Loss for 100k Data

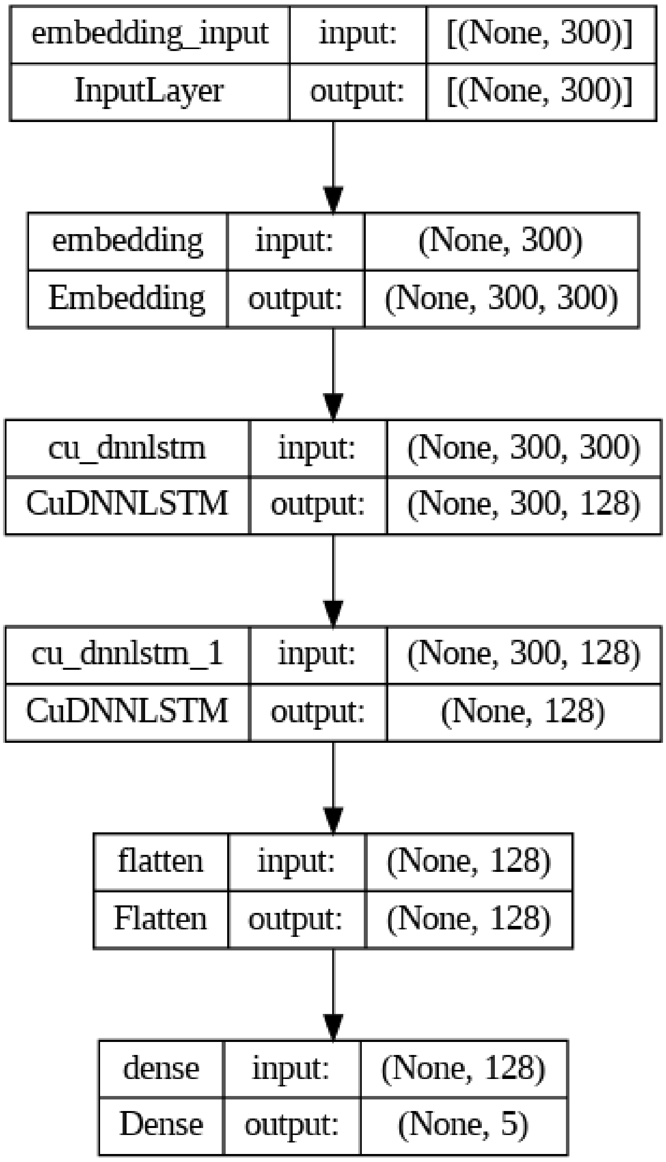

Figure 4.13: Double Layered LSTM Architecture(100K Data)

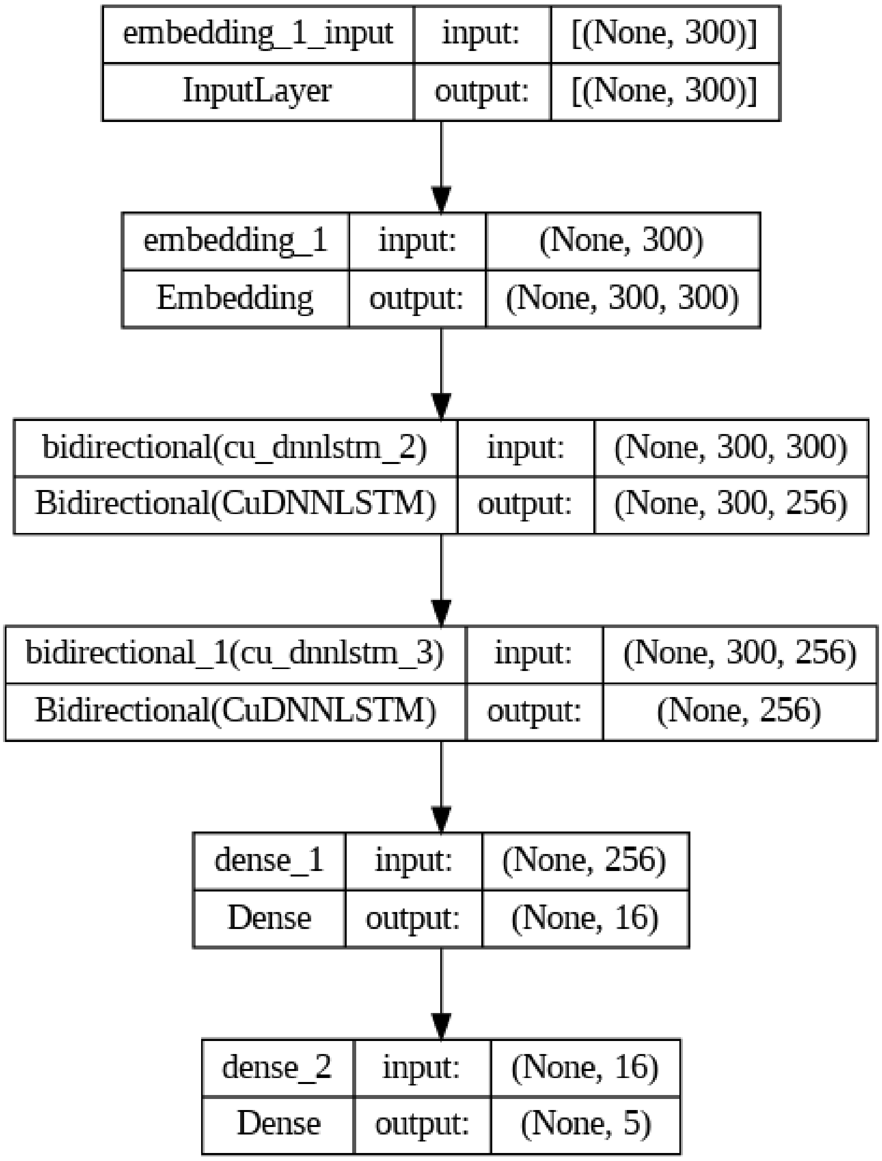

Figure 4.14: Double Layered BiLSTM Architecture(100K Data)

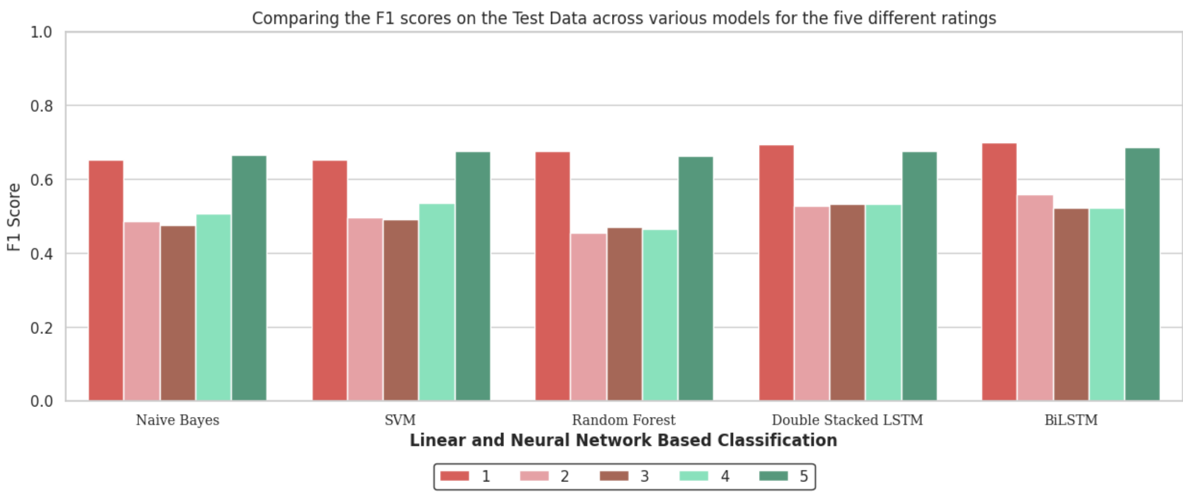

Figure 4.15: F1-score Comparison Across All Models(100K Data)

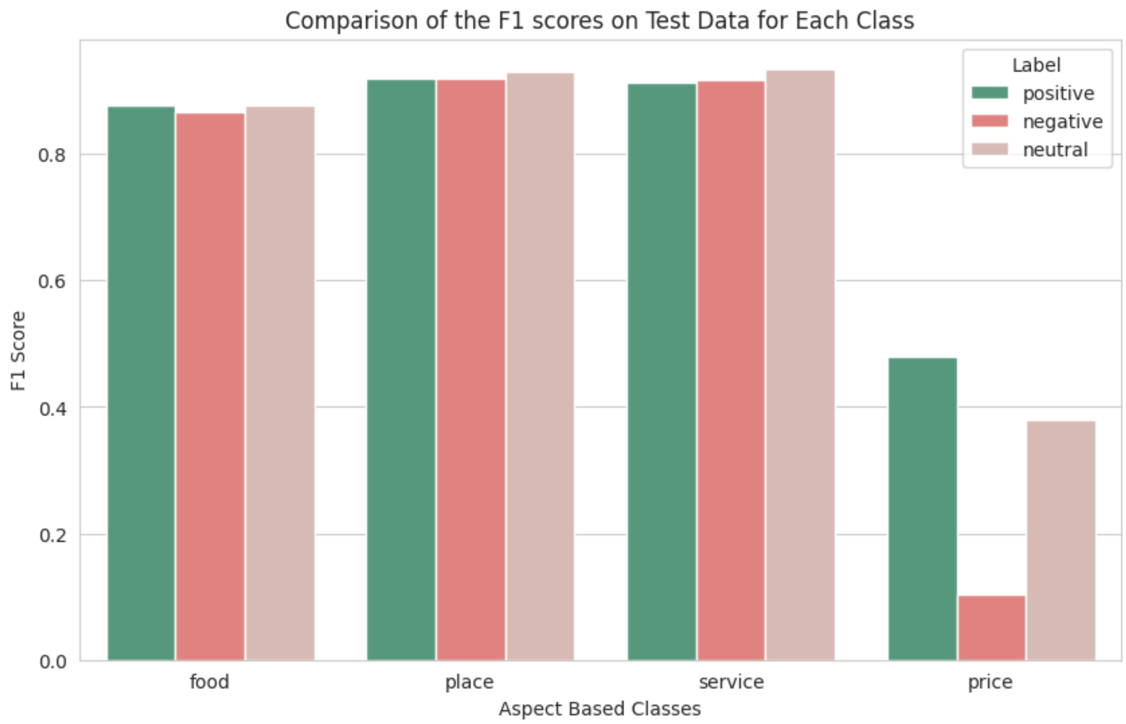

Figure 4.16: Comparison of the F1-scores on Test Data for Each Aspects

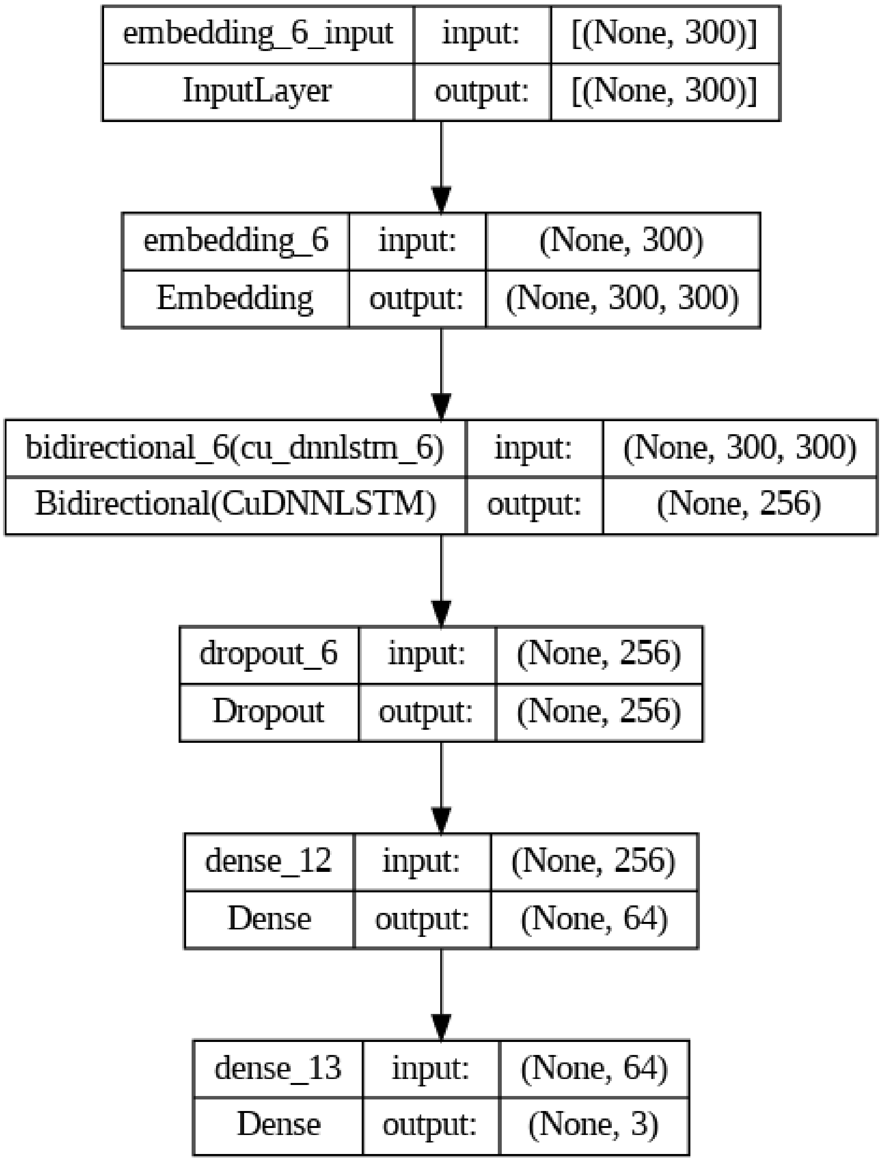

Figure 4.17: BiLSTM Architecture for ABOM

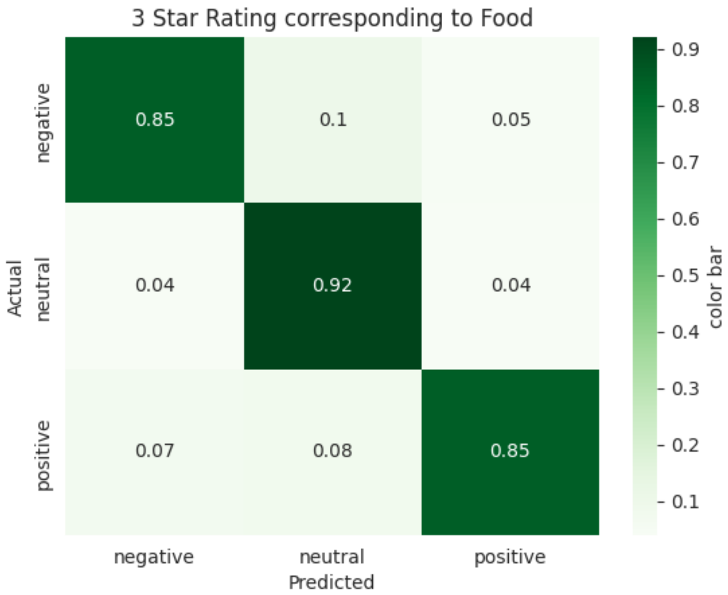

Figure 4.18: Confusion Matrix for BiLSTM(Food)

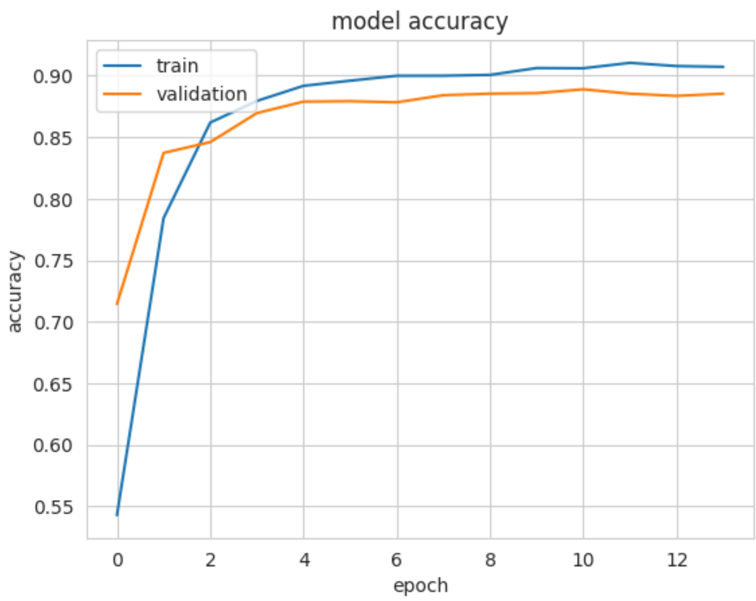

Figure 4.19: BiLSTM Model Accuracy for Food

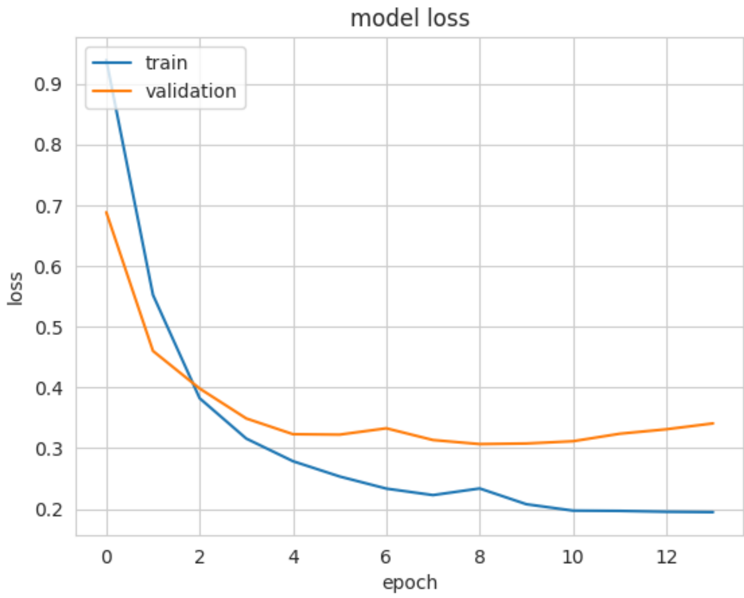

Figure 4.20: BiLSTM Model Loss for Food

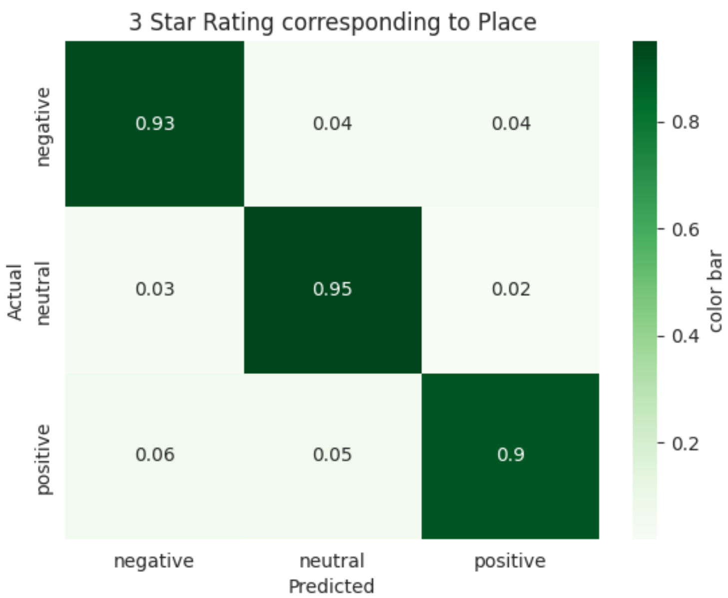

Figure 4.21: Confusion Matrix for BiLSTM(Place)

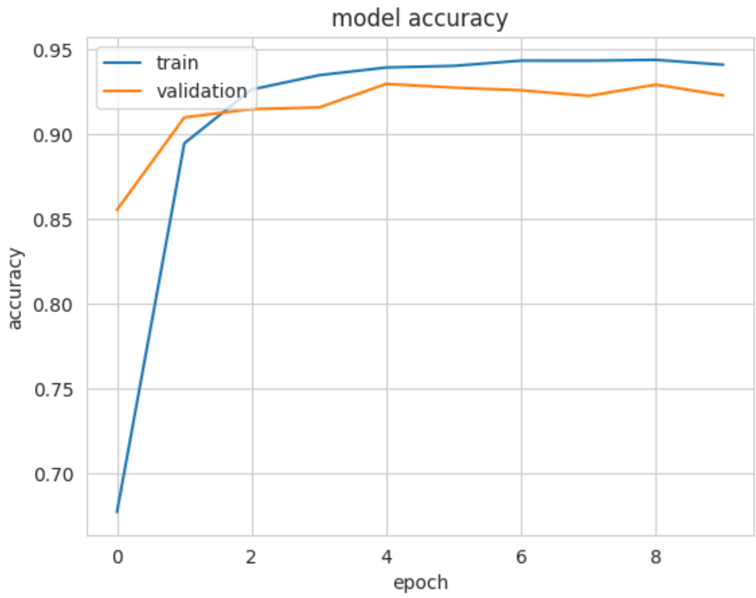

Figure 4.22: BiLSTM Model Accuracy for Place

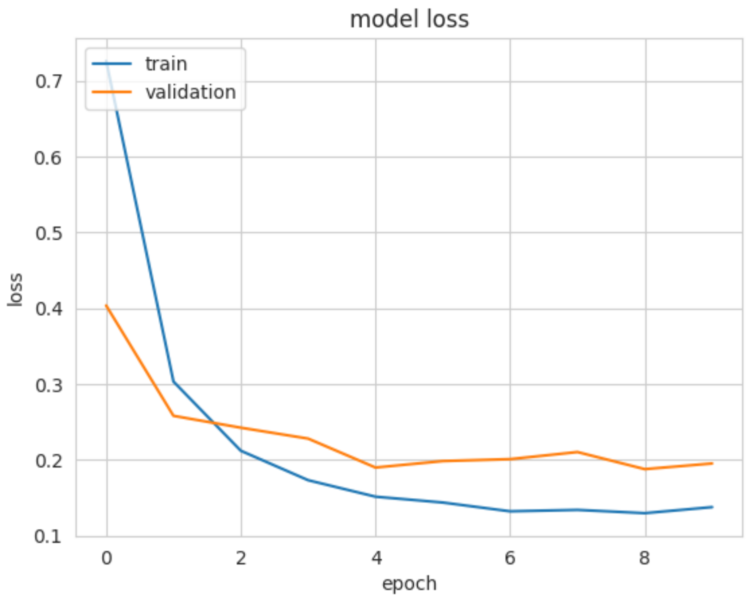

Figure 4.23: BiLSTM Model Loss for Place

Figure 4.24: Confusion Matrix for BiLSTM(Service)

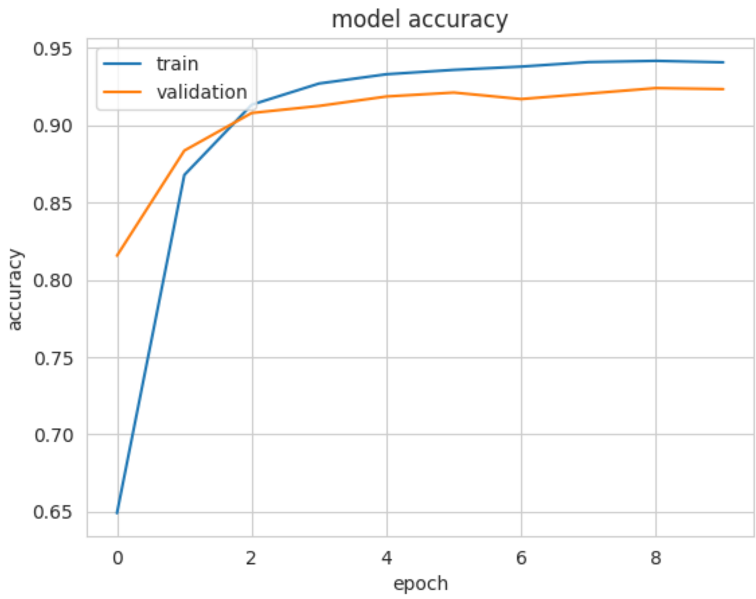

Figure 4.25: BiLSTM Model Accuracy for Service

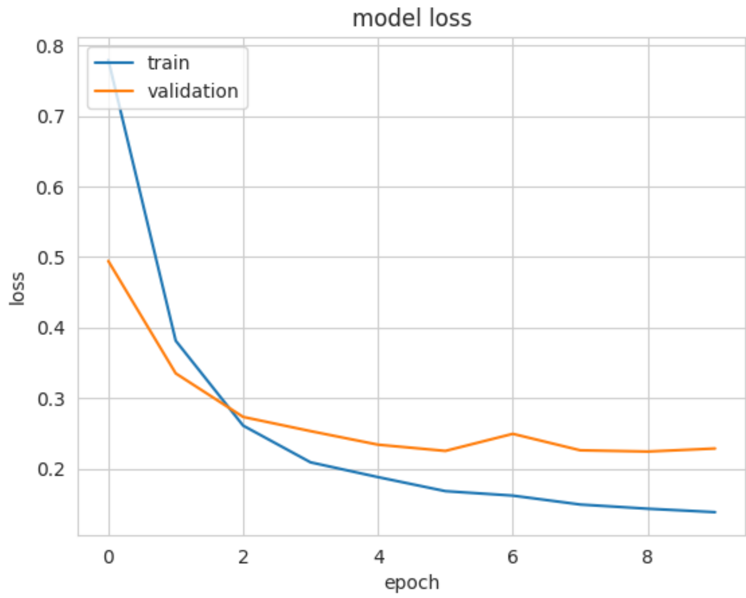

Figure 4.26: BiLSTM Model Loss for Service

Figure 4.27: Confusion Matrix for BiLSTM(Price)

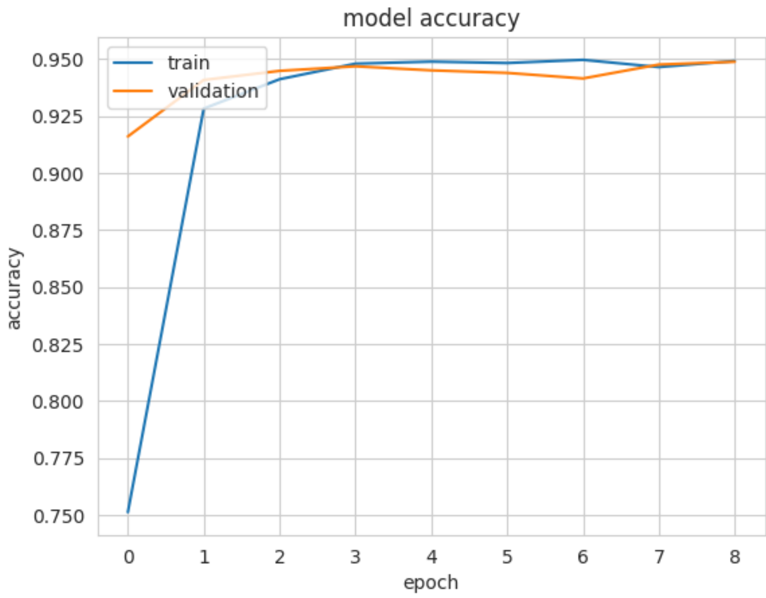

Figure 4.28: BiLSTM Model Accuracy for Price

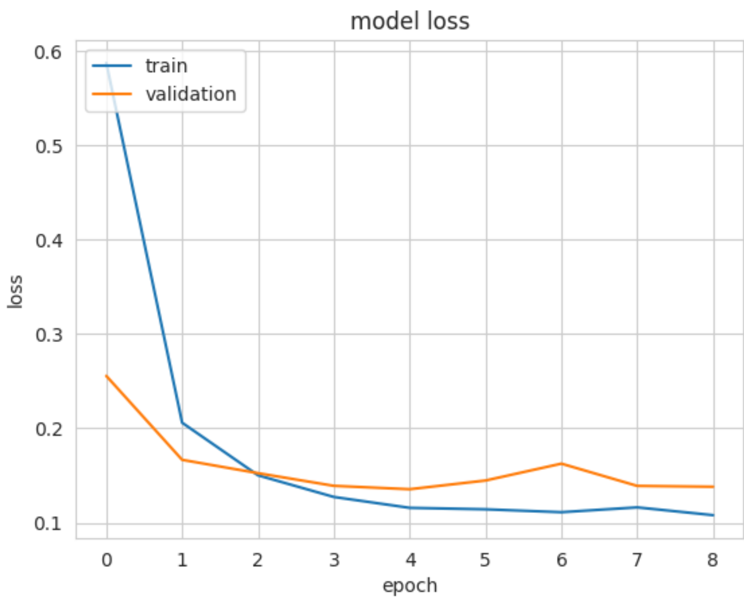

Figure 4.29: BiLSTM Model Loss for Price

Chapter 5

Challenges, Limitations and Future Works

5.1 Challenges

5.1.1 Computational Resources

The initial challenge faced was the computational limitations when using Support Vector Machines (SVM) on a large dataset. The SVM model requires substantial computational power and time, which may not be accessible to all researchers. This highlights the need for efficient algorithms that can handle large datasets without extensive resource requirements.

5.1.2 Class Imbalance

The original dataset exhibited a significant class imbalance, with the majority of reviews being 5-star ratings. This class imbalance can skew the model's performance and make it biased toward predicting positive sentiments. Addressing class imbalance is a crucial challenge in sentiment analysis.

5.1.3 Aspect Labeling

Labeling aspects like Food, Place, Service, and Price manually can be subjective and time-consuming. This introduces the potential for labeling errors and may limit the quality of the training data, especially when dealing with a vast dataset.

5.1.4 Limited Aspect Data

The aspect-based sentiment analysis model struggled with predicting sentiments for the Price aspect, likely due to a lack of sufficient labeled data. Collecting more annotated data for this specific aspect is a challenge, as it requires manual effort and domain expertise.

5.1.5 Data Augmentation

While data augmentation improved the performance of the BiLSTM model for most aspects, it did not effectively address the challenges associated with the Price aspect. This highlights the need for specialized techniques to augment data for underrepresented aspects.

5.2 Limitations

5.2.1 Model Performance

Despite various attempts and techniques, achieving high accuracy and F1 scores across all aspects and rating categories remains a challenge. The limitations of the models suggest that further optimization or the exploration of alternative algorithms may be necessary.

5.2.2 Subjectivity

The aspect labeling process relies on human judgment and interpretation of reviews, which can be subjective and prone to errors. Additionally, the aspect categories chosen may not cover all relevant aspects important to customers, limiting the comprehensiveness of the analysis.

5.2.3 Generalizability

The performance of the models may be specific to the dataset used for training and testing. Generalizing the models to other restaurant review datasets or different domains could be challenging and may require additional adaptation.

5.3 Future Work

5.3.1 Efficient Algorithms

Research can focus on developing sentiment analysis algorithms that are computationally efficient, particularly for large datasets. This can make sentiment analysis more accessible to researchers with limited computational resources.

5.3.2 Handling Class Imbalance

Investigate techniques for handling class imbalance, such as oversampling minority classes or using different evaluation metrics that account for class distribution, like weighted F1-score.

5.3.3 Aspect Labeling Automation

Explore the possibility of automating the aspect labeling process using natural language processing techniques. This could reduce labeling errors and expedite the process.

5.3.4 More Data for Underrepresented Aspects

Collect more annotated data specifically for underrepresented aspects like Price. Crowdsourcing or domain-specific forums might be useful for obtaining additional labeled samples.

5.3.5 Model Ensemble

Experiment with model ensembles that combine the strengths of multiple algorithms, potentially mitigating the limitations of individual models and enhancing overall performance.

5.3.6 Fine-tuning for Price Aspect

Focus on fine-tuning the model specifically for the Price aspect, considering alternative model architectures or more advanced techniques tailored to handle this challenging aspect.

5.3.7 Domain Adaptation

Investigate domain adaptation techniques to improve the generalizability of the models to different types of restaurants and cuisines.

Incorporating these suggestions into our research paper will demonstrate a comprehensive understanding of the challenges, limitations, and potential directions for future work in aspect-based opinion mining on restaurant reviews.

Chapter 6

Conclusion

Data mining, computational linguistics, and natural language processing are all parts of opinion mining. Based on the review, it decides if the consumer has a favorable or unfavorable impression. Large amounts of feedback from customers about any restaurant may be handled using the technology, which also offers improved legitimacy. This article focuses on putting into practice an aspect-based opinion miner in consumer domains like restaurant reviews, which automatically identifies significant aspects and opinions of a restaurant through reviewing reviews, then creates a sentiment profile of each restaurant, which can then be used to compare and choose restaurants in a specific location by any customer. There may be room for improvement in this essay. The same, as well as an examination of other varieties, is what our next study seeks to offer. In order to increase the precision of opinion mining, our further study seeks to incorporate the aforementioned as well as the analysis of various sentence types, such as comparative and conditional sentences. For improved outcomes, classifiers can also be used to examine the suggested system.

Bibliography

[1] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of machine Learning research*, vol. 3, no. Jan, pp. 993–1022, 2003.

[2] M. Hoffman, F. Bach, and D. Blei, "Online learning for latent dirichlet allocation," *advances in neural information processing systems*, vol. 23, 2010.

[3] E. Cambria, B. Schuller, B. Liu, H. Wang, and C. Havasi, "Guest editorial special issue on concept-level opinion and sentiment analysis," *IEEE Intelligent Systems Magazine, Special Issue on Concept-Level Opinion and Sentiment Analysis*, vol. 28, pp–15, 2012.

[4] B. Liu, "Sentiment analysis and opinion mining," *Synthesis lectures on human language technologies*, vol. 5, no. 1, pp. 1–167, 2012.

[5] M. Lucas and D. Downey, "Scaling semi-supervised naive Bayes with feature marginals," in *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Sofia, Bulgaria: Association for Computational Linguistics, Aug. 2013, pp. 343–351. [Online]. Available: https://aclanthology.org/P13-1034.

[6] T. Wang and G. Hirst, "Applying a naive Bayes similarity measure to word sense disambiguation," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Baltimore, Maryland: Association for Computational Linguistics, Jun. 2014, pp. 531–537. DOI: 10.3115/v1/P14-2087. [Online]. Available: https://aclanthology.org/P14-2087.

[7] N. Hardeniya, J. Perkins, D. Chopra, N. Joshi, and I. Mathur, *Natural language processing: python and NLTK*. Packt Publishing Ltd, 2016.

[8] S. Poria, E. Cambria, and A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," *Knowledge-Based Systems*, vol. 108, pp. 42–49, 2016.

[9] S. Poria, E. Cambria, and A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network," *Knowledge-Based Systems*, vol. 108, pp. 42–49, 2016.

[10] S. Ruder, P. Ghaffari, and J. G. Breslin, "A hierarchical model of reviews for aspect-based sentiment analysis," *arXiv preprint arXiv:1609.02745*, 2016.

[11] K. Bauman, B. Liu, and A. Tuzhilin, "Aspect based recommendations: Recommending items with the most valuable aspects based on user reviews," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017, pp. 717–725.

[12] M. Pavlinek and V. Podgorelec, "Text classification method based on self-training and lda topic models," *Expert Systems with Applications*, vol. 80, pp. 83–93, 2017.

[13] S. Ray, "Naive bayes classifier explained: Applications and practice problems of naive bayes classifier," *Analytics Vidhya*, 2017.
[14] P. Dwivedi, "Nlp: Extracting the main topics from your dataset using lda in minutes," *Towards data science. Available online: https://towardsdatascience.com/nlp-extracting-the-main-topics-from-your-dataset-using-lda-in-minutes-21486f5aa925.(accessed on 11 March 2021). Search in*, 2018.
[15] M. Peng, Q. Zhang, Y.-g. Jiang, and X.-J. Huang, "Cross-domain sentiment classification with target domain specific information," in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2018, pp. 2505–2513.
[16] M. Abbas, K. A. Memon, A. A. Jamali, S. Memon, and A. Ahmed, "Multinomial naive bayes classification model for sentiment analysis," *IJCSNS Int. J. Comput. Sci. Netw. Secur*, vol. 19, no. 3, p. 62, 2019.
[17] W.-H. Khong, L.-K. Soon, and H.-N. Goh, "Hybrid models for aspects extraction without labelled dataset," in *proceedings of the second workshop on fact extraction and VERification (FEVER)*, 2019, pp. 63–68.
[18] G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, "Sentiment analysis of comment texts based on bilstm," *IEEE Access*, vol. 7, pp. 51 522–51 532, 2019. DOI: 10.1109/ACCESS.2019.2909919.
[19] M. R. Bhat, M. A. Kundroo, T. A. Tarray, and B. Agarwal, "Deep lda: A new way to topic model," *Journal of Information and Optimization Sciences*, vol. 41, no. 3, pp. 823–834, 2020.
[20] G. O. Diaz, X. Zhang, and V. Ng, "Aspect-based sentiment analysis as fine-grained opinion mining," in *Proceedings of the Twelfth Language Resources and Evaluation Conference*, 2020, pp. 6804–6811.
[21] C. Gong, J. Yu, and R. Xia, "Unified feature and instance based domain adaptation for aspect-based sentiment analysis," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020, pp. 7035–7045.
[22] B. D. P. Statista, "Number of social network users worldwide from 2017 to 2025," *URL: https://www. statista. com/statistics/278414/number-of-worldwide-social-network-users/(: 06.02. 2022)*, 2020.
[23] H. Xu, B. Liu, L. Shu, and P. S. Yu, "Dombert: Domain-oriented language model for aspect-based sentiment analysis," *arXiv preprint arXiv:2004.13816*, 2020.
[24] L. Xu, L. Bing, W. Lu, and F. Huang, "Aspect sentiment classification with aspect-specific opinion spans," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020, pp. 3561–3567.
[25] H. Cai, R. Xia, and J. Yu, "Aspect-category-opinion-sentiment quadruple extraction with implicit aspects and opinions," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2021, pp. 340–350.
[26] U. Chauhan and A. Shah, "Topic modeling using latent dirichlet allocation: A survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 7, pp. 1–35, 2021.
[27] S. Sazzed, "A hybrid approach of opinion mining and comparative linguistic analysis of restaurant reviews," in *Proceedings of the International Conference*

on Recent Advances in Natural Language Processing (RANLP 2021), 2021, pp. 1281–1288.
[28] N. Seth, "Part 2: Topic modeling and latent dirichlet allocation (lda) using gensim and sklearn," *Analytics Vidhya*, 2021.
[29] I. L.-R. Botana, V. Bolón-Canedo, B. Guijarro-Berdiñas, and A. Alonso-Betanzos, *Explain and conquer: Personalised text-based reviews to achieve transparency*, 2022. arXiv: 2205.01759 [cs.LG].
[30] S. Parubets, "Digital penetration as a driver of countries'economic growth," *Publishing House "Baltija Publishing"*, 2022.
[31] M. Gerlich, W. Elsayed, and K. Sokolovskiy, "Artificial intelligence as toolset for analysis of public opinion and social interaction in marketing: Identification of micro and nano influencers," *Frontiers in Communication*, vol. 8, p. 1 075 654, 2023.
[32] [Online]. Available: https://datareportal.com/social-media-users%5C%7D,%20journal=%7BDataReportal%7D.
[33] [Online]. Available: https://wisdomml.in/what-are-stopwords-in-nlp-and-why-we-should-remove-them/#:~:text=Stop%20words%20removal%20is%20the,with%20less%20number%20of%20features.
[34] [Online]. Available: https://www.techtarget.com/searchenterpriseai/definition/lemmatization#:~:text=Likewise%2C%20in%20NLP%2C%20lemmatization%20helps,determine%20meaning%20and%20emotional%20tone..
[35] [Online]. Available: https://saturncloud.io/glossary/stemming/#:~:text=What%20is%20stemming%3F,classification%2C%20and%20other%20NLP%20tasks.
[36] [Online]. Available: https://www.analyticsvidhya.com/blog/2022/02/text-data-augmentation-in-natural-language-processing-with-texattack/.
[37] [Online]. Available: https://nlpaug.readthedocs.io/en/latest/.
[38] [Online]. Available: https://towardsdatascience.com/data-augmentation-in-nlp-2801a34dfc28.
[39] [Online]. Available: https://github.com/makcedward/nlpaug.
[40] [Online]. Available: https://www.ibm.com/topics/naive-bayes.
[41] [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.feature_extraction.text.TfidfVectorizer.html.
[42] [Online]. Available: https://scikit-learn.org/stable/modules/naive_bayes.html.
[43] [Online]. Available: https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c.
[44] [Online]. Available: https://medium.com/@cmukesh8688/tf-idf-vectorizer-scikit-learn-dbc0244a911a.
[45] [Online]. Available: https://www.ibm.com/topics/random-forest.
[46] [Online]. Available: https://www.section.io/engineering-education/introduction-to-random-forest-in-machine-learning/.
[47] [Online]. Available: https://towardsdatascience.com/understanding-random-forest-58381e0602d2.
[48] [Online]. Available: https://www.javatpoint.com/machine-learning-random-forest-algorithm.
[49] [Online]. Available: https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/.

[50] [Online]. Available: https://www.datacamp.com/tutorial/random-forests-classifier-python.
[51] [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html.
[52] [Online]. Available: https://www.ibm.com/docs/en/spss-modeler/saas?topic=models-about-svm.
[53] [Online]. Available: https://www.ibm.com/docs/en/spss-modeler/saas?topic=models-how-svm-works.
[54] [Online]. Available: https://www.ibm.com/docs/en/spss-modeler/18.2.2?topic=nodes-support-vector-machine-models.
[55] [Online]. Available: https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/.
[56] [Online]. Available: https://scikit-learn.org/stable/auto_examples/svm/plot_rbf_parameters.html#sphx-glr-auto-examples-svm-plot-rbf-parameters-py.
[57] [Online]. Available: https://scikit-learn.org/stable/auto_examples/svm/plot_svm_nonlinear.html#sphx-glr-auto-examples-svm-plot-svm-nonlinear-py.
[58] [Online]. Available: https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm.
[59] [Online]. Available: https://scikit-learn.org/stable/modules/svm.html.
[60] [Online]. Available: https://www.analyticsvidhya.com/blog/2021/10/support-vector-machinessvm-a-complete-guide-for-beginners/.
[61] [Online]. Available: https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47.
[62] [Online]. Available: https://www.tensorflow.org/api_docs/python/tf/keras/layers/LSTM.
[63] [Online]. Available: https://towardsdatascience.com/light-on-math-ml-intuitive-guide-to-understanding-glove-embeddings-b13b4f19c010.
[64] [Online]. Available: https://nlp.stanford.edu/projects/glove/.
[65] [Online]. Available: https://keras.io/api/layers/recurrent_layers/lstm/.
[66] [Online]. Available: https://pytorch.org/docs/stable/generated/torch.nn.LSTM.html.
[67] [Online]. Available: https://pytext.readthedocs.io/en/stable/configs/pytext.models.representations.stacked_bidirectional_rnn.StackedBidirectionalRNN.Config.html.