

Aspect-Based Sentiment Analysis for Polarity Estimation of Customer Reviews on Twitter

Ameen Banjar¹, Zohair Ahmed², Ali Daud¹, Rabeeh Ayaz Abbasi³ and Hussain Dawood^{4,*}

¹Department of Information Systems and Technology, College of Computer Science and Engineering, University of Jeddah, Jeddah, 21589, Saudi Arabia

²Air University, Islamabad, 44000, Pakistan

³Department of Computer Science, Islamabad, 44000, Pakistan

⁴Department of Computer and Network Engineering, College of Computer Science and Engineering, University of Jeddah, Jeddah, 21589 Saudi Arabia

*Corresponding Author: Hussain Dawood. Email: hdaoud@uj.edu.sa

Received: 07 September 2020; Accepted: 13 December 2020

Abstract: Most consumers read online reviews written by different users before making purchase decisions, where each opinion expresses some sentiment. Therefore, sentiment analysis is currently a hot topic of research. In particular, aspect-based sentiment analysis concerns the exploration of emotions, opinions and facts that are expressed by people, usually in the form of polarity. It is crucial to consider polarity calculations and not simply categorize reviews as positive, negative, or neutral. Currently, the available lexicon-based method accuracy is affected by limited coverage. Several of the available polarity estimation techniques are too general and may not reflect the aspect/topic in question if reviews contain a wide range of information about different topics. This paper presents a model for the polarity estimation of customer reviews using aspect-based sentiment analysis (ABSA-PER). ABSA-PER has three major phases: data preprocessing, aspect co-occurrence calculation (CAC) and polarity estimation. A multi-domain sentiment dataset, Twitter dataset, and trust pilot forum dataset (developed by us by defined judgement rules) are used to verify ABSA-PER. Experimental outcomes show that ABSA-PER achieves better accuracy, i.e., 85.7% accuracy for aspect extraction and 86.5% accuracy in terms of polarity estimation, than that of the baseline methods.

Keywords: Natural language processing; sentiment analysis; aspect co-occurrence calculation; sentiment polarity; customer reviews; twitter

1 Introduction

The web is a vast source of information and knowledge discovery regarding different entities that can be evaluated by analysis (positive, negative, and neutral sentiments of users). Sentiment analysis is an emerging field and has been a significant task performed on social network datasets using Natural Language processing (NLP) since the mid-2000s [1]. E commerce has provided companies and users with many opportunities to explore online content to make selling and



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

buying decisions, specifically enabling companies to manage their production decisions based on the selling patterns of products and services [2]. There are many methods to analyse sentiments using natural language processing (NLP) techniques, such as employee opinion mining, machine learning, deep learning, and computational linguistics. Several types of datasets have been explored, such as online product reviews, tweets about politics, education and economic affairs, opinions in blogs or discussion forums, and other methods. Several practical applications of sentiment analysis include product selling and buying decisions, the prediction of company sales, and online market development strategies [3]. A polarity scale is employed to analyse the sentiments or user opinions for different domains. Different levels of granularity have been investigated, such as context-aware sliding windows and document-level, sentence-level or aspect-based sentiment analysis [4]. Sentiment analysis focuses on finding the polarity in textual content to uncover the hidden semantics of people's opinions.

Extracting sentiments from a text with opinion mining by using NLP is one of the fields of artificial intelligence. Some researchers have stated that subjectivity analysis is the leading driver of opinion mining [5,6]. However, initially, subjectivity analysis was a sub-task of sentiment analysis and opinion mining [7]. Consider the following: 'According to my friend, it is a greenish shirt, but for me, the colour of the shirt is yellow.' The previous sentence shows that one is convinced that the colour of the shirt is yellow, but to another person, the colour of the shirt is green. The sentence in the previous example is both a subjective and an objective sentence. The sentence 'The shirt is green, a colour that is more eye-catching' is a subjective sentence with sentiment and emotions. Additionally, 'My favourite yellowish shirt is sold out' is an objective sentence with a sentiment [3].

According to the above discussion, the search for sentiments involves four items. These are the sentiment and target objective expressed in a sentence, where h represents the holder and t represents time. These four items are expressed in the form of a quadruple as $(s; g; h; t)$ [8]. Here, the first two items, i.e., the sentiment and target objective, are the most important in sentiment analysis. Two additional terms, i.e., entity and aspect, are also significant in sentiment analysis. An entity can be anything with an independent existence. For example, in mobile reviews, mobile devices can be taken as an entity, while other things such as price and colour are the aspects of mobile devices. Aspect-level sentiment analysis is a vast field; it not only judges the sentiment and emotions that are related to a particular entity but also explores the aspects related to entities [9,10]. There is also a predetermined list of aspects available to use in some techniques. Usually, aspects present in this predetermined list are general-purpose aspects. Discovering aspects from a text is known as the extraction of aspects.

On the other hand, 'I could not attend the call because of a signal problem' is an example of an implicit sentence. Here, someone expects to attend a call; however, this expectation is a negative sentiment. There are many emotions expressed by human beings, such as happiness, anger, and sadness. Languages other than English can apply aspect-based sentiment analysis, such as Arabic and Bangla [11,12].

Document-level sentiment analysis classifies textual documents based on general sentiments of the whole document, e.g., product reviews, comments on tweets, and opinions in blogs or online forums. There is one type of emotion that classifies a document (positive, negative, or neutral). However, sentiment analysis classifies a sentence/phrase, where many sentences in a document can be classified into different types of sentiments, keeping in mind that not all sentences are subjective [2]. When a user provides an opinion about one aspect of an entity, it does not mean that he/she likes or dislikes all other aspects of that entity as well. Consider the following: "The

Canon camera is amazing; it is better than the Samsung camera.” The preceding sentence is an example expressing both positive and negative opinions about two different entities in one sentence. This type of sentiment is confusing and should not be generalized to both products.

Aspect-level sentiment analysis is essential for obtaining fine-grained opinions [3]. It is also called aspect-based opinion mining (ABOM). It is employed to investigate and extract entities and opinions about them. It involves a three-step process: first, extract aspects of a specific entity, then extract opinion words and determine the polarities of the entities, and finally identify the links among the opinions and aspects. If we apply ABOM to the same example, the process output will be “Canon camera,” with the corresponding opinion “amazing,” and then a link between the opinion “amazing” and “Canon camera” instead of “Samsung camera,” based on the correct mapping of relationship. The correct opinion regarding a specific entity is identified and then mapped to the correct entity. Additionally, presenting multiple entities and multiple opinions in one sentence is quite challenging.

The proposed technique is accomplished in three sub-phases: data preprocessing, aspect term extraction, and the polarity estimation of reviews. Data come from the preprocessing step, and then aspect mining or aspect extraction is performed. In this step, all significant aspects are extracted and called features, which reflect user interest. To find the sentiments in any review or opinion, NLP part-of-speech (POS) tagging is performed. POS tagging defines consistent tags and describes how to use a term in a review and the heaviness of the weight estimate polarity of the aspect or feature.

Overall, the contributions made in this paper are as follows:

1. Calculation of the polarity intensity of a review aspect instead of classifying it as either positive, negative, or neutral.
2. Ranking of the identified aspects, which can be achieved using feature ranking methods.
3. A new dataset development, named the trust pilot (forum dataset), which uses defined judgment rules.

2 Related Work

The related work is divided into two subcategories, i.e., aspect detection and sentiment analysis, based on the focus of this work.

2.1 Aspect Detection

The detection of aspects from text data is an essential task in aspect-based sentiment analysis. Aspect detection techniques are categorized as frequency-based methods, syntax-based methods, unsupervised machine learning methods, supervised machine learning methods and hybrid methods [8,10], as shown in Fig. 1.

2.1.1 Frequency-Based Methods

Generally, frequently used words are taken as aspects. These small sets of words are preferably used instead of the rest of the dictionary.

The most well-known approach is the use of frequency-based methods to detect aspects [5]. First, they find all combination of nouns. Nouns sometimes do not have successive values. They must be present in the same text. Then, they are used to judge the screen size of aspects. The primary purpose of using rules is to overcome or remove combinations where nouns occur near each other. One additional purpose is to remove single-word aspects that appear as a part of the

multi-word aspect. The primary reason behind this preprocessing is to find infrequent aspects that are mostly unrelated to any specific aspect.

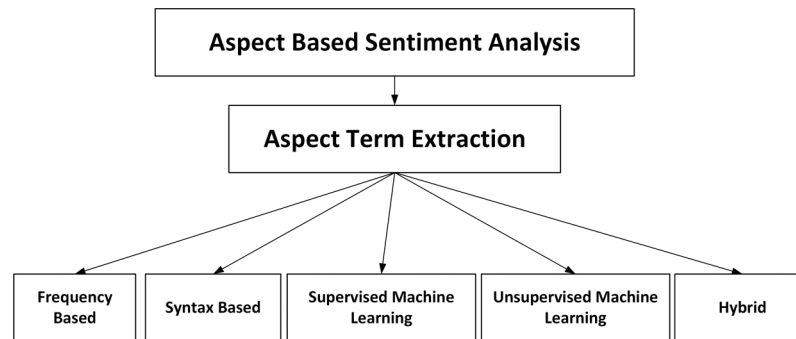


Figure 1: Aspect identification methods

Aspect-based sentiment analysis overcomes this issue in which noun phrases that appear with high frequency are mistakenly considered as aspects [10]. It takes a corpus of 100 million words of spoken and written conversational English. Then, it produces results by comparing the frequency of a prospective aspect with baseline statistics gathered from the corpus. The baseline frequency is very similar to a bigram, which frequently occurs in reviews as an aspect. These non-features are those that slow down the processing of techniques.

2.1.2 *Syntax-Based Methods*

Syntax-based methods work by finding the aspects of syntactical relations. The difference between frequency-based methods and syntax-based methods is that frequency-based methods work by finding the frequencies of aspects, while syntax-based methods find aspects of syntactical relations. The problem of aspect extraction and sentiment lexicon expansion (in its improved methodology) was solved by the proposed algorithm of double propagation [10]. It takes features similar to aspect detection and parallels sentiment words. A speciality of this method is its ability to find more sentiment words using known aspects. These algorithms continuously work for all the different sentiment words or targets. A constructed set of rules based on grammatical relations is used to find the sentiment words. This method needs only a small corpus to accomplish its task.

2.1.3 *Supervised Machine Learning Methods*

Supervised machine learning is a type of machine learning for testing only labelled data and validating the model. The training of any algorithm is necessary for supervised machine learning. A large dataset is usually needed to train these algorithms. Machine learning [13] methods work based on features; thus, features are the main driver of supervised methods. These features are better than part-of-speech features or a bag of words; usually, these features are more generalized.

Aspect detection is considered a labelling problem [11,14], which addresses a linear-chain conditional random field (CRF). Use of the CRF in NLP and to process a whole sentence is common. In the preprocessing of an algorithm, the context of each word is automatically taken and then used to derive multiple features. These features include the actual word, POS tag and relation with other words and sentiment expressions. If the word is a noun, then it is closest to a sentiment expression.

2.1.4 *Unsupervised Machine Learning Methods*

Unsupervised machine learning or cluster analysis is a type of machine learning carried out to find undetected patterns in a dataset without using any labelled data. There is no training required for unsupervised machine learning methods, and groups/clusters of entities are usually formed using distance measures.

Latent Dirichlet allocation (LDA) is one of the most common approaches for aspect detection [8] among the unsupervised machine learning methods. LDA is very similar to probabilistic latent semantic analysis, in which a latent layer is used to model word and document semantic relationships. Both LDA and LSA can be used for the same purpose, but the difference between them is that LDA uses the Dirichlet prior for the topic distribution, while LSA uses a uniform topic distribution. LDA provides the direct link between topic names and aspects or entities using a bag-of-words approach when modelling the topics and documents.

In MG-LDA, there is a dynamic set of local topics and a fixed set of global topics [10]. A document is modelled in such a way that MG-LDA covers a specific number of adjacent sentences to find the local topics. Several local topics vary from document to document. These windows overlap, and by using this overlap, they generate a particular word. This word can be the most sampled from many windows. This phenomenon solves the problem of unigram words or aspects.

2.1.5 *Hybrid Methods*

Whenever two methods or techniques are combined, hybrid methods are formed. Hybrid methods can be categorized into two main types: serial hybridization and parallel hybridization. In serial hybridization, the output of one phase becomes the input for the next phase. In other words, the output of one technique becomes the input of the next technique.

Here, aspect pointwise mutual information is used, and it is provided to the naïve Bayes classifier as its input to yield an output in the form of explicit aspects [15]. More examples of serial hybridization are those in which clusters of noun phrases are produced using a Dice similarity measure. Then, those aspects become the input of support vector machines (SVMs) or any other supervised machine learning method to make a final decision regarding whether they are aspects.

Parallel hybridization deals with two or more methods used to find a complementary set of aspects [11]. Parallel hybridization in the MAXENT classifier finds frequent aspects. Here, sample data and a rule-based method are used. Both use frequency information and syntactic patterns to find less frequent aspects. On the basis of this idea, aspect detection explores the available data, which serve as a back-up for cases in rule-based methods, where sufficient data are not available.

2.2 *Sentiment Analysis*

Sentiment analysis calculates the sentiment score for each review or sentence. It is also known as precise and accurate sentiment analysis. There are many approaches to calculate the sentiment scores for each aspect. Some methods are dictionary-based, corpus-based or phrase-level [16–20]. The details of each approach are given in Fig. 2.

2.2.1 *Dictionary-Based Methods*

Dictionary-based methods are efficient over a specific domain of study. Usually, sentiment words are adjectives. The standard dictionary methods are SENTICNET, SENTIFUL, SENTIWORDNET, and WORDNET [21].

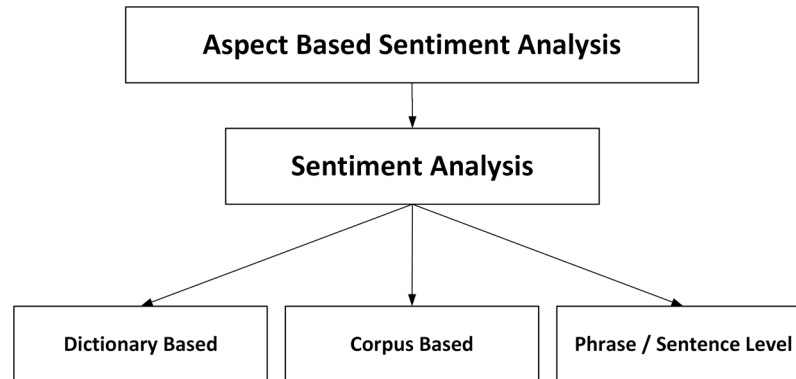


Figure 2: Sentiment analysis methods

A dictionary is generated for each adjective that occurs in a sentence and will be presented in a sentiment class [2]. After five sentiment words, a negation word appears, and then its polarity is changed. In this case, a sentiment class is judged using voting. However, for the same positive and negative words, a different approach can be used. In that case, each adjective sentiment is associated with the closest aspect based on the word distance. Sentiment or opinion words that are attached to the same aspects are more valuable.

2.2.2 *Lexicon (Corpus)-Based Methods*

The lexicon-based method uses an adjective set provided by some online sources, such as Epinions.com. A technique involving the use of a star rating is proposed; here, all adjectives are mapped to a particular star rating [22]. If a situation where a sentiment word does not exist in this set occurs, then the word is taken from the WORDNET synonym graph. A breadth-first search is used to find the two closest adjectives (opinion words) present in the rated list of Epinions.com. After that, a distance weighted nearest neighbour algorithm is used. The purpose of this algorithm is to assign a weighted average to the rating of two neighbours as the estimated rating of the current sentiment word (adjective).

2.2.3 *Phrase-Level/Sentence-Level*

Several approaches discussed in the literature calculate one sentiment score for each sentence, either positive, negative, or neutral. These approaches usually fail when a sentence has multiple sentiment aspects. Consequently, a technique is proposed to overcome this issue [23]. Here, sentences are considered in the form of segments. Each segment is attached to one of the aspects found in the sentence. After that, a sentiment dictionary or lexicon is used to calculate the score of each segment, and an aspect score pair is produced that contains the overall scores or polarity of an aspect within a particular review/comment.

3 Problem Framework

After studying the existing literature regarding aspect extraction and polarity estimation, we can conclude that there exists a need to design a refined technique for polarity estimation in sentiment analysis.

3.1 Problem Formulation

Consider a set of aspects extracted from different customer reviews, for example, opinions presented on reviewing forums or in tweets.

her

.

Aspects of first review: $A_{11}, A_{12}, A_{13}, \dots, A_{1m}$.

Aspects of second review: $A_{21}, A_{22}, A_{23}, \dots, A_{2m}$.

.

.

.

Aspects of n th, review: $A_{n1}, A_{n2}, A_{n3}, \dots, A_{nm}$.

The total number of aspects of all the reviews:

$$R(A) = \sum_{i=1}^n \sum_{j=1}^m (A_{ij}) \quad (1)$$

where A_{ri} is a particular aspect corresponding to review r . To calculate the polarity of each aspect, a model is built for polarity estimation of the normalized relevance of each aspect and the polarity of each aspect, named PER_{OCCR} :

$$PER_{OCCR}(A_r) = \sum_{j=1}^m PER_{OCCR}(A_{rj}) \quad (2)$$

where the norm is the aspect relevance score after normalization. The sum of all scores for the polarity of reviews can be calculated as:

$$PER_{OCCR}(A) = \sum_{i=1}^n \left(\sum_{j=1}^m PER_{OCCR}(A_{ij}) \right) \quad (3)$$

3.2 Problem Statement

Most of the existing techniques are lexicon-based, and their accuracy is affected by limited coverage. Various techniques in practice for polarity estimation about a particular product are too general (either positive, negative, or neutral) and may not indicate the topic clearly, mainly when the category contains much diverse information. Although tweet length is minimal, Twitter can convey a story about different occurrences or events in a precise way. Moreover, tweets sometimes contain a mixture of languages and may not follow the grammatical rules or conventions of English or other languages.

4 Proposed Approach

The proposed approach is composed of three major phases, i.e., data preprocessing, aspect term extraction and the polarity estimation of reviews. The following subsections briefly discuss each phase. Fig. 3 shows a conceptual model of our proposed approach for polarity estimation in customer reviews using aspect-based sentiment analysis.

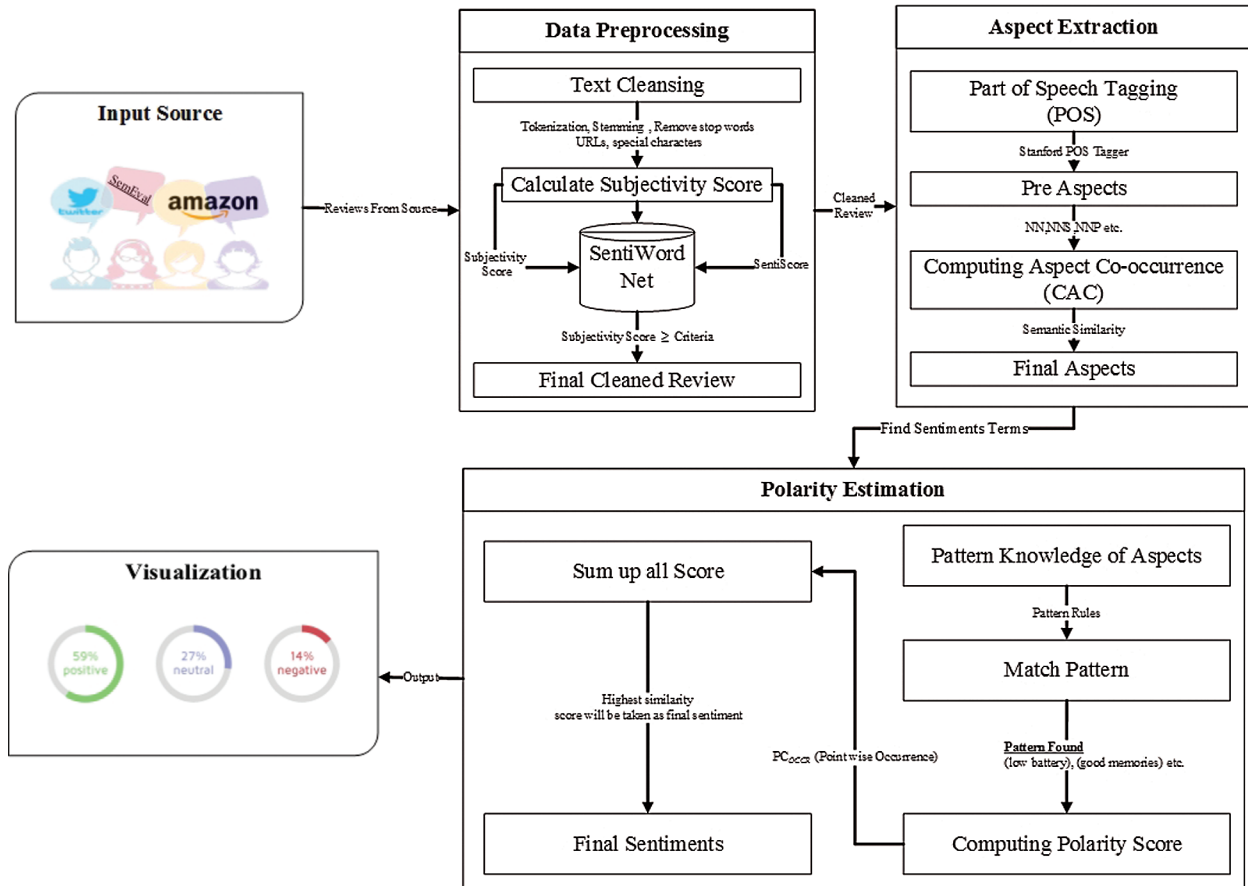


Figure 3: Proposed framework

4.1 Data Preprocessing

Twitter data are obtained using the Twitter API based on the following search criteria: origin name, hashtags, time and “Where on Earth ID” (WoeID) [24]. For text cleansing, stemming is performed, and stop words are eliminated. Reviews are split into sentences and phrases. Tab. 1 shows the data preprocessing of tweets.

In the data preprocessing segment, the paper takes reviews and tweets of any product and generates an output after preprocessing. The output contains a clean form of these tweeted reviews. The following standard rules are applied in data preprocessing:

1. Any full post or review that contains a hyperlink is eliminated because it might be for an announcement, ad, commercial or junk message.
2. All retweets are removed because of redundancy. Posts starting with RT are ignored.
3. Hashtags are used to associate tweets with topics. Mentions refer to usernames on Twitter.
4. Non-English words are removed, for example, invalid characters and digits.

Table 1: Data preprocessing of tweets

S. No	Original tweet	Kept	Processed tweet
1	Leading cinematic presents offline inspecting for iPhone and iPad http://t.co/crtbKLyfTX	No	–
2	@Jackson together since I have a colleague who transfers files using her iPhone	Yes	Together since I have a colleague who transfers files using her iPhone

4.2 Subjectivity Segment

In this research, any statements with sentiments that have some positive or negative opinions are subjective statements, while statements with no sentiments are objective statements. Objective statements are universal truths or facts. For example, “I love the world” is considered in subjectivity, but “The sun rises in the east” must be considered as an objective statement. Therefore, in the subjectivity segment, to achieve this, we need to obtain subjective reviews. After applying preprocessing to any review, we find the subjectivity of those reviews using SENTIWORDNET [2,3]. SENTIWORDNET is a text-based word list that is categorized by the part-of-speech key letters and contains positive, negative, and objective scores of each term. Few words are displayed using SENTIWORDNET. Only positive and negative scores are used to verify that any review is subjective or objective. Tab. 2 provides the positive, negative, and objective scores of some terms, and Fig. 4 provides the term subjectivity and Senti score using SENTIWORDNET.

Table 2: Terms along with their scores on SENTIWORDNET

Word	Positive	Negative	Objective
Glad	0.984	0.00	0.236
Low	0.246	0.86	0.236
Bad	0.00	0.536	0.486
Best	0.86	0.00	0.36
Horrible	0.00	0.486	0.736

4.3 Calculating the Subjectivity of Posts

Let pos_i be the positive score of word w_i from SENTIWORDNET and neg_i be the negative score. Eq. (4) is used to obtain the subjective score from a review or post. The subjectivity score for review r is as follows:

$$R(r) = \frac{\sum_{i=1}^n (pos_i + neg_i)}{n} \quad (4)$$

where n is the total number of words/terms in review R . Basically, this equation sums up the positive and negative weights of each term of review R and the accumulation of all words that occurred in the review or post. To calculate the normalized subjectivity score, it is required to

divide by n . If $pos_i = neg_i$, then the overall score will be 0, even though the sentence is subjective. For example, consider the following two posts:

R1: For those who go once and do not enjoy it, all I can say is that they just do not get it.

R2: The staff was so horrible to us.

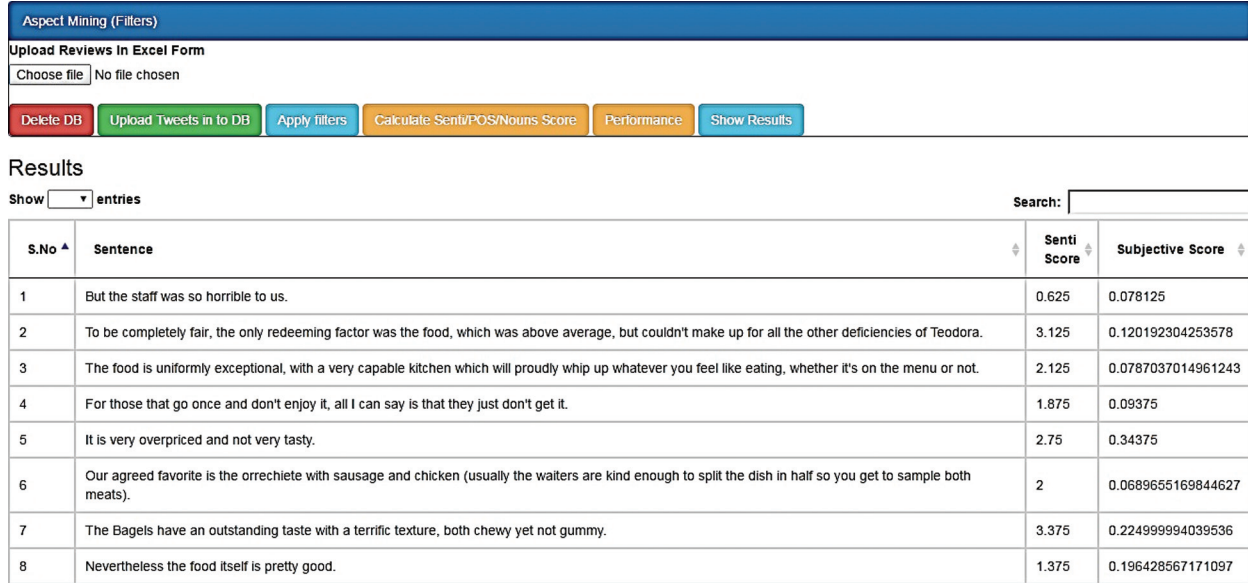


Figure 4: Subjectivity and Senti scores obtained using SENTIWORDNET

In the above example, R2 has more sentiment than R1 regarding any food service, but after adding up all the weights, R1 scores higher than R2 because the number of words is maximum in R1. Therefore, the outcome that R1 reflects stronger sentiment than R2 is false. Dividing by the total number of words is helpful for normalization, and this normalized weight is helpful for further processing. We build a tool to calculate the subjectivity and Senti score using SENTIWORDNET. Algorithm 1 provides the details of the calculation of the subjectivity score of reviews.

4.4 Aspect Term Extraction

Aspect identification from text is known as aspect extraction. It can also be considered an information extraction task, but in the case of sentiment analysis, some definite characteristics of the current problem can help in extraction as well. One of the most critical pieces of information is opinions or emotions. It is essential to judge the sentiment and opinion expressions and their targets related to a sentence. In many situations, opinions or sentiment expressions can be considered in two ways. For example, in the sentence “I love my country,” love is a sentiment word and country is an aspect. Some of the critical approaches include extraction based on frequent nouns and noun phrases, extraction by manipulating opinion and objective associations, or extraction by using supervised machine learning techniques. This research proposes a method named aspect co-occurrence calculation (CAC) to determine the refined aspects of posts with the help of the POS.

Algorithm 1: Calculating the subjectivity score of reviews**Input:** Preprocessed reviews, R from the preprocessing module**Output:** Subjective posts, SP.**Start****IF** (WR > 3) **THEN** **WHILE** (WR.startswith! = "RT") **DO** **WHILE** (WR.contains(! = URL)) **DO** **FOR** Reviews R in pre-processed reviews **DO**

Get the pos_score and neg_score of each word, w from SENTIWORDNET

Get the subjective score for R

IF (subjective score ≥ 0.05) **THEN**

Add review, R, to collection of subjective posts, SP

ELSE

Discard

ELSE

Discard

End**4.4.1 Part-of-Speech (POS) Tagging**

POS tagging is performed to identify how a term is used in a review. The Stanford POS tagger is one of the best POS taggers for assigning POS tag metadata in reviews or tweets [23], examples of which are provided in Fig. 5.

Tag	Description	Example	Tag	Description	Example
CC	Coordin. Conjunction	<i>and, but, or</i>	SYM	Symbol	<i>+, %, &</i>
CD	Cardinal number	<i>one, two, three</i>	TO	"to"	<i>to</i>
DT	Determiner	<i>a, the</i>	UH	Interjection	<i>ah, oops</i>
EX	Existential 'there'	<i>there</i>	VB	Verb, base form	<i>eat</i>
FW	Foreign word	<i>mea culpa</i>	VBD	Verb, past tense	<i>ate</i>
IN	Preposition/sub-conj	<i>of, in, by</i>	VBG	Verb, gerund	<i>eating</i>
JJ	Adjective	<i>yellow</i>	VBN	Verb, past participle	<i>eaten</i>
JJR	Adj., comparative	<i>bigger</i>	VBP	Verb, non-3sg pres	<i>eat</i>
JJS	Adj., superlative	<i>wildest</i>	VBZ	Verb, 3sg pres	<i>eats</i>
LS	List item marker	<i>1, 2, One</i>	WDT	Wh-determiner	<i>which, that</i>
MD	Modal	<i>can, should</i>	WP	Wh-pronoun	<i>what, who</i>
NN	Noun, sing. or mass	<i>llama</i>	WP\$	Possessive wh-	<i>whose</i>
NNS	Noun, plural	<i>llamas</i>	WRB	Wh-adverb	<i>how, where</i>
NNP	Proper noun, singular	<i>IBM</i>	\$	Dollar sign	<i>\$</i>
NNPS	Proper noun, plural	<i>Carolinas</i>	#	Pound sign	<i>#</i>
PDT	Predeterminer	<i>all, both</i>	"	Left quote	<i>(' or ")</i>
POS	Possessive ending	<i>'s</i>	"	Right quote	<i>(' or ")</i>
PP	Personal pronoun	<i>I, you, he</i>	(Left parenthesis	<i>([, ({ , <</i>
PP\$	Possessive pronoun	<i>your, one's</i>)	Right parenthesis	<i>(] ,) , ></i>
RB	Adverb	<i>quickly, never</i>	,	Comma	<i>,</i>
RBR	Adverb, comparative	<i>faster</i>	.	Sentence-final punc	<i>(. ! ?)</i>
RBS	Adverb, superlative	<i>fastest</i>	:	Mid-sentence punc	<i>(: ; ... --)</i>
RP	Particle	<i>up, off</i>			

Figure 5: Descriptive part of speech [25]

A POS tagger finds labels or words in a sentence along with the grammatical relationships present in a sentence. For example, each word in a sentence can be tagged as a verb (VB), noun (NN), noun phrase (NNP), proper noun (NNP), adjective (JJ), or article (DT).

4.4.2 Computing the Aspect Co-Occurrence (CAC)

The aspect co-occurrence is used to compute the semantic similarity between words/terms used in a review sentence and those that are selected by the user to represent an aspect [25]. When the value of $CAC(x,y)$ is relatively small, words can be considered to have a more significant semantic similarity.

As an example, in this review, we detected two candidate aspects; here, we compute the semantic similarity between them. “Mobile” received 130 hits; “Battery” received 260 hits; and the accumulative for “Mobile” and “Battery” together received 80 hits in the reviews. According to the formula $CAC(x,y) = 0$, x is the same as y . Therefore, $f(x) = 130$, $f(y) = 260$, $(f(x),f(y)) = 80$, and $N = 190$ when the calculation value is 1.63. CAC is defined in Eq. (5).

$$CAC(t_1, t_2) = \sum_{i=1}^n \frac{f(t1) + f(t2) - f(t_1, t_2)}{N} \tag{5}$$

Here, N is the total number of words/terms in reviews, and x and y represent words used to compute the similarity. In addition, $f(x)$ and $f(y)$ are the numbers of occurrences containing x and y , respectively.

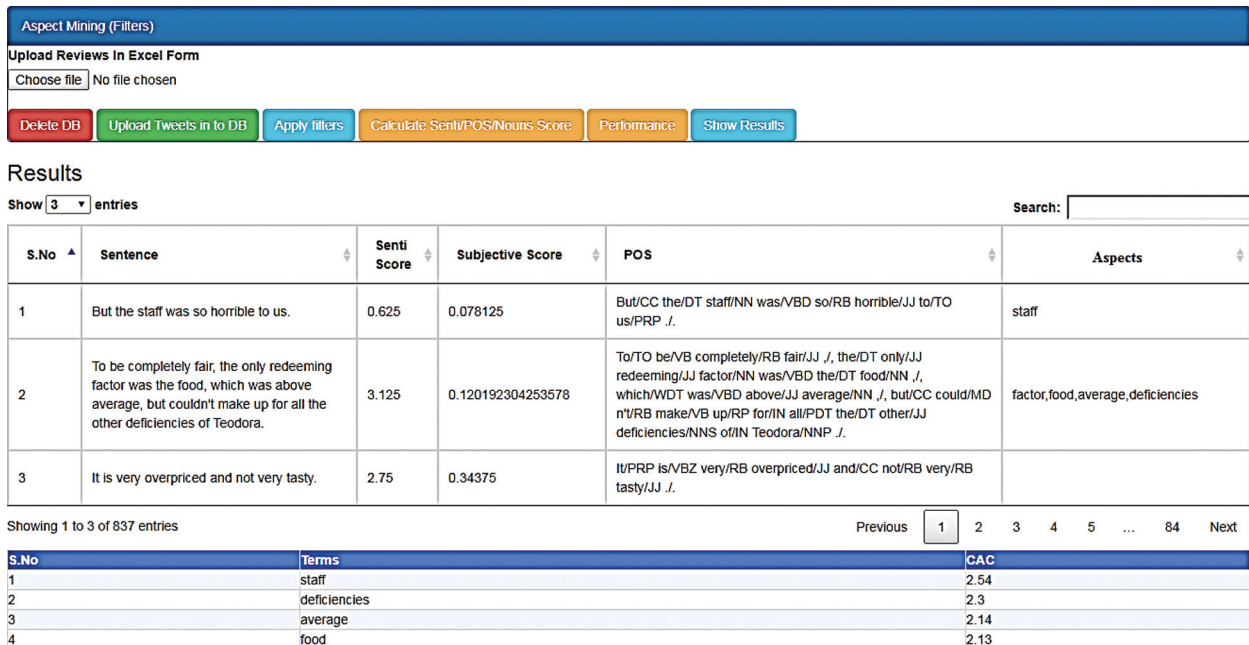


Figure 6: Aspect extraction using the CAC

This association measure can be utilized to identify the most accurate aspects of a particular category. Finding an association between two different terms using the CAC does not require any background knowledge or analysis of the problem domain. Instead, it automatically analyses all

features through occurrence or frequency in the dataset. In this phase, an aspect matrix is created that provides each aspect along with the corresponding class/category.

In the aspect matrix, all aspects of the relevant class are represented as a vector consisting of the co-occurrence (number of occurrences of a feature with its class) along with its relevance score. Aspects with a minimum score greater than a standard threshold are discarded. Our method's second-phase results are shown in Fig. 6. Details of the aspect term extraction are provided step by step in Algorithm 2.

Algorithm 2: Aspects of term extraction (CAC)

Input: Subjective reviews, R, from the subjective module

Output: Aspect terms

Start

FOR Review, R in SR **DO**

 T = R.Split (' ')

FOR T in stop_words **DO**

 Regex.Replace(T, stop_words)

 LOAD_POS_tagger

 = new MaxentTagger(@"\wsj-0-18-bidirectional-nodistsim.tagger")

 R = MaxentTagger.tokenizeText(T)

FOR SR in R **DO**

 POS = edu.stanford.nlp.ling.Sentence.list(taggedSentence)

 pattern = @"\/w*/NN/b|\/w*/NNS/b|\/w*/NNPS/b";

 MatchCollection matches = Regex.Matches(POS, pattern)

FOR matches in pattern **DO**

 PAT = matches

 Calculate_CAC (A, B, AB)

 N = Total no. of aspects

 CAC_AB = No. of occurrences of both terms

 CAC_A = No. of occurrences of term A

 CAC_B = No. of occurrences of term B

 CAC (A, B) = ((CAC_A + CAC_B) - CAC_AB)/N

 FAT = Maximum of CAC(A, B)

End

End

4.5 Polarity Estimation

Polarity calculation and estimation are part of sentiment judgment. The most common types of polarities are positive, negative, and neutral.

4.5.1 Sentiment Detection

Detecting important sentiments and aspects simultaneously from a dataset is an essential task in sentiment analysis. These features are used to find opinion and sentiment words appearing as adjectives/adverbs. Usually, adjectives are taken as opinion words; if adjectives or opinions come close to a feature in the text, then this adjective will be taken as a full opinion word. Fig. 7 shows the extracted phrases.

- The strap is terrible and can only be attached to the camera with much effort.
- After acquiring approximately 700 images, I found the ability of this camera to capture images to be unbelievable.
- It comes with a rechargeable battery that does not seem to last all that long.

Pattern	1 st Term	2 nd Term	3 rd Term
Pattern 1	JJ	NN/NNS	—
Pattern 2	JJ	NN/NNS	NN/NNS
Pattern 3	RB/RBR/RBS	JJ	—
Pattern 4	RB/RBR/RBS	JJ/RB/RBR/RBS	NN/NNS
Pattern 5	RB/RBR/RBS	VBN/VBD	—
Pattern 6	RB/RBR/RBS	RB/RBR/RBS	JJ
Pattern 7	VBN/VBD	NN/NNS	—
Pattern 8	VBN/VBD	RB/RBR/RBS	—

Figure 7: Extracted phrase patterns

The judgement term “terrible” is related to the closest aspect “strap” in review number one. Similarly, the judgement term “unbelievable” is related to the closest aspect “images” in the next review example. Opinion words/phrases are mainly adjectives/adverbs that are used to qualify product features with noun(s) phrases. The closest adjectives can be extracted as opinion words/phrases based on the availability of features in the sentences. For review number three, the feature “battery” was not found to be related to the closest adjective with the opinion word “long.”

Adverbs and adjectives are good indicators of subjectivity and opinions. Therefore, phrases containing noun(s), verbs, adverbs, and adjectives that signify opinions should be extracted. Verbs such as like, dislike, love, appreciate, recommend, and prefer and adverbs such as really, overall, absolutely, never, always, not, and well are considered opinion words in this work.

Consequently, we fetch 2/3 consecutive terms after the POS tagging of several patterns. We collect all opinionated phrases of mostly 2/3 words such as (ADJ, NN), (ADJ, NN, NN), (AD-VERB, ADJ), (ADVERB, ADJ, NN), and (VRB, NN) onward after processing the POS-tagged reviews. [Tab. 3](#) offers some examples of opinion phrases. The patterns obtained are later used to match and find opinion words/phrases from tweets and reviews after POS tagging is performed. Although the patterns are quite comprehensive, some of the expected opinion words/phrases were not detected from the dataset. Most adverbs/adjectives extracted by applying these patterns signify opinions regarding the nearest noun words/phrases.

[Tab. 3](#) provides a few examples of applying sentiment detection to the extracted aspect, for example, (low battery), (good celebrations), (remarkable camera), (extremely happy), (very relaxed), (really annoying), (totally incredible), (accomplished well), which were successfully detected by our method.

4.6 Polarity Weight Calculation

Generally, context words are taken as aspects. A sentiment dictionary of words or lexicons is developed for each type of domain separately for identifying opinions from the text in

SENTIWORDNET. The units of resources are verbs and adjectives, such as ⟨No⟩ + ⟨adjective + verb⟩. All the sentiment units are mapped or linked by calculating similarity scores. PC_{OCCR} (pointwise occurrence) is used to calculate the similarity among different units.

Table 3: Examples of sentiment detection

POS Tags	Sentiments
(Adjective, noun)	(Low battery), (good memories), (awesome camera), and so forth
(Adjective, noun, noun)	(High-quality pictures)
(Adverb, adjective)	(Extremely pleased), (very easy), (really annoying), (absolutely amazing), and so forth
(Adverb, adjective, noun)	(Very compact camera), (very good pictures), and so forth
(Adverb, verb)	(Personally recommend) (not so bad), and so forth
(Verb, noun)	(Recommend camera), (appreciate picture), and so forth
(Verb, adverb)	(Perform well)

To calculate positive and negative PC_{OCCR} , we used Eq. (6) for positive polarity and Eq. (7) for negative polarity.

$$P_{+ve}(w_i, +ve) = \log_2 \frac{\sum_i^n (w_i, +ve) \cdot N}{\sum_i^n (w_i) \cdot \sum_i^n (+ve)} \quad (6)$$

$$P_{-ve}(w_i, -ve) = \log_2 \frac{\sum_i^n (w_i, -ve) \cdot N}{\sum_i^n (w_i) \cdot \sum_i^n (-ve)} \quad (7)$$

Eq. (8) is employed on a given text from the complete dataset to calculate the PC_{OCCR} scores, from which the highest similarity score is selected as a final sentiment score. Tab. 4 provides example results for different terms, and step-by-step details of the computation of polarity estimation scores are provided in Algorithm 3.

Table 4: Example results

Multi-domain amazon dataset Aspect	Subjectivity	Polarity estimation			Results
	Polar %	Positive (+ve)	Negative (-ve)	Neutral	Sentiment
Awesome movie	70.2	0.702	0.177	0.121	Positive
Great screen	68.6	0.686	0.221	0.093	Positive
Reasonable price	59.2	0.592	0.299	0.109	Positive
Low battery	50.3	0.132	0.503	0.365	Negative
Helpful service	68.9	0.273	0.038	0.689	Neutral

Algorithm 3: Computing the polarity estimation scores**Input:** Final aspect terms, T, from the CAC model**Output:** Pointwise occurrence, PC (OCCR), score with sentiment**Start**

```

FOR Aspect, T, in set of aspects DO
  FOR add patterns in pattern[i] DO
    Pattern[i] = "JJ," "NN/NNS," "null"
    Pattern[i] = "JJ," "NN/NNS," "NN/NNS"
    Pattern[i] = "RB/RBR/RBS," "JJ," "null"
    Pattern[i] = "RB/RBR/RBS," "JJ/RB/RBR/RBS," "NN/NNS"
    Pattern[i] = "RB/RBR/RBS," "VBN/VBD," "null"
    Pattern[i] = "RB/RBR/RBS," "RB/RBR/RBS," "JJ"
    Pattern[i] = "VBN/VBD," "NN/NNS," "null"
    Pattern[i] = "VBN/VBD," "RB/RBR/RBS," "null"
    MatchPattern(w, Pattern[i])
    IF (Pattern_Found.Count > 0) THEN
      Compute_Positive_Score(w, pos_i)
      FOR w in Positive pos_i DO
        | w, pos_i = log2((Count(pos_i))/(w*Count(pos_i)))*N
      Compute_Negative_Score(w, neg_i)
      FOR w in Negative neg_i DO
        | w, neg_i = log2((Count(neg_i))/(w*Count(neg_i)))*N
      Compute_PC(OCCR)_Score(pos_i, pos_n)
      FOR w in pos_i && neg_i DO
        | T_Score = pos_i - neg_i
        | Sum_T_Score+ = T_Score
      T_Score = Sum_T_Score;

      IF (MAX(T_Score).Contains(pos_i)) THEN
        | Final_Sentiment = Positive
      ELSE
        IF (MAX(T_Score).Contains(neg_i)) THEN
          | Final_Sentiment = Negative
        ELSE
          Final_Sentiment = Neutral
      ELSE MatchPattern(w, Pattern[i])
  End
End

```

For the polarity calculation, the following notations are used:

$$\sum_i^n (w_i, +ve): \text{Number of aspect terms } w_i \text{ in positive reviews.}$$

$$\sum_i^n (w_i, -ve): \text{Number of aspect terms } w_i \text{ in negative reviews.}$$

N : Total number of tokens in the dataset.

$\sum_i^n (w_i)$: Number of aspect terms w_i in all reviews.

$\sum_i^n (+ve)$: Total number of terms in positive reviews.

$\sum_i^n (-ve)$: Total number of terms in negative reviews.

$$T_{score}(w_i) = P_{+ve}(w_i, +ve) - P_{-ve}(w_i, -ve) \quad (8)$$

5 Experiments

In this section, we discuss the experimental setup and evaluation of our proposed method.

5.1 Datasets

For the evaluation, we used three different types of datasets: tweets, customer reviews and online forums. The distribution of datasets is shown in [Tab. 5](#).

Table 5: Distribution of datasets

Domain	Customer review		Twitter	Trust pilot forum
Dataset	Books	DVDs	Tweets	
Reviews/tweets	1250	1500	2000	300
No. of reviews/tweets after preprocessing	1052	1236	1675	250
No. of aspects identified	776	1021	1296	137
Positive	246	359	507	32
Negative	167	203	313	42
Neutral	363	459	476	63

5.1.1 Tweets

We utilize the SemEval-2017 task 4 sentiment analysis in Twitter; here, 50,000 tweets [26] of goods are used. We took English tweets from Twitter using the WOEID [6,27].

5.1.2 Customer Reviews

The multi-domain sentiment dataset [26] contains product reviews taken from META-SHARE. We gathered a domain-related dataset (Books, DVDs). It includes the core aspects identified, aspect polarities and overall sentence polarities. Each domain has several thousand reviews, but the exact number varies by domain [24].

5.1.3 Trust Pilot Forum

Trust pilot [28] is a source of online reviews for businesses and online transaction areas. This research collects some reviews randomly from this forum and computes results to evaluate our proposed method. For judgements, we share data with product experts who have used these products before. They rate the aspects of these products using a rating scale of 1, 2, or 3. However, a majority vote determines the final decision. If all judges have a different opinion

(i.e., $\text{judge1} \neq \text{judge2} \neq \text{judge3} = 0$), then the review is not considered. [Tab. 6](#) shows the results of the judgement procedure. The rating scale is as follows:

Scale value 1: Not a related aspect of the product.

Scale value 2: Minor related aspect of the product.

Scale value 3: Related aspect of the product.

Table 6: Sample judgement results of the aspects of trust pilot forum reviews

Aspect	Judge 1 score	Judge 2 score	Judge 3 score	Final score (majority basis)	Final decision
Display	1	3	1	1	Not relevant
Screen	3	3	2	3	Relevant
Case	2	1	3	0	No result
Charger	2	2	3	2	Minor relevance

5.2 Baselines

We use different baseline methods for the experimental comparisons. The baseline methods are provided in Sections 5.6–5.8.

5.3 Performance Measures

This research uses precision, recall, f-measure, and accuracy to calculate the relevance and irrelevance of the mined aspects or features and opinions from reviews. [Eqs. \(9\)–\(12\)](#) show each of the performance measures.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (9)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (10)$$

$$F\text{-Measure} = 2 \cdot \frac{(\text{Precision} \cdot \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (11)$$

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{(\text{True Positive} + \text{False Negative} + \text{False Positive} + \text{True Negative})} \quad (12)$$

5.4 Results and Discussion

This research compares our method with SWIMS [29], Diego Terrana [27] and manually annotated data. For the manually labelled dataset, we label the review ratings regarding how appropriate the aspects are to the products using the abovementioned rating scale.

5.5 Qualitative Example

In this section, we explain with an example where our method works fine compared to the baselines and why.

Review Example: “The transportation facility was good, but the vehicle was too old when we reached the hotel; the accommodation was better than their transportation.”

In the above example, a wide range of information may not reflect the aspect strongly. The person is talking about one service being too bad compared to other services and another service being much better than the previous one.

The baseline methods use dataset-based general sentiments; when a complex sentence structuring is present, they produce majority-based results or could consider them neutral [27,29].

In our model, first, we calculate the subjectivity score to check whether the sentence has sentiment or not. Second, the aspect co-occurrence is calculated, and the aspects are ranked to obtain the top (with the maximum scores) aspects. Finally, polarity estimation is used to calculate the polarity of each aspect and then summed to find the intensity of polarity. Considering the above review Senti score of 2.125, the subjective score is 0.965. Next, in the aspect extraction phase, the most relevant aspects are “vehicle,” “hotel”, “transportation,” and “accommodation.” During the polarity estimation phase, we find the sentiments of these aspects (transportation: good, vehicle: old, hotel: better, and accommodation: better), after which we calculate the intensity (good: 0.85, old: 0.69, better: 0.75). Therefore, the final result of the above review is the intensities of the sentiments (Positive: 0.57, Negative: 0.29, and Neutral: 0.14)

5.6 Comparison (Customer Reviews)

In our proposed method, named aspect-based sentiment analysis—polarity estimation of reviews (ABSA-PER), we use the new aspect term extraction method CAC and rank them. Tab. 7 shows a comparison of the aspect term extraction after sentiment term detection is applied to calculate the polarity estimation. Tab. 8 presents an analysis of the polarity estimation, which is discussed in detail in Section 3. Fig. 8 visualizes a comparison of ABSA-PER aspect extraction with existing state-of-the-art approaches [29].

Table 7: Comparison of aspect term extraction with existing approaches

Multi-domain dataset		SWIMS—10-fold baseline approach			SWIMS [30] baseline approach			ABSA-PER proposed approach		
		P	R	F	P	R	F	P	R	F
Amazon	DVD	76.20	78.80	77.39	80.00	84.00	81.95	81.35	84.62	82.95
	Books	76.83	81.80	79.23	80.39	82.00	81.19	82.93	83.79	83.35

Table 8: Comparison of sentiment polarity with the state-of-the-art approach

Dataset	Label	Steinberger et al. [31]			ABSA-PER proposed approach		
		P	R	F	P	R	F
Amazon	# Positive reviews	72.12	75.3	73.671	81.24	82.69	81.95
	# Negative reviews	74.07	66.04	69.83	83.35	78.48	80.84
	# Neutral reviews	71.35	76.35	73.65	85.78	84.59	85.18

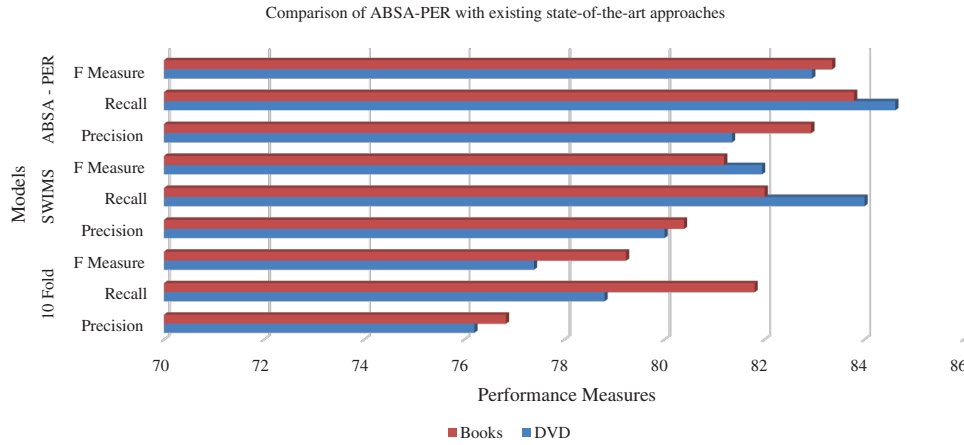


Figure 8: Visual comparison of our approach with existing state-of-the-art approaches

5.7 Comparison (Twitter)

Our proposed technique is compared with the baseline method of Terrana et al. [27]. In this comparison, initially, a Twitter dataset is built, and preprocessing is applied to subjective segments to finalize the cleansing of tweets. Then, we apply aspect term extraction using the CAC. The categorization is made using the polarity estimation model. Experimental outcomes reveal that our methodology reduces human involvement and the sub-processes. Tab. 9 shows the polarity estimation on the Twitter dataset and a comparison with the baseline. Fig. 9 shows our model on polarity estimation of the Twitter dataset and its comparison with the baseline.

Table 9: Polarity estimation on a Twitter dataset

Twitter (class label)	Terrana et al. [27] baseline approach	ABSA-PER proposed approach
	Accuracy	Accuracy
No. of aspects identified	0.741	0.895
No. of positive reviews	0.761	0.879
No. of negative reviews	0.863	0.947
No. of neutral reviews	0.977	0.872
Average accuracy %	83.55	89.82

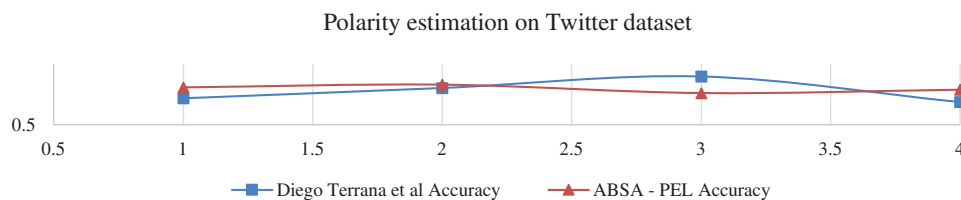


Figure 9: Visualization of polarity estimation on the Twitter dataset

5.8 Comparison (Gold Standard)

We do not use any baseline method here; rather, we use human annotators. We collect random reviews from forums and select the top N correct reviews, with which several experts perform annotations. They convert these reviews into readable sentences, obtain useful aspects or features for the product of interest, and classify them into sentiment classes (positive, negative and neutral). After that, we use the same forum dataset for our proposed ABSA-PER and find that the results are in good agreement. We also compute accuracy with respect to the human annotator results. Tab. 10 shows the difference between the results of our proposed method and the human annotation results.

Table 10: Experimental results for human annotations

Dataset	No. of reviews	No. of sentences	No. of aspects extracted %	Accuracy %
Our approach	300	216	113	87

5.9 Overall Performance of Our Proposed Technique

An overall comparison of our technique on different datasets [27,29] is provided in Fig. 10. The results show a clear improvement compared to the baseline methods, and good accuracy is achieved on the forum dataset as well in comparison to the human-annotated data.

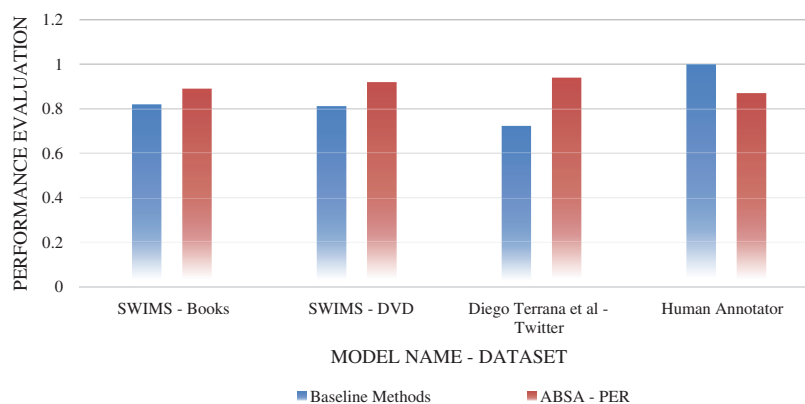


Figure 10: Overall performance of the proposed technique (ABSA-PER)

6 Conclusions and Future Work

In today's research, determining sentiment polarity can be considered one of the significant tasks in opinion mining. Polarity estimation has standard methods to recognize aspects or features, which are regularly in use and have gained much recognition, e.g., Screen, Memory and Battery. It can be concluded from the results and discussion that theme-specific polarity estimation in which a diverse type of information is present is essential. It is also evident that the exploitation of aspect-based sentiment analysis provides better insights into review and tweet data. In social media, tweets can be examined for the identification of a specific occasion, which may be an emerging event. There are distinct possibilities to improve this work. For example, multi-word

clustering enables the exploration of reviews in a better way, e.g., “hard” is one word, and “disk” is another, but their combination (“hard disk”) has completely different semantics. A Vector representation of these multi-words may help improve the performance of different approaches. Furthermore, objectivity feature exploration for tweets and reviews can improve the correctness of aspects and improve accuracy.

Acknowledgement: The authors acknowledge with thanks the University technical and financial support.

Funding Statement: This work was funded by the *University of Jeddah*, Saudi Arabia, under Grant No. (UJ-12-18-DR).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] N. Shelke, S. Deshpande and V. Thakare, “Domain independent approach for aspect-oriented sentiment analysis for product reviews,” in *Advances in Intelligent Systems and Computing*, Singapore: Springer, pp. 651–659, 2017.
- [2] F. H. Khan, U. Qamar and S. Bashir, “A semi-supervised approach to sentiment analysis using revised sentiment strength based on SentiWordNet,” *Knowledge and Information Systems*, vol. 51, no. 3, pp. 851–872, 2017.
- [3] K. Sarawgi and V. Pathak, “Opinion mining: Aspect level sentiment analysis using SentiWordNet and amazon web services,” *International Journal of Computer Applications*, vol. 158, no. 6, pp. 31–36, 2017.
- [4] M. A. Masood, R. A. Abbasi and N. W. Keong, “Context-aware sliding window for sentiment classification,” *IEEE Access*, vol. 8, pp. 4870–4884, 2020.
- [5] B. Kama, M. Ozturk, P. Karagoz, I. H. Toroslu and O. Ozay, “A web search enhanced feature extraction method for aspect-based sentiment analysis for Turkish informal texts,” *Big Data Analytics and Knowledge Discovery*, LNCS, vol. 9829, pp. 225–238, 2016.
- [6] H. Saif, Y. He, M. Fernandez and H. Alani, “Contextual semantics for sentiment analysis of Twitter,” *Information Processing Management Journal*, vol. 52, no. 1, pp. 5–19, 2016.
- [7] V. Gayakwad and A. Tiwari, “Improved technique of sentiment classification for objective word,” *International Journal of Emerging Research in Management Technology*, vol. 5, no. 6, pp. 105–112, 2016.
- [8] K. Schouten and F. Frasinca, “Survey on aspect-level sentiment analysis,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 28, no. 3, pp. 813–830, 2016.
- [9] M. D. Molina-González, E. Martínez Camara, M. T. Martín-Valdivia and L. A. Urena Lopez, “A Spanish semantic orientation approach to domain adaptation for polarity classification,” *Information Processing Management*, vol. 15, no. 4, pp. 520–531, 2015.
- [10] K. Ravi and V. Ravi, “A survey on opinion mining and sentiment analysis: Tasks, approaches and applications,” *Knowledge Based System*, vol. 89, pp. 14–46, 2015.
- [11] M. Al-Smadi, M. Al-Ayyoub, Y. Jararweh and O. Qawasmeh, “Enhancing aspect-based sentiment analysis of Arabic hotels’ reviews using morphological, syntactic and semantic features,” *Information Processing & Management*, vol. 56, no. 2, pp. 308–319, 2019.
- [12] M. Rahman and E. Kumar Dey, “Datasets for aspect-based sentiment analysis in Bangla and its baseline evaluation,” *Data*, vol. 3, no. 2, pp. 15, 2018.
- [13] H. H. Do, P. W. C. Prasad, A. Maag and A. Alsadoon, “Deep learning for aspect-based sentiment analysis: A comparative review,” *Expert Systems with Applications*, vol. 118, pp. 272–299, 2019.
- [14] H. Cho, S. Kim, J. Lee and J. S. Lee, “Data-driven integration of multiple sentiment dictionaries for lexicon-based sentiment classification of product reviews,” *Knowledge Based System*, vol. 71, pp. 61–71, 2014.

- [15] P. Takala, P. Malo, A. Sinha and O. Ahlgren, “Gold-standard for topic-specific sentiment analysis of economic texts,” in *Proc. of the Ninth Int. Conf. on Language Resources and Evaluation*, Reykjavik, Iceland, pp. 2152–2157, 2014.
- [16] A. Montejo-Raez, M. C. Díaz-Galiano, F. Martínez-Santiago and L. A. Ureña-López, “Crowd explicit sentiment analysis,” *Knowledge Based System*, vol. 69, pp. 134–139, 2014.
- [17] F. Bravo-Marquez, M. Mendoza and B. Poblete, “Meta-level sentiment models for big social data analysis,” *Knowledge Based System*, vol. 69, pp. 86–99, 2014.
- [18] S. Zhou, Q. Chen, X. Wang and X. Li, “Hybrid deep belief networks for semi-supervised sentiment classification,” in *25th Int. Conf. on Computational Linguistics Technical Papers*, Dublin, Ireland, pp. 1341–1349, 2014.
- [19] S. Poria, E. Cambria, G. Winterstein and G. B. Huang, “Sentic patterns: Dependency-based rules for concept-level sentiment analysis,” *Knowledge Based System*, vol. 69, pp. 45–63, 2014.
- [20] A. Weichselbraun, S. Gindl and A. Scharl, “Enriching semantic knowledge bases for opinion mining in big data applications,” *Knowledge-Based Systems*, vol. 69, pp. 78–85, 2014.
- [21] C. M. Faisal, A. Daud, F. Imran and S. Rho, “A novel framework for social web forums’ thread ranking based on semantics and post quality features,” *The Journal of Supercomputing*, vol. 72, no. 11, pp. 4276–4295, 2016.
- [22] J. Smailović, M. Grčar, N. Lavrač and M. Žnidaršič, “Stream-based active learning for sentiment analysis in the financial domain,” *Information Sciences*, vol. 285, pp. 181–203, 2014.
- [23] H. U. Khan, A. Daud, U. Ishfaq, T. Amjad, N. Aljohani *et al.*, “Modelling to identify influential bloggers in the blogosphere: A survey,” *Computers in Human Behavior*, vol. 68, pp. 64–82, 2017.
- [24] M. A. Jarwar, R. A. Abbasi, M. Mushtaq, O. Maqbool, N. R. Aljohani *et al.*, “CommuniMents: A framework for detecting community based sentiments for events,” *International Journal on Semantic Web and Information Systems*, vol. 13, no. 2, pp. 87–108, 2017.
- [25] P. M. B. Vitányi and R. L. Cilibrasi, “Normalized web distance and word similarity,” *arXiv*, pp. 317–338, 2010.
- [26] “Tweet object—Twitter developers. (n.d.),” 2017. [Online]. Available: <https://dev.twitter.com/overview/api/tweets>.
- [27] D. Terrana, A. Augello and G. Pilato, “Automatic unsupervised polarity detection on a Twitter data stream,” in *IEEE Int. Conf. on Semantic Computing*, Newport Beach, CA, USA, pp. 128–134, 2014.
- [28] “Trustpilot API documentation (n.d.),” 2017. [Online]. Available: <https://developers.trustpilot.com/>.
- [29] F. H. Khan, U. Qamar and S. Bashir, “SWIMS: Semi-supervised subjective feature weighting and intelligent model selection for sentiment analysis,” *Knowledge-Based Systems*, vol. 100, pp. 97–111, 2016.
- [30] “Amazon developer services (n.d.),” 2017. [Online]. Available: <https://developer.amazon.com/>.
- [31] J. Steinberger, T. Brychcín and M. Konkol, “Aspect-level sentiment analysis in Czech,” in *Proc. of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, Baltimore, Maryland, pp. 22–30, 2014.