BANGLA ASPECT-BASED SENTIMENT ANALYSIS BY SUPERVISED LEARNING BASED ON ASPECT TERM EXTRACTION

Forhad An Naim
Student Id: 012162011

A Thesis

in

The Department

of

Computer Science and Engineering

Presented in Partial Fulfillment of the Requirements

For the Degree of Master of Science in Computer Science and Engineering

United International University

Dhaka, Bangladesh

October 2020

Approval Certificate

This thesis titled **Bangla Aspect Based Sentiment Analysis by Supervised Learning Based on Aspect Term Extraction** submitted by Forhad An Naim, Student ID: 012162011, has been accepted as Satisfactory in fulfillment of the requirement for the degree of Master of Science in Computer Science and Engineering on 15 October, 2020.

Board of Examiners

1.
_______________________________ Supervisor

Dr. Mohammad Nurul Huda
Professor and Director, MSCSE Program
Department of Computer Science & Engineering
United International University
Dhaka 1212, Bangladesh

2.
_______________________________ Head Examiner

Dr. Swakkhar Shatabda
Associate Professor and Undergraduate Program Coordinator
Department of Computer Science & Engineering
United International University
Dhaka 1212, Bangladesh

3.
_______________________________ Examiner-I

Dr. Md. Motaharul Islam
Professor
Department of Computer Science & Engineering
United International University
Dhaka 1212, Bangladesh

4.
_______________________________ Examiner-II

Dr. Md. Saddam Hossain Mukta
Assistant Professor
Department of Computer Science & Engineering
United International University
Dhaka 1212, Bangladesh

5.
_______________________________ Ex-Officio

Dr. Salekul Islam
Professor & Head of the Dept.
Department of Computer Science & Engineering
United International University
Dhaka 1212, Bangladesh

Declaration

This is to certify that the work entitled **Bangla Aspect Based Sentiment Analysis by Supervised Learning Based on Aspect Term Extraction** is the outcome of the research carried out by me under the supervision of Dr. Mohammad Nurul Huda, Professor and Director, MSCSE Program.

Forhad An Naim
012162011
MSCSE Program
Department of Computer Science & Engineering
United International University
Dhaka 1212, Bangladesh

In my capacity as supervisor of the candidate's thesis, I certify that the above statements are true to the best of my knowledge.

Dr. Mohammad Nurul Huda
Professor and Director, MSCSE Program
Department of Computer Science & Engineering
United International University
Dhaka 1212, Bangladesh

Abstract

Sentiment Analysis is the process of retrieving human sentiments from the text. Aspect-based sentiment analysis goes one step ahead than sentiment analysis by automatically assigning sentiments to specific aspect terms. It is a text analysis technique that extracts and separates each aspect term and identifies the sentiment polarity associated with each aspect term. Bangla is the seventh most native language in the world. Almost 230 million people are spoken in Bangla. So, sentiment analysis in the Bangla language is considered as crucial and well-timed research topics. Recently, aspect-based sentiment analysis is progressing because of identifying fine-grained opinion polarity associated with specific aspect terms. But due to lack of proper resources like annotated dataset, corpora, etc. Aspect-based sentiment analysis is a complicated task. In this thesis, we have used two publicly available datasets named cricket and restaurant. To perform aspect category extraction known as one of ASBS's task, we have conducted our experiments based on two recent studies from 2018 and 2020. In those studies, researchers used some conventional supervised learning algorithms as 2018 [SVM, RF, KNN] and 2020 [SVM, RF, KNN, NB, LR]. In our work, after pre-processing the dataset, we applied a new technique named PSPWA (Priority Sentence Part Weight Assignment) on the dataset. After that, we used a few conventional supervised learning algorithms (SVM, KNN, RF, LR, and NB) to demonstrate results. Whereas our dataset is imbalanced, we considered F1-score as a performance measure factor. Then we compared our results with the previous research works on the same dataset. In the cricket dataset, SVM, KNN, LR, NB performed better than two existing works during the experiment and resulting F1 score of 46%, 31%, 43%, 32%. In the restaurant dataset, SVM, LR, NB performed better than two existing works during the experiment and resulting in an F1 score of 49%, 44%, 34%. For both cricket and restaurant dataset, SVM achieved the best F1 score between all algorithms and scored 46% and 49% respectively.

Acknowledgement

By the grace of almighty Allah, we have got the courage, patience, and strength to accomplish the thesis. We would like to express our deep gratitude to our supervisor Dr. Mohammad Nurul Huda for his valuable and constructive suggestions. We would also extend our gratitude to all examiners for their advice and recommendations.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

This chapter describes the aspect-based sentimental analysis. It also narrates the necessity of sentimental analysis and the difference between traditional sentimental analysis and aspect-based sentimental analysis. The objective of the thesis and contribution of the thesis is also described in this chapter.

1.1 Aspect-based Sentiment Analysis – An Introduction

Sentiment analysis extracts meaningful information from text identifying the text as a negative, neutral, or positive opinion. aspect-based sentiment analysis is a text analysis technique that converts text into aspects (service of a restaurant, food of a restaurant). After that, the aspect-based sentiment analysis assigns sentiment for each aspect term (positive, negative, or neutral). To construct the aspect-based sentiment analysis model, researchers can follow a two-step approach. The first step is aspect term extraction and the second step is aspect-specific sentiment extraction. As a first step, the system extracts aspects from a given text. Secondly, the system identifies the sentiment polarity associated with each aspect term.

1.2 Problem Definition

The number of social media users is increasing enormously. At least 3.96 billion people across the world are using social media today, equating to 51 percent of the total global population. On average, they are spending 2 hours and 22 minutes using social media each day [1]. Alike, online shopping is popularizing day by day. An estimation of 1.79 billion people buys digital goods worldwide in 2019. The forecast says the number of online buyers will increase by over 2.14 billion in 2021 [2]. The peoples on the internet are from a different region, cast, culture, and language. As a result, a large number of online content are producing every day in various aspects. According to Forbes, 2.5 quintillion bytes of data are creating each day at the current pace [3]. So, it is quite impossible to read all content and extract the user’s emotion, sentiment from those contents. So, sentiment analysis will be the right option extracting the hidden sentiments or emotions of the content.

Aspect-based sentimental analysis (ABSA) provides fine-grained information on different aspects of the content. In 2020, the world met an unexpected situation due to the Covid-19 pandemic. Peoples have to stay at home to stay safe. But necessity has no bound.

In this situation, peoples purchase their food from the restaurant, grocery from the super shop, medicine from medicine shop through e-Commerce. Telemedicine became popular due to the pandemic. A survey found people are spending on an average 10-30% more online and e-commerce sales increased significantly (grocery 250%, restaurants food 18.88%, hygienic good 300%) [4]. So, a significant number of documents, reviews, comments generate daily through the web, mobile apps, social media, etc. around the world in various languages.

Before purchasing goods or settling any decision on the internet, people are very much aware of checking or verifying reviews, comments, any related documents, social posts, etc. So, sentiment analysis (SA) on online data is a very crucial and well-timed research topic. But sometimes, traditional SA fails to provide correct insight about reviews or comments. An example of an online restaurant review illustrates the scenario.

"The food was good. But the manager how took my phone call was harsh. The delivery man had no idea locating customer addresses using Google Maps. He was goofy."
Traditional SA predicts the overall sentiment of this review as a negative polarity. Form behavioral context overall sentiment score is ok. But people are more concerned about the quality of food rather than the behavior of the restaurant manager or delivery man when ordering food from online. SA fails to serve the goal correctly. What about aspect-based sentiment analysis (ABSA)?

Table 1: Sample example how ABSA model works on senetnce

Sentence	Aspect Category	Category Polarity
The food was good. But the manager how took my phone call was harsh. The delivery man had no idea locating customer address using Google map. He was goofy.	Food	Positive
	Service	Negative

ABSA performs a fine-grained analysis that identifies each aspect of the given documents or reviews and calculates the polarity score associated with each aspect term. This level of analysis is capable of discovering complex opinions from reviews [5].

In the ABSA task, firstly identifies major categories discussed in the sentence. Secondly takes information from the previous step and determines the polarity of each aspect category [6]. After performing ABSA tasks for review, the ABSA model predicts two-aspect categories as food and service. The model determines the sentiment polarity of food is positive & service is negative.

The output of the ABSA will provide a brief idea about the restaurant. It describes the food of the restaurant is good though the service is poor. The restaurant is suitable for good food instead of a better service. Standard SA classifies the overall sentiment of a text but fails to identify some important information such as entity, subject, or, aspect. ABSA model extracts fine-grained information from the review.

1.3 Thesis Objective

The goal of this thesis is to extract or pull out the fine-grained information from Bangla reviews using ABSA. Though, ABSA is a more complex task because of identifying both aspects and sentiments [7]. The study extracts the aspect terms from the given datasets. During the experimental period, a new technique named PSAWA (Priority Sentence Part Weight Assignment) applied to the dataset after preprocessing the datasets. After that, we used several conventional supervised learning algorithms such as Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF), K-Nearest Neighbors (KNN), and Naive Bayes (NB) to demonstrate results. Then we compared our results with the previous research works 2018 [5] and 2020 [6] on the same dataset. The purpose of the study is to improve the F1 score than the previous research works.

1.4 Thesis Contributions

- The study aims to describe the process of constructing an ABSA model based on aspect term extraction using conventional machine learning algorithms.
- In this thesis, a new technique named PSPWA (Priority Sentence Part Weight Assignment) has introduced to improve the performance of the proposed ABSA model using conventional machine learning (ML) algorithms.

- The study aims to give a new technique named PSPWA algorithm for future researchers on the field of ABSA to improve the performance of the existing ABSA model or build a new robust ABSA model using a weighted input matrix by the PSPWA algorithm.

1.5 Thesis Organization

The thesis is organized as follows:

Chapter 2: Background and Literature Reviews: This chapter describes the background of aspect-based sentimental analysis. It gives a summary of the previous aspect -based sentiment analysis and mentions the importance and obstacles of this research.

Chapter 3: Proposed Method: This chapter describes the data collection process and data pre-processing process. It also describes the PSPWA technique and supervised learning methods.

Chapter 4: Evaluation and Comparison of Results: This chapter represents the experimental and comparison results to evaluate the proposed methodology.

Chapter 5: Conclusion & Future Works: It concludes the research work and describes future plans.

Chapter 2

Background and Literature Reviews

This chapter describes the background of sentimental analysis and aspect based sentimental analysis. It gives a summary of the previous aspect-based sentiment analysis researches and mentions the importance and obstacles of previous researches.

2.1 Background

Sentiment analysis is the process of extracting human emotion or opinion from any text. Sentiment analysis uses data mining techniques to extract data for analysis to determine the opinion of a document or collection of documents, like blog posts, e-commerce product reviews, news articles, and social media feeds like tweets and status updates. Sentiment analysis is called: also opinion mining which is a computational study of reviews, sentiments, opinions, evaluations, attitudes, subjective, views, emotions, etc., expressed in the text.

Bangla is treating as the seventh most spoken language over the world. Almost 200 million people use Bangla to express their expression and emotion [8]. Generally, people give their reviews after purchasing online products in their mother tongue. So a large number of reviews (comments, hashtags, and social media content) generate every day in Bangla.

Sentiment analysis uses for different purposes. There is no bound. The main use of this technology is on social networks and e-commerce sites. It allows obtaining a summary from people's view behind particular topics and emotional response to a product, brand, event, etc. According to the 2018 Study, 94% of the customer’s trust in online sentiment reviews [9]. Nowadays, the applications of sentiment analysis are wide and powerful. The ability to extract insights from social data is a practice that is being widely adopted by organizations across the world. As an example, Donald Trump used sentiment analysis to gauge public opinion to campaign messages and policy announcements using social media ahead of the 2017 presidential election [10].

Generally, sentiment analysis aims to detect the emotional polarity of the whole document, essay or sentence - if the sentence is positive, negative, or neutral. But, apparently, document or essay expresses different emotions in different aspects. So, traditional sentiment analysis sometimes fails to serve the purposes.

ABSA performs a fine-grained analysis that identifies each aspect of the given documents or reviews and calculates the polarity score associated with each aspect. ABSA extracts each emotion polarity associated with each aspect. So ABSA analysis of document or essay shows better results than traditional SA. The task of the ABSA model is divided into four sub-tasks to identify the polarity of each aspect presents in the document, essay, or sentence [11].

Aspect Term Extraction: ATE identifies all possible aspect terms present in each review or sentence. It can be seen as an information extractor and an aspect can be expressed as a noun, adverb, and adjective [12].

Aspect Term Polarity: ATP determines the polarity of each aspect terms such as positive, negative, neutral or conflict.

Aspect Category Detection: ATD identifies major categories discussed in each sentence or review. Aspect categories are typically coarser than the aspect terms as defined in Aspect Term Extraction [12].

Aspect Category Polarity: ACP takes the information from the previous task ACD and determines the polarity of each aspect category discussed in each sentence such as positive, negative or neutral.

ABSA is one of the critical problems to be addressed by computer. Sometimes determining many entities, features, opinions are very tough for computer machines even impossible while it is very easy for a human. Here, two examples of unmanageable situations for machines about sentiment analysis.

- Dealing with trolls or sarcasm is a very complex task. It is challenging to determine the opposite or real meaning. Sometimes sarcasm shows positive polarity but in reality, it expresses positive polarity. In the process of sentiment analysis, machines get confused to determine the correct polarity.

- Human language is diverse. Grammatical errors, misspellings, slangs are another challenge in the process of extracting accurate human sentiment from text automatically. In particular, the Bangla language is more complex to analyze because of its complicated sentence structure.

Table 2: Sample items of cricket and restaurant dataset

Dataset	Sentence	Aspect Category	Polarity
Cricket	ব্যাক টু ব্যাক হাফসেঞ্চুরি অভিনন্দন সাকিব ভাই	batting	Positive
Restaurant	সেখানে কোন পরিবেশ ছিল না,সেবা ছিল অনুপযুক্ত,এবং খাদ্যটি ছিল গড়ের নিচে।	food	Negative

Machine learning requires a corpus collection or dataset. Machine learning has two types of classifiers as supervised learning and unsupervised learning. Supervised learning requires a labeled dataset but unsupervised learning labeled dataset is optional. Labeling the dataset is a huge task. If we take a huge amount of such texts and feed some machine learning algorithms (like neural networks or SVM or RF) with them, it will learn how to recognize sentiment automatically.

2.2 Literature Reviews

Different conventional learning approach has developed to introduce ABSA in several languages such as English, Hindi, Bangla, etc. Current methods are categorized based on their proposed algorithms and models as language rule, sequential, topic model, deep learning, and more [13].

In early 2012, Bing Liu was first introduced ABSA specifically in his research works. He included a chapter on ABSA specifically in his thesis. It describes ABSA and its methods. It also expresses many sub problems that arise from the main problem such as dealing with explicit and implicit aspects. The study narrates aspects that are expressed by nouns or noun phrases are explicit aspects and all other expressions that indicate aspects are called implicit aspects [14]

In SemiVal-2014, the task of the ABSA model is divided into four sub-tasks to identify the aspects of given target entities and the sentiments expressed for each aspect. The study defines sub-tasks as Aspect Term Extraction, Aspect Term Polarity, Aspect Category Detection, and Aspect category polarity [11].

In 2018, a study of ABSA conducted on Amazon products. The authors built an ABSA model based on feature level sentimental analysis. Firstly features are extracted from the sentences. After that, the sentiment is classified associated with each feature, and individual polarity scores are assigned [15]. Other authors built an ABSA model using Convolutional Neural Network (CNN) and bi-directional LSTM. They used a feed forward neural network for aspect category classification, Conditional Random Field for opinion target expression extraction, and Convolutional Neural Network (CNN) for sentiment polarity classification [16].

In 2019, the authors researched on sentiment analysis tasks and challenges. The authors divided the ABSA task into five subtasks. The study utters named entity extraction, coreference Resolution, domain Dependency, sentiment polarity determination, and subjectivity classification are the challenges of ABSA [17]. The same year, other authors built an ABSA model based on LSTM cell by using SenticNet as an external knowledge to improve accuracy. The model called multi-attentive LSTM (MA-LSTM). The study claims the ABSA model was the first work for using multi-attention LSTM with external knowledge [18]. A research conducted to implement the ABSA model using Lexicon-Based, Machine learning, and Hybrid techniques. The study shows Lexicon-Based presents less performance than Machine Learning, Machine learning requires training labeled data than Lexicon-Based, and Hybrid technique provides highly accurate results [19].

In 2020, researchers of the School of Software Engineering, Xi'an Jiaotong University, conducted a study to find out the issues and challenges of ABSA. The study narrates that the core challenges of ABSA tasks are aspect extraction and sentiment polarity determination that could not be handled through a single solution. Instead, researchers should divide the core challenge into sub-task and sub-challenges to resolve the core challenge [20].

In [21] [22], the authors researched on ABSA model in the Arabic Language. They used some conventional techniques to build the ABSA model such as a combination of lexicon with rule-based models, conditional random fields (CRF), bi-directional LSTM. The studies achieved significant output in the context of the Arabic ABSA model.

In [23], the authors introduced sentence-level ABSA in the Serbian language. The study provided a base for further research on ABSA for the Serbian language hat is well under-resourced and under-researched in this area.

ABSA model also developed in the Hindi language. In [24], the authors uttered their research works was the first attempt to build the ABSA model in the Indian language. They provided a benchmark platform by creating an annotated dataset that treated as a baseline model for further research. In [25], the authors put special focuses on code mix Indian languages due to the non-availability of language and vocabulary (linguistic and lexical) tools and annotated resources. In [26], the authors built a sentence level ABSA model in Malayam language.

In 2018, the authors first researched to build the ABSA model in Bangla language. They provided two publicly available datasets for further analyzing ABSA due to due to the lack of available datasets for ABSA. The study defines a baseline approach for the subtask of aspect category extraction. They developed the ABSA model based on conventional supervised learning algorithms [5]. The same year, they also developed the ABSA model based on Convolutional Neural Network. The study claims CNN performs better than supervised learning algorithms [27].

In 2019, the authors developed the ABSA model through stacked auto-encoders in Bangla text. The introduced several classification techniques based on stacked auto-encoders and achieved better aspect classification performance concerning the state-of-the-art [28]. The same year, other authors conducted Sentiment Analysis using the ABSA dataset based on Recurrent Neural Network (RNN) with Long-Short-Term-Memory (LSTM). The study achieved a significant result than the previous study [29].

In 2020 [6], the authors developed an ABSA model using two publicly available datasets of 2018 [5]. They used several conventional supervised learning algorithms to build the ABSA model. The study narrates it achieved better F1 scores than the previous study of 2018 [5].

Chapter 3

Proposed Method

This chapter describes the data collection process and data pre-processing process. It also describes the complete PSPWA technique with the following steps as calculate priority length, prepare stem library, find the greatest common part, find the least common part, and weight assignment. It also explains the rules, advantages, and disadvantages of several classification models.

3.1 Dataset Collection

The dataset that has been used in this thesis is publicly available datasets which is the first time implemented ABSA model in Bengali [5] by Md. Atikur Rahman and Emon Kumar Dey. Those two datasets named are cricket dataset and restaurant dataset. Cricket dataset has five aspect categories (batting, bowling, team management, team, and other) and three aspect polarities (positive, negative, and neutral). Restaurant dataset has five aspect categories (food, service, price, ambiance, and anecdotes/miscellaneous) and four aspect polarities (positive, negative, neutral, and conflict). (https://github.com/AtikRahman/Bangla_ABSA_Datasets).

Table 3: Overall summary of both dataset

Dataset	No of sentences	Aspect Category	Polarity
Cricket	2979	Batting Bowling Team Management Team Other	Positive (19%) Negative (72%) Neutral (9%)
Restaurant	2059	Food Service Price Ambiance Miscellaneous	Positive (59%) Negative (23%) Neutral (6%) Conflict (12%)

3.2 Preprocessing

We preprocessed the dataset to make it useful for experiments. First of all, we removed stop words from dataset such as dari('।'), comma(','), colon(';'), etc. We corrected the misspelled sentences manually for a smooth analysis. For tokenization, we used KERAS tokenizer [30]. After tokenizing the dataset we checked each token to determine whether it is unigram text, emoji or emoticon. If it is emoji or emoticon we decoded the emoji or emoticon in the alphabetic form using emoji library [31] for smooth training on the dataset or corpus collection.

3.3 Priority Sentence Part Weight Assignment (PSPWA)

3.3.1 Evolution of PSPWA

In the ABSA model, the aspect terms rely on the NOUNS present in the sentence. On the other hand, the polarity of sentences relies on the adjective and adverb in the sentences [32]. To extract aspect terms from the text, we have used NOUNS from the dataset. Then identified the most similar NOUNS belonging to the given aspect categories. To extract aspect categories, we only considered NOUNS because a NOUN is a word that functions as the name of some specific object or set of objects. We have concluded that NOUNS are the priority word in our proposed method in terms of aspect extraction. And the parts of the sentence are holding NOUNS are priority sentence parts. We improved our previous research work algorithm (PWWA) to build a new technique PSPWA [33].

3.3.2 Discovering Priority Word & Sentence Part

The basic sentence pattern in Bengali is subject + object +verb (SOV) Example: আমি(S) ভাত (O) খাই(V). As our proposed model suggests, we had to identify the NOUNS presents in the text to extract aspect terms. After analyzing the dataset manually, we found that generally, the NOUNS are present either in the subject or object in Bangla sentence. We divided each Bangla sentence into three parts as the subject, the object, and the other. We concluded subject and object as priority sentence parts. So the words present in the subject part and the object part are the priority words. And priority words respective positions are priority positions. We put three times higher weight for each word present in the subject part and the object part than the other part in each sentence. After analyzing the dataset manually, we have seen the subject or object part consist of one first word in a

sentence of length two. In a sentence of length three, the subject or object part consists of two first words. And the other part consists of two or three words in a sentence of length greater than three. So, we considered the words present in subject and object part named as priority sentence parts for the PSPWA algorithm. We proposed an equation to find out the subject and object part based on the length of a sentence. The proposed equation returned an integer value as a priority length. We only considered the words for the PSPWA algorithm between the first and the priority length position in a sentence.
The equation of priority length, PL is below.

$$PL = MIN(LEN\ (SENTENCE), 3) + \frac{LEN(SENTENCE)}{3} - 1 \qquad (1)$$

Table 4 : Example of splitting sentence in two sentence part (Priority part and Other) using priority length (PL)

Sentence	**Priority Part**	**Other**	**Aspect Category**
বাংলাদেশের জাতীয় দলের সবাই ভালো খেলে	বাংলাদেশের জাতীয় দলের সবাই	ভালো খেলে	Team
বাংলাদেশের ব্যাটিং বিপর্যয় ।	বাংলাদেশের ব্যাটিং বিপর্যয়		Batting
খাদ্য ভাল হিসাবে খুব সাশ্রয়ী মূল্যের	খাদ্য ভাল হিসাবে খুব	সাশ্রয়ী মূল্যের	Food

3.3.3 Weight Assignment

The most crucial step to develop the PSPWA method is weight assignment. So, we assigned a priority counter for each unigram word for the subject and the object part. The value of the priority counter increased depending on some criteria. Below we described each criterion gradually. At first, we made a dictionary for each word in the datasets. The key to the dictionary is each word in sentences. The value of the dictionary is a decimal value represented as a word position in the sentence.
Example:

- Sentence, S: বাংলাদেশের জাতীয় দলের সবাই ভালো খেলে
- Dictionary, D: {“বাংলাদেশের”:0, “জাতীয়”:1,” দলের”:2,” সবাই”:3,

”ভালো”:4,” : খেলে ”:5}

Stem/Root Form of Word:

Root words/stems are base forms of words to which affixes (suffix, prefix, etc.) can be attached but expressed same context. Consider the following group of words বাংলাদেশ, বাংলাদেশী, বাংলাদেশের. The appropriate stem for these words is বাংলাদেশ.

Greatest Common Part:

১. বাংলাদেশ আজকে জিতবে ২. আমি বাংলাদেশ সাপোর্ট করি

৩. বাংলাদেশী হিসেবে আমি গর্বিত ৪. বাংলাদেশের প্রান সাকিব আল হাসান

In above four sentences, all stems of "বাংলাদেশ" presents either in priority part. Here, greatest common part is subject.

Least Common Part:

In above four sentences, all stems of "বাংলাদেশ" presents either in priority part. Here, least common part is other.

According to our previous discussion, we considered the first four-word for assigning priority counter. Then we applied the PSPWA assignment on each word of the subject and object for each dataset item. We converted words of the sentences to its stem because words from the same stem are identical in terms of context [34]. Previously we discuss that the unigram word stays on the subject and the object part in the sentences considered as priority words for aspect terms extraction. The initial priority counter of each word is zero. Initially, $PC_{\text{বাংলাদেশের}} = 0$. If any word presents in the subject or object then we increased priority counter by one. For the above sentence, the word 'বাংলাদেশের" presents in the priority position. So, increment $PC_{\text{বাংলাদেশের}}$ by one, $PC_{\text{বাংলাদেশের}} = PC_{\text{বাংলাদেশের}} + 1$.

Then we found out the part (subject, object, other) of the sentence for all stems of that word ("বাংলাদেশের") in each sentence. We determined the greatest common part of sentences in the dataset for all stems of that word. If the sentence part of the word ("বাংলাদেশের") and the greatest common part of sentences matched then we increased the priority counter of that word by one. The sentence's part of the word ("বাংলাদেশের") is subject and the greatest common part is also subject. So, increment $PC_{\text{বাংলাদেশের}}$ by one, $PC_{\text{বাংলাদেশের}} = PC_{\text{বাংলাদেশের}} + 1$.

After that, we determined the least common part of sentences in the dataset for all stems of that word. If the sentence part of the word ("বাংলাদেশের") and the least common part of sentences not matched.

Then we increased the priority counter of that word by one. The sentence's part of the word ("বাংলাদেশের") is subject and the least common part is a other. So, increment $PC_{\text{বাংলাদেশের}}$ by one, $PC_{\text{বাংলাদেশের}} = PC_{\text{বাংলাদেশের}} + 1$.

After completing the previous processes, we check the current priority counter of the word. If the priority counter value is still less than two then we marked it as less important. We assigned value (0.25) for less important words. If the priority counter value of that word is greater or equal two then we marked it as important. Then we assigned weight three times greater than less important word weight. We assigned value (0.75) for important words. Here, the priority counter of $PC_{\text{বাংলাদেশের}}$ is three. So, $W_{\text{বাংলাদেশের}} = 0.75$

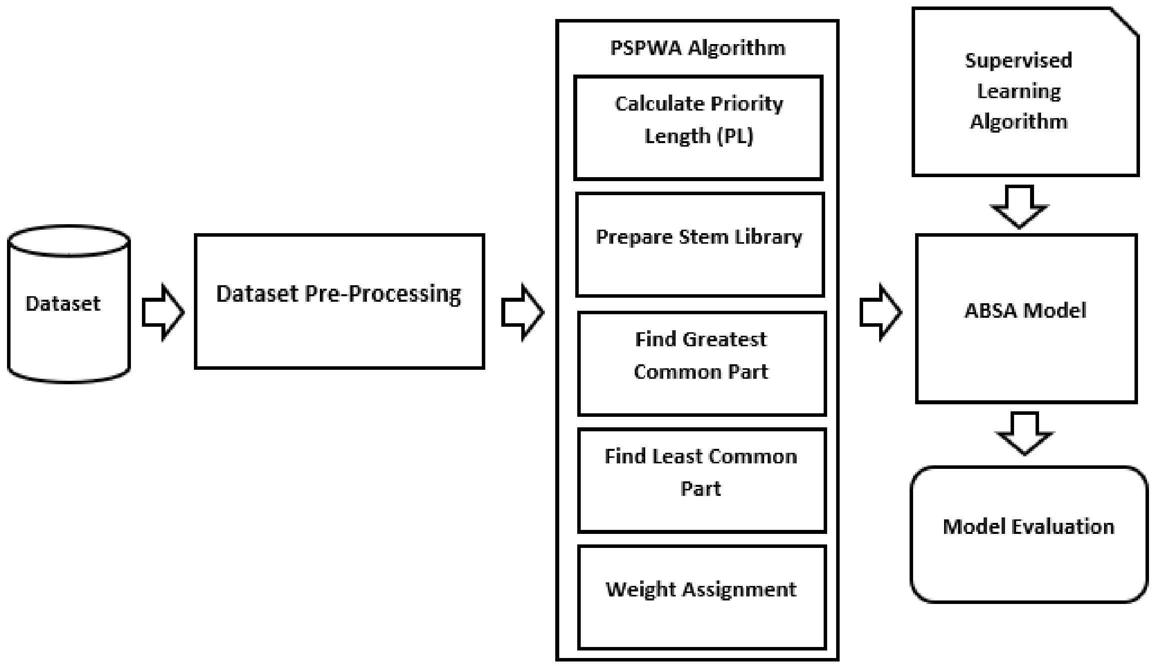

Figure 1: Flow chart of the proposed model

3.3.4 PSPWA Algorithm

Input: A tokenize dataset

Output: A weighted dataset

1. Initial dataset, DS []
2. $Priority_{Part\_OF\_SENTENFE}$ ← {SUBJECT_PART, OBJECT_PART}
3. Make dictionary from dataset DS,
   D {WORD, PART_OF_SENTENCE} ← DICT (DS)
4. REPEAT until last item of dataset, DS
5. FOR each sentence S, in each item of dataset, DS
6. FOR each word W until PL of S
7. Initial priority counter for each PC_W ← 0
8. Prepare Stem dictionary from dictionary D of all stems of W,
   ST {EACH_STEM, PART_OF_SENTENCE}← DICT (W, D)
9. Prepare array of each stem part of sentence, in ST,
   A [] ← FOR each value in dictionary, ST
10. Greatest common part of stems,
    Max_{ST} ← MAX (A)
11. Least common part of stems,
    Min_{ST} ← MIN (A)
12. IF sentence part of W in S ∈ $Priority_{Part\_OF\_SENTENFE}$ then,
    INCREMENT (PC_W)
13. IF sentence part of W in S ==Max_{ST} then,
    INCREMENT (PC_W)
14. IF sentence part of W in S != Min_{ST} then,
    INCREMENT (PC_W)
15. IF $PC_W \geq 2$ then, W_i← 0.75
16. ELSE W_i←0.25
17. Go to step 4 until end of DS

3.4 Methodology

We have used several supervised learning algorithms such as K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Logistic Regression (LR), Random Forest (RF) and Naïve Bayes (NB).

K-Nearest Neighbors:

KNN is widely used supervised learning algorithm which can be used for both classifications as well as regression predictive problems. However, it is mainly used for classification related problems. KNN works by finding the distances between the data point and all the training data points, selecting the specified number training data points (K) closest to the data point, then votes for the most frequent label. Follow below steps to implement the algorithm.

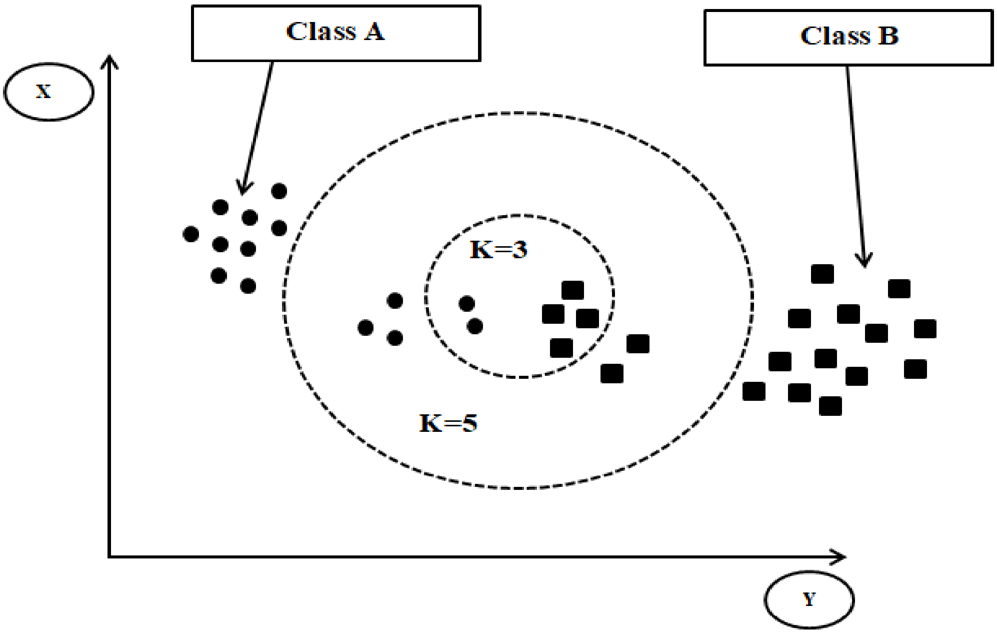

Figure 2: Sample of KNN classification based on differtent K neighbors

KNN Algorithm:

1. Load the data
2. Initialize K, as nearest data points
3. Repeat for each test data
4. Calculate the distances between the test data and the each row of training data using distance metric.
5. Sort the calculated distances in ascending order based on distance values
6. Choose the top k rows from the sorted array
7. Assign the predicted class based on the most frequent class of these rows
8. End

Support Vector Machine:

Support vector machines are power supervised learning algorithm uses for classification, regression and outliner detection. It can solve both linear and non-liner problems. Even with a limited amount of data it works better than other algorithms. The basic idea of SVM is very simple. It creates a line or hyper line which separates the data point into class. Support vectors are the data points that are closest to the hyperplane. Hyperplane is the decision line which is divided set of datasets into different classes. Margin is defined as the space between two lines on the closet data points of different classes.

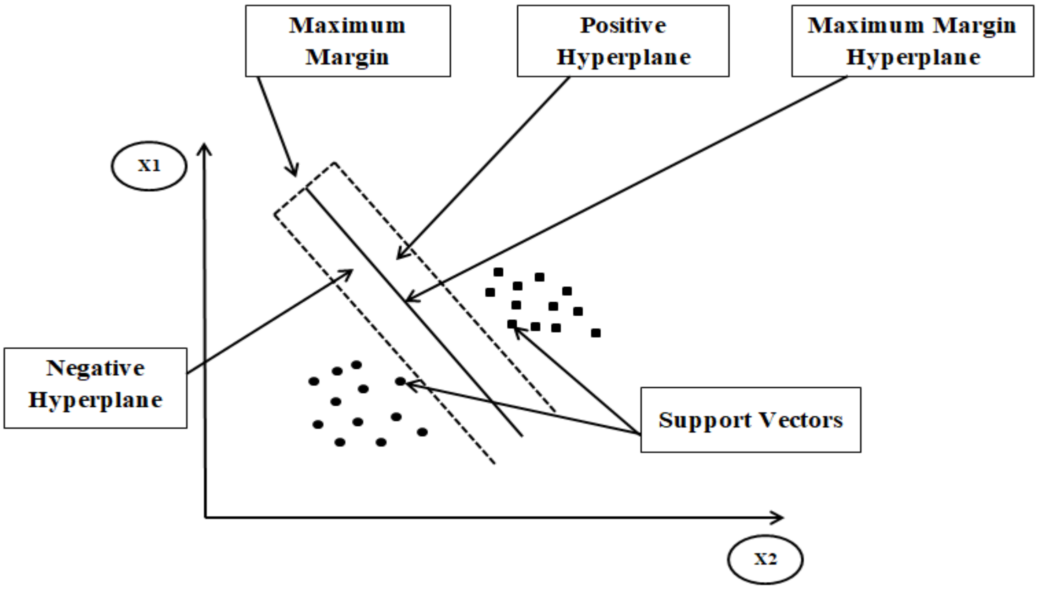

Figure 3: Sample of SVM classification using hyperplane

Advantages of SVM:

1. Accuracy
2. Performs better on limited dataset
3. Effective on datasets with multiple features
4. Different kernel functions

Disadvantages of SVM:

1. Performs worse on large dataset
2. Higher training time on datasets with multiple features

Logistic Regression:

Logistic Regression is a popular supervised learning algorithm used for binary classification. Multinomial Logistic Regression is a form of Logistic Regression which is used to predict a multi class classification. In this case, LR uses softmax function instead of sigmoid function. The softmax function squashes all values to the range [0,1] and the sum of the elements is 1.

$$\text{softmax}\ (x_i) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \tag{2}$$

Cross entropy is a measure of how different 2 probability distributions are to each other. If p and q are discrete we have:

$$\text{H (p,q)} = -\sum_{x} p(x) \log q(x) \tag{3}$$

This function has a range of [0,inf] and is equal to 0 when p=q and infinity when p is very small compared to q or vice versa.

Advantages of LR:

1. LR is easy to implement and efficient to train
2. LR performs better on linearly separable dataset
3. Logistic regression is less prone to over-fitting but it can overfit in high dimensional datasets.

Disadvantages of LR:

1. LR leads to over fit if number of observations are lesser than the number of features
2. LR can only be used to predict discrete functions

Random Forest:

Random Forest is a supervised learning algorithm which is used for both classification and regression problem. Forest is made up of trees and more trees made more robust forest. At first, RF creates decision tree on data points. Then it predicts class label for each data point. Finally, selects the best solution using voting. It is an ensemble method which is better than a single decision tree because it reduces the over-fitting by averaging the result. Follow below steps to implement the algorithm.

RF Algorithm:

1. Load the data
2. Select random samples from the given dataset
3. Construct a decision tree for every sample
4. Predict the class label from every decision tree
5. Performs vote for every prediction result
6. Select the most voted prediction result

7. End

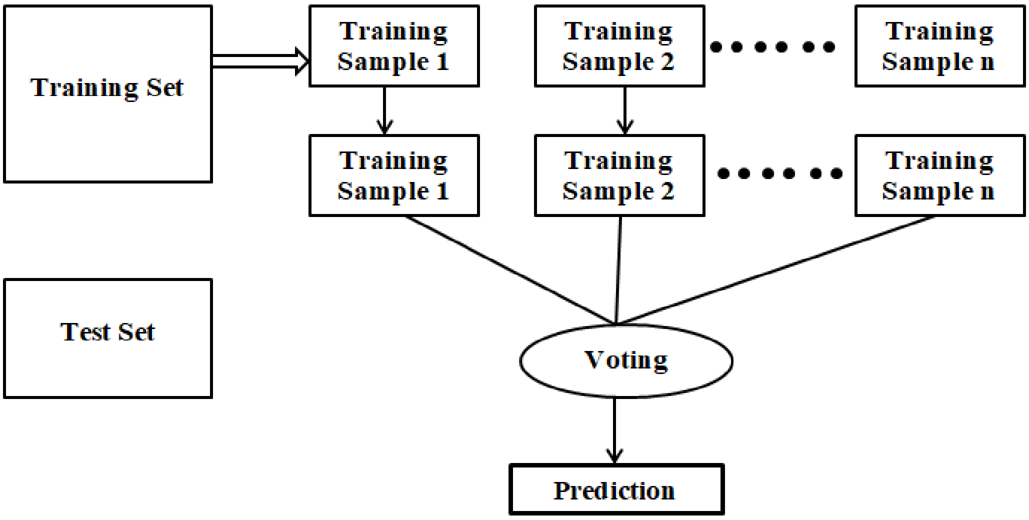

Figure 4: Sample RF working diagram

Naïve Bayes:

Bayes theorem works on conditional probability. Conditional probability is the probability that something will happen, given that something else has already occurred. Below is the formula for calculating the conditional probability.

$$P\,(H|E) = \frac{P\,(H|E) \times P\,(H)}{P\,(E)} \qquad (4)$$

NB is a probabilistic classification algorithm based on Bayes theorem. It predicts membership probabilities for each class such as the probability that given record or data point belongs to a particular class.

Advantages of NB:

1. Predicts class of test data set easily and performs better in multi class prediction
2. Naïve Bayes performs better compare to other models in terms of independent holds assumption
4. Performs better in case of categorical input variables compared to numerical variable(s)

Disadvantages of NB:

1. NB known as bad estimator, so the probability outputs are not taken too seriously
2. Another limitation of NB is the assumption of independent predictors
3. Considers all the features to be unrelated, so it cannot learn the relationship between features

Chapter 4

Evaluation and Comparison of Results

This chapter represents the experimental and comparison results to evaluate the proposed methodology. Table 5 represents the result analysis of the proposed model for both the cricket and the restaurant. Table 6 and 7 represents the comparison between the proposed model and the previous two studies for cricket and restaurant respectively.

4.1 Training & Test Data Set

Earlier, we said that we used a publicly available dataset for analysis. We split the dataset into two parts as a training dataset and test dataset. We used 80% of the dataset as a training dataset and 20% of the dataset as a test dataset

4.2 Experimental Classification Methods

We have used several traditional supervised learning methods for the experiment. Those are Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Logistic Regression (LR), Rain Forest (RF), and Naïve Bayes (NB). In the next part, we will demonstrate a comparison.

4.3 Method Evaluation

Table 5 represents the experimental results of f1-score, precision, and recall of the ABSA classification analysis using some traditional supervised learning algorithms such as SVM, KNN, LR, RF, and NB identifying aspect categories for both datasets. The True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) are four parameters to measure the terms f1-score, precision, and recall.

True Positive (TP) defines that output is positive but the classification model predicts it as positive.

True Negative (TN) defines that output is negative but the classification model predicts it as negative.

False Positive (FP) defines that output is negative and the classification model predicts it as positive.

False Negative (FN) defines that output is positive and the classification model predicts it as negative.

Precision is the number of sentences in the dataset set that is correctly identified by the classification algorithm from the total sentences in the dataset that are clustered by the classification algorithm for a particular class. That is,

$$\text{Precision (P)} = \frac{TP}{TP+FP} \tag{5}$$

The recall is the number of sentences in the dataset that is correctly identified by the classification algorithm from the total sentences in the dataset that are correctly clustered for a particular class. That is,

$$\text{Recall (R)} = \frac{TP}{TP+FN} \tag{6}$$

F1 Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives into account. Intuitively it is not as easy to understand as accuracy, but F1 is usually more useful than accuracy, especially if you have an uneven class distribution.

$$\text{F1 Score (F)} = 2 \times \frac{R \times P}{R+P} \tag{7}$$

Table 5: Experimental results of the proposed method for both dataset

Dataset	**Classifier**	**Precision**	**Recall**	**F1-Score**
Cricket	KNN	0.33	0.38	0.31
	SVM	0.49	0.49	**0.46**
	LR	0.45	0.53	0.43
	NB	0.36	0.48	0.32
	RF	0.41	0.57	0.41
Restaurant	KNN	0.39	0.46	0.39
	SVM	0.49	0.54	**0.49**
	LR	0.44	0.55	0.44
	NB	0.37	0.47	0.34
	RF	0.35	0.50	0.35

The results demonstrate, for both Cricket and Restaurant dataset Support Vector Machine (SVM) gave the highest precision score, which is 49%. But in contrast, for the cricket dataset, Random Forest (RF) gave the highest recall, which is 57%. And for the restaurant dataset, Logistic Regression (LR) gave the highest recall, which is 55%. For the Cricket dataset, the precision score of KNN, LR, RF, NB are 33%, 45%, 36%, and 41%. For the Restaurant dataset, the precision score of KNN, LR, RF, NB are 39%, 44%, 37%, and 35%. On the other hand, For the Cricket dataset, the recall score of KNN, SVM, LR, NB are 38%, 49%, 53%, and 48%. For the Restaurant dataset, the recall score of KNN, SVM, RF, NB are 46%, 54%, 47%, and 50%.

F1-score is more useful than others in terms of measuring algorithm performance when the dataset is imbalanced. As our publicly available dataset is imbalanced, we used F1-score to measure the performance of the proposed model. For both Cricket and Restaurant dataset, Support Vector Machine (SVM) gave the highest F1-score, which is 46% and 49%. For the Cricket dataset, the F1-score of KNN, LR, RF, NB are 31%, 43%, 32%, and 41%. For the Restaurant dataset, the F1-score of KNN, LR, RF, NB are 39%, 44%, 34%, and 35%.

Table 6: Model performance comparison of cricket dataset

	Classifier	**Precision**	**Recall**	**F1 Score**
2018 [5]	SVM	0.71	0.22	0.35
	KNN	0.45	0.21	0.25
	RF	0.60	0.27	**0.37**
2020 [6]	SVM	0.40	0.35	0.35
	KNN	0.27	0.27	0.27
	RF	0.39	0.36	**0.37**
	LR	0.41	0.34	0.34
	NB	0.23	0.27	0.18
Proposed Method	SVM	0.49	0.49	**0.46**
	KNN	0.33	0.38	0.31
	RF	0.35	0.50	0.35
	LR	0.45	0.53	0.43
	NB	0.36	0.48	0.32

Table 6 represents the comparison results of the cricket dataset. F1-score is the measurement term to compare the performance of the methods as the dataset is imbalanced. 2018 [5] used SVM, KNN, RF algorithms in their studies. In that study, SVM achieved the best F1-score, which is 37%. 2020 [2] used SVM, KNN, LR, RF, and NB algorithms in their studies. In that study, RF achieved the best F1-score, which is 37%. In our proposed method, we used the same algorithms as 2020 [6]. In our study, SVM achieved the best F1-score, which is 46%.

Table 7 represents the comparison results of the restaurant dataset. In 2018 [5], KNN achieved the best F1-score, which is 44%. In 2020 [6], LR achieved the best F1-score, which is 43%. In our study, SVM achieved the best F1-score, which is 49%.

Table 7: Model performance comparison of restaurant dataset

	Classifier	**Precision**	**Recall**	**F1 Score**
2018 [5]	SVM	0.77	0.30	0.38
	KNN	0.54	0.34	**0.42**
	RF	0.64	0.26	0.33
2020 [6]	SVM	0.79	0.30	0.39
	KNN	0.39	0.38	0.38
	RF	0.70	0.27	0.35
	LR	0.42	0.43	**0.43**
	NB	0.25	0.26	0.17
Proposed Method	SVM	0.49	0.54	**0.49**
	KNN	0.39	0.46	0.39
	RF	0.35	0.50	0.35
	LR	0.44	0.55	0.44
	NB	0.37	0.47	0.34

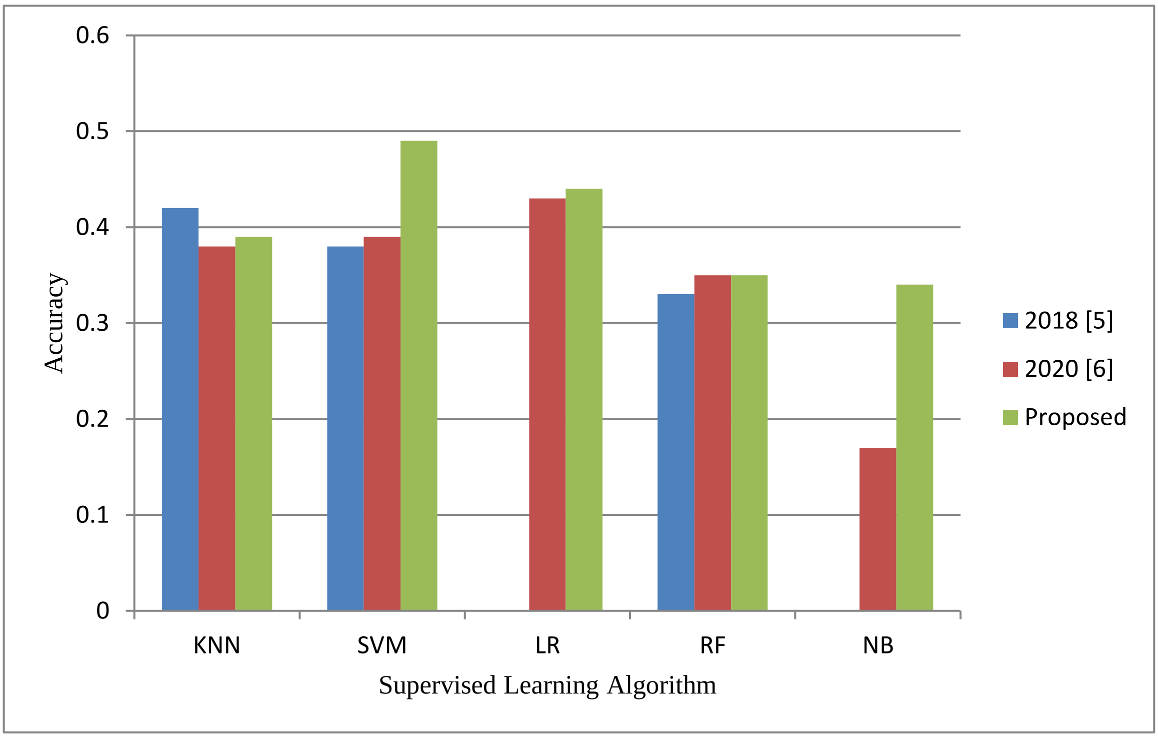

Figure 5 : F1-score comparison chart of cricket dataset

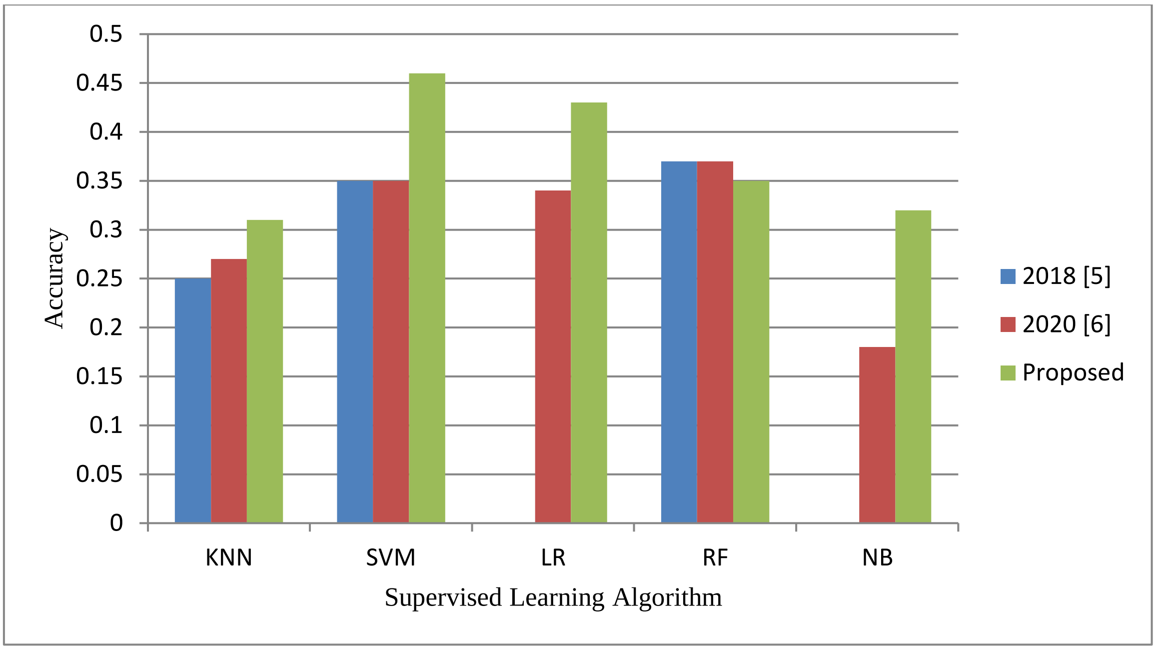

Figure 6: F1-score comparison chart of restaurant dataset

Chapter 5

Conclusion & Future Works

It concludes the research work and describes future plans.

5.1 Conclusion

In this research, we presented a new technique named PSPWA for extracting the aspect terms from the given dataset. The proposed method performs statistical and numerical analysis of the given dataset. The PSPWA technique improved the F1-score when applied to the dataset. For both datasets, our proposed method achieved a better F1-score than 2018 [5] and 2020 [6]. The study aims to give a new technique named PSPWA algorithm for future researchers on the field of ABSA to improve the performance of the existing ABSA model or build a new robust ABSA model using a weighted input matrix by the PSPWA algorithm.

5.2 Future Works

ABSA model divides into two parts as aspect terms extraction and aspect polarity determination. In this research, we did the first part of the ABSA model named aspect terms extraction by several conventional supervised learning. But next, we will try a Convolutional Neural Network (CNN) to compare performance with the current method. In the future, we will also work on identifying sentiment associated with each aspect.

5.3 Limitations

Because of our limitations, we have not constructed a full phase ABSA model. We have constructed a partial ABSA model to identify the aspect terms from the given dataset. A full phase ABSA model identifies the aspect terms and determines the polarity of associated with each aspect term.

References

[1] https://datareportal.com/social-media-users, (Accessed on July 28, 2020)

[2] https://www.statista.com/statistics/251666/number-of-digital-buyers-worldwide/, (Accessed on August 7, 2020)

[3] https://www.forbes.com/sites/bernardmarr/2018/05/21/how-much-data-do-we-ceate-every-day-the-mind-blowing-stats-everyone-should-read/, (Accessed on July 29, 2020)

[4] https://www.bigcommerce.com/blog/covid-19-ecommerce/#product-categories-shifting-during-covid-19, (Accessed on August 7, 2020)

[5] Md. Atikur Rahman and Emon Kumar Dey. Datasets for aspect-based sentiment analysis in bangla dataset. MDPI Journals. doi: 10.3390/data3020015, 2018.

[6] Sabrina Haque, Tasnim Rahman, Asif khan Shakir, Md. Shohel Arman, Khalid Been Badruzzaman Biplob, Farhan Anan Himu and Dipta Das. Aspect Based Sentiment Analysis In Bangla Dataset Based On Aspect Term Extraction. 2nd International Conference on Cyber Security and Computer Science, Febryary 2020.

[7] Mickel Hoang, Oskar Alija Bihorac and Jacobo Rouces. Aspect-Based Sentiment Analysis Using BERT. Nordic Conference on Computational Linguistics (NoDaLiDa), September 2019.

[8] https://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers, , (Accessed on July 29, 2020)

[9] https://www.retailtimes.co.uk/94-of-consumers-check-online-reviews-before-they-buy-anything-feefo-research-finds/, (Accessed on August 8, 2020)

[10] https://www.politico.com/story/2017/10/20/trump-social-media-election-244009, (Accessed on August 8, 2020)

[11] Maria Pontiki, Dimitrios Galanis, John Pavlopoulos, Haris Papageorgiou, Ion Androutsopoulos and Suresh Manandhar. SemEval-2014 Task 4: Aspect Based Sentiment Analysis. In: Proceedings of the 8th International Workshop on Semantic Evaluation, Association for Computational Linguistics, Dublin Ireland, pp 27-35, 2014.

[12] Naveen Kumar Laskari and Suresh Kumar Sanampudi. Aspect Based Sentiment Analysis Survey. IOSR Journal of Computer Engineering (IOSR-JCE),Volume 18, Issue 2, Ver. I (Mar-Apr. 2016), PP 24-28, 2016.

[13] Zohreh Madhoushi, Abdul Razak Hamdan and Suhaila Zainudin. Aspect-Based Sentiment Analysis Methods In Recent Years.Asia-Pacific Journal Of Information Technology And Multimedia, Vol. 7 No. 2 : 79 – 96, June 2019.

[14] Bing Liu. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012.

[15] Sindhu C, Swapneel Niraj Deo, Yash Mukati, Gona Sravanthi and Shubhranshu Malhotra.. Aspect Based Sentiment Analysis Of Amazon Product Reviews.International Journal of Pure and Applied Mathematics, Volume 118 No. 22 2018, 151-157, June 2018.

[16] Alson Cahyadi, Masayu Leylia Khodra. Aspect-Based Sentiment Analysis Using Convolutional Neural Network and Bidirectional Long Short-Term Memory.5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA), August 2018.

[17] Shubham V. Pandey and A. V. Deorankar. A Study of Sentiment Analysis Task and It's Challenges. IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), February 2019.

[18] Win Lei Kay Khin, Nyein Thwet Thwet Aung.Applying Deep Learning Approach to Targeted Aspect-based Sentiment Analysis for Restaurant Domain. International Conference on Advanced Information Technologies (ICAIT), November 2019.

[19] Neha Nandal, Jyoti Pruthi, Amit Choudhary. Aspect Based Sentiment Analysis Approaches with Mining ofReviews: A Comparative Study.International Journal of Recent Technology and Engineering (IJRTE)ISSN: 2277-3878, Volume-7, Issue-6, March 2019

[20] Ambreen Nazir, Yuan Rao, Lianwei Wu and Ling Sun. Issues and Challenges of Aspect-based Sentiment Analysis: A Comprehensive Survey IEEE Transactions on Affective Computing, PP(99):1-1, January 2020.

[21] Saad A. Al-Zboon, Saja Khaled Tawalbeh, Mohammad AL-Smadi and Yaser Jararweh. Using Bidirectional Long Short-Term Memory and Conditional Random Fields for Labeling Arabic Named Entities: A Comparative Study. The Fifth International Conference on Social Networks Analysis, Management and Security(SNAMS-2018), At Valencia, Spain. October 15-18, 2018.

[22] Sufyan Areed, Omar Alqaryouti, Bilal Siyam and Khaled Shaalan. Aspect-Based Sentiment Analysisfor Arabic Government Reviews. Recent Advances in NLP: The Case of Arabic Language, Studies in Computational Intelligence 874, January 2020.

[23] Nikola Nikolic , Olivera Grljevic and Aleksandar Kovacevic. Aspect-based sentiment analysis of reviews in the domain of higher education. The Electronic Library, DOI: 10.1108/EL-06-2019-0140, January 2020.

[24] Md Shad Akhtar, Asif Ekbal, and Pushpak Bhattacharyya. Aspect Based Sentiment Analysis:Category Detection and SentimentClassification for Hindi. Computational Linguistics and Intelligent Text Processing, pp.246-257, January 2018.

[25] Gazi Imtiyaz Ahmad, Jimmy Singla and Nikita Nikita. Review on Sentiment Analysis of Indian Languages with a Special Focus on Code Mixed Indian Languages. International Conference on Automation, Computational and Technology Management (ICACTM), April 2019.

[26] P. K. Thulasi and K Usha. Aspect polarity recognition of movie and product reviews in Malayalam. International Conference on Next Generation Intelligent Systems (ICNGIS), September 2016.

[27] Md. Atikur Rahman and Emon Kumar Dey. Aspect Extraction from Bangla Reviews usingConvolutional Neural Network. 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision& Pattern Recognition (icIVPR), 2018.

[28] Matteo Bodini. Aspect Extraction from Bangla Reviews Through Stacked Auto-Encoders. MDPI Journals. doi:10.3390/data4030121, 2019.

[29] Md. Ferdous Wahid, Md. Jahid Hasan and Md. Shahin Alom. Cricket Sentiment Analysis from Bangla Text Using Recurrent Neural Network with Long Short Term Memory Model. International Conference on Bangla Speech and Language Processing (ICBSLP), 27-28 September 2019.

[30] https://keras.io/preprocessing/text/, (Accessed on July 2, 2020)

[31] https://pypi.org/project/emoji/, (Accessed on June 30, 2020)

[32] Md. Shah Jalal Jamil, Forhad An Naim, Bulbul Ahamed and Mohammad Nurul Huda. Customer Review Analysis by Hybrid Unsupervised Learning Applying Weight on Priority Data. International Joint Conference on Advances in Computational Intelligence, November 2020

[33] Shaika Chowdhury and Wasifa Chowdhury Sentiment Analysis for Bangla Microblog Posts, Brac University, 2013.

[34] Tapashee Tabassum Urmi, A Corpus Based Unsupervised Bangla Word Stemming Using N-Gram Language Model, Shahjalal University of Science of Technology, May 2016.

Appendix A

A comparison table represents analytical results.

We make a confusion matrix based on True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). TP states that output is positive and the classification model predicts it as positive, TN states that output is positive and the classification model predicts it as negative. FP states that output is negative and the classification model predicts it as positive and FN states that output is negative and the classification model predicts it as negative.

SVM - Support Vector Machine

KNN - k-Nearest Neighbors

LR - Logistic Regression

RF – Random Forest

NB – Naïve Bayes