



# HAS: Hybrid Analysis of Sentiments for the perspective of customer review summarization

Gagandeep Kaur<sup>1,2</sup> · Amit Sharma<sup>3</sup>

Received: 5 August 2021 / Accepted: 19 January 2022 / Published online: 20 February 2022  
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

The reviews posted online by the end-users can help the business owners obtain a fair evaluation of their products/services and take the necessary steps. However, due to the large volume of online reviews being generated from time to time, it becomes challenging for business owners to track each review. The Customer Review Summarization (CRS) model that can present the summarized information and offer businesses with significant acumens to understand the reason behind customers' choices and behavior, would therefore be desirable. We propose the Hybrid Analysis of Sentiments (HAS) for the perspective of effective CRS in this paper. The HAS consists of steps like pre-processing, feature extraction, and review classification. The pre-processing phase removes the unwanted data from the text reviews using Natural Language Processing (NLP) based on different pre-processing functions. For efficient feature extraction, the hybrid mechanism consisting of aspect-related features and review-related features is proposed to build the unique feature vector for each customer review. Review classification is performed using different supervised classifiers like Support Vector Machine (SVM), Naïve Bayes, and Random Forest. The experimental results show that HAS efficiently performed the sentiment analysis and outperformed the existing state-of-the-art techniques with an F1 score of 92.2%.

**Keywords** Aspect category extraction · Customer review summarization · Feature extraction · Hybrid features · Sentiment analysis · Supervised classification

## 1 Introduction

The emergence of Web 2.0 (Vila and Ribeiro-Soriano 2014; Baxter and Connolly 2014) and the Internet of Things (IoT) (Mahajan and Badarla 2018, 2019, 2021; Mahajan et al. 2021) enabled web users to express their opinions on numerous social media platforms and e-commerce portals. It has resulted in an exponential increase in reviews, which can help make informed decisions about a product, brand, national or international event. The Customer Review

Summarization (CRS) has become a vital requirement for business owners to enhance their services and products by evaluating the reviews. Since the arrival of the Covid-19 threat, there has been a tremendous increase in online delivery of things like food, electronic products, clothes, etc. (Bafna and Toshniwal 2013; Khan and Jeong 2016). It has led to many online reviews being posted by customers from time to time. Such reviews contain vital information concerning the customer's needs and satisfaction. The business such as restaurants, e-commerce, and movies use such reviews to enhance their productivity by systematically studying the product-specific reviews. Additionally, the reviews assist in the exhaustive analysis of competitors. However, it is challenging to mine reviews due to the volume, variety, veracity, and velocity at which they are constantly generated online.

Researchers have been working intensively on developing modeling frameworks that can help mine these corpora of information using Artificial Intelligence (AI), Natural Language Processing (NLP), Sentiment Analysis, Information Extraction, and Information Retrieval (Hanni et al. 2016). At

✉ Gagandeep Kaur  
gagandeep.kaur@sitpune.edu.in

Amit Sharma  
profamitsharma@gmail.com

<sup>1</sup> Department of CSE, Lovely Professional University, Punjab, India

<sup>2</sup> Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University), Pune, India

<sup>3</sup> School of Computer Applications, Lovely Professional University, Punjab, India

present various methods have been introduced since the last decade, however, large volume, accuracy, and scalability is a major challenge to performing the CRS. The main goal of this paper is to present the automatic, robust, and efficient customer review summarization model for efficient sentiment analysis using a hybrid approach.

The functionality of CRS proposed in this paper mainly consists of two phases: Review Mining (RM) and Review Summarization (RS). In RM, the reviews received from each user are classified as positive, negative, or neutral. In RS, according to the outcome of the RS step, a concise summary gets automatically generated for corresponding reviews.

The scope of this paper is limited to the accurate and robust RM for effective CRS. The Review Mining phase (RM) phase is also called sentiment analysis. Sentiment analysis includes various tasks such as detection of subjectivity, polarity, sentiment magnitude, and type of emotion (Ramírez-Tinoco et al. 2017). Subjectivity detection is the subtask wherein the text gets classified as subjective (positive or negative sentiment) or objective (neutral, factual, or without any sentiment). Polarity detection is the most popular subtask and includes an overall classification of text as either positive or negative, without considering the strength of sentiment. Next, sentiment magnitude refers to the detection of a more in-depth emotion, in other words, it is the extraction of feelings like how much hatred, sadness, and happiness are there inside the comment. The final subtask involves the detection of the type of emotion including sad, anger, excitement, happiness, and so on (Lighthart et al. 2021).

Sentiment analysis can be performed at the word level, phrase level, sentence level, and document level (Singh et al. 2016). First comes the sentiment analysis at the word level, which includes determining the sentiment of a person, product or its aspect, brand, or any other entity. For example, the review, “I am happy with iPhone’s battery’s performance” shows a positive sentiment for the aspect battery of the entity iPhone. Next, phrase-level sentiment analysis involves detecting the sentiment of multi-word expressions. Further, sentence-level sentiment analysis includes detecting the overall sentiment of a sentence. Finally, document-level analysis determines overall sentiment using average or weighted methods on one or more sentences (Hussein 2018; Schouten and Frasincar 2016).

The last decade has witnessed many techniques and models for the task of sentiment analysis. The most widely employed approaches include machine learning for text classification and rule-based lexical methods for text labeling (Schouten and Frasincar 2016; Moussa et al. 2018; Trilla and Alias 2013; Liu et al. 2012a, b; Liu et al. 2012a, b; Pannala et al. 2016). Both the approaches have shown great success in traditional text sources, formal language, and well-defined domains, where pre-labeled data is available for

training or lexicon is sufficient to cover the variety and range of sentiment-bearing words in the given corpus. However, these methods are incapable of matching up to the volume, velocity, and variety of informal and unstructured data being generated constantly on the internet.

Recently, improvement in the performance of sentiment analysis methods using machine learning has been observed due to the introduction of different types of feature extraction techniques which play a vital role in accurate sentiment analysis (Shahana and Omman 2015; Parlar et al. 2018). The feature extraction phase suffers from various challenges such as detecting ambiguous and unreliable features for classification. To overcome such challenges, the hybrid model of feature extraction is required to be designed for an accurate outcome. Along with sentence or review specific features, specific product/topic, its aspect, emoticons should also be considered for clear and reliable feature extraction.

For illustration, consider the review “iPhone holds the charge for a longer period, heavy on the pocket though”, including implicit aspects (charge, pocket) and sentiment-bearing word relations. Overall sentiment seems to be neutral, but aspect-based sentiments have positive and negative polarities. Moreover, these implicit aspects are not well-defined and not expressed as popular synonyms or standard forms. One solution to this problem is combining very similar aspects to form attributes, followed by attribute-sentiment analysis. It is also called aspect-based sentiment analysis (Kok et al. 2018; Hossain et al. 2020; Chakraborty et al. 2020; Hitkul et al. 2020). Recently various aspect-based sentiment analysis methods have been proposed, but a more sophisticated mechanism is required for CRS that can effectively capture the implicit word relations, detect similar aspects, and handle special terms and ambiguities.

In this paper, we propose the novel mechanism of sentiment analysis using the robust Hybrid Analysis of Sentiments (HAS). We focus on the problem of performing efficient and effective RM in this paper. The HAS consists of steps including pre-processing, features extraction, and classification. We used a simple and effective pre-processing technique to remove the stop words, meaningless words, numbers, etc. We proposed the robust and effective hybrid feature extraction technique by combining the aspect-related features and review-related features into a hybrid feature vector for the input pre-processed reviews. The hybrid feature extraction approach can address ambiguity and reliability-related issues as well for sentiment analysis. First, the different review-related features are extracted and represented in numerical form. Then, the aspect-related features are extracted by detecting the aspect terms and forming the features vector using co-occurrence frequencies. Finally, different supervised machine learning techniques are applied to classify the hybrid feature vector either into positive or negative classes.

Section 2 presents a brief review of related work and contributions of this paper, Sect. 3 presents the methodology of HAS, Sect. 4 presents the experimental results and discussion, and Sect. 5 presents the conclusion and future work.

## 2 Related work

Online review mining and summarization have received significant interest from researchers due to their importance for business productivity. Accurate sentiment analysis from big-size datasets is a challenging task because of the presence of sarcasm, noisy data, emoticons, and unreliable features. As the proposed HAS method focuses on the efficient analysis of sentiments, therefore, this section includes some recent studies and categorizes them under two domains: Sentiment Analysis (SA) and Aspect-based Sentiment Analysis (ABSA). We reviewed such methods according to the methodology used, i.e., machine learning and rule-based approaches. The research motivation and contributions of this paper are discussed at the end.

### 2.1 SA methods

This section presents the review of some recent work carried out in the domain of SA (Wang et al. 2020; Hao et al. 2020; Zhu et al. 2021; Naresh and Krishna 2021; Singh et al. 2021; Munuswamy et al. 2021; Ayyub et al. 2020; Oyeboode et al. 2020; Iqbal et al. 2019; Khan et al. 2020).

Wang et al. (2020) proposed the SentiDiff algorithm for Twitter data SA. The authors first analyzed sentiment diffusion using the sentiment reversal method and then discovered interesting characteristics from it. The inter-relationships among sentiment diffusion patterns and textual information were used for predicting the sentiment polarities in an input twitter message. The stochastic word embedding mechanism is recently introduced by Hao et al. (2020) for cross-domain sentiment encoding. The authors explored the simple mapping of occurrence information and word polarity and encoded that information for accurate analysis with minimum computational efforts.

The SentiVec mechanism proposed by Zhu et al. (2021) for SA uses the kernel optimization technique for word embedding. The authors used supervised learning in the first phase and unsupervised learning models in the second phase. The Twitter data classification model proposed by Naresh and Krishna (2021) uses the optimization-based machine learning technique. The model consists of steps such as data pre-processing, feature extraction, and sentiment classification. The sentiment analysis of the real-world twitter datasets on Covid-19 is performed by Singh et al. (2021) using the Bidirectional Encoder Representations from Transformers (BERT).

The sentiment dictionary-based approach for sentiment analysis of end-users is proposed by Munuswamy et al. (2021). According to the sentiment prediction, the automatic recommendation is provided to end-users for product purchase. The authors used the n-gram feature extraction technique and Support Vector Machine (SVM) as the prediction method. Ayyub et al. (2020) carried out the investigation study to explore the various feature extraction and machine learning techniques for SA. They performed an empirical study on traditional machine learning algorithms, deep learning techniques, and ensemble classifiers.

Another investigation study using the real-world dataset is presented by Oyeboode et al. (2020) for the sentiment analysis using machine learning techniques. The authors collected 88,125 reviews to evaluate mental health apps. Iqbal et al. (2019) proposed a Genetic Algorithm (GA) based feature reduction mechanism to bridge the gap between the machine learning and lexicon-based techniques for improving scalability and accuracy. The simple mechanism for SA is proposed by Khan et al. (2020) using the Bag of Words (BoW) feature extraction and Naïve Bayes (NB) classifier. The proposed model classifies the movie reviews either as positive or negative. The authors explored the semantic graph-based technique for the summarization of classified reviews.

### 2.2 ABSA methods

Due to the presence of sarcasm, emotions, feelings, and opinion-related features, accurate and reliable features representations become an important research problem for reviews mining. Aspect terms extraction for feature formation has received significant attention from researchers as it remarkably enhances the SA accuracy. Recently, several studies have been proposed on ABSA (Schouten et al. 2018; Alqaryouti et al. 2019; Wang 2021; Kumar et al. 2020; Nandal et al. 2020; Li et al. 2020; Prathi et al. 2020; Alamanda 2020; Shams et al. 2020; Bie and Yang 2021).

The supervised, as well as unsupervised techniques, have been utilized by Schouten et al. (2018) to extract the aspects for SA. The authors have applied the association rule mining technique on the co-occurrence frequency data acquired from a corpus to determine aspect types. The ABSA method proposed by Alqaryouti et al. (2019) used the hybrid mechanism of consolidating the rules and domain lexicons to analyze the elements of apps reviews. The proposed method extracts the vital aspects from the input reviews and classifies them into sentiments.

The novel approach of Cross-Lingual Sentiment Classification (CLSC) is proposed by Wang (2021) for sentiment classification. The aspect, opinion, and sentiment classification model has been designed using the unsupervised machine learning technique. The authors utilized the coarse alignment technique for precision latent features extraction.

Another ABSA mechanism proposed by Kumar et al. (2020) uses the ontology-based approach that consists of steps such as semantic features extraction using ontologies, Word2Vec conversion, and Convolutional Neural Network (CNN).

The simple technique of ABSA proposed by Nandal et al. (2020) extracts the item features and uses aspect-based sentiment classification. The novel framework called the SEML (SEmi-supervised Multi-task Learning) framework has been proposed by Li et al. (2020) for ABSA. The authors applied Cross-View Training (CVT) for learning of semi-supervised sequence and bidirectional Recurrent Neural Network (RNN). The cold-start problem of the ABSA technique has been solved by Prathi et al. (2020) by computing the automatic sentiments from the input reviews for dynamic aspects extraction.

The polarity classification and extraction of sentiment from the input reviews for efficient ABSA is proposed by Alamanda (2020). The interesting polarity aspects are automatically extracted according to customers' preferences using deep learning and machine learning techniques. The LISA (Language Independent aspect-based SA) recently proposed by Shams et al. (2020) uses three coarse-grained phases. The prior domain knowledge is extracted in the first phase. In the second phase, the document is divided into different aspects and then the probability is calculated in the third phase.

Most recently the study that investigated the ABSA mechanism and designed novel MTMVN (MultiTask MultiView Network) is presented by Bie and Yang (2021). The authors considered the ABSA as the main task and considered the two subtasks: aspect term mining and aspect opinion prediction as supporting tasks. The representation acquired from the branch network of the primary task is considered as the global view, while the representations of the two subtasks are regarded as two local views with distinct prominences. Multitask learning facilitates the primary task by considering precise aspect boundary information and opinion polarity information. The authors further optimized the performance of the model by augmenting the correlations between the views beneath the concept of multiview learning.

Other than these works, deep learning has received significant attention for sentiment analysis (Lu et al. 2021; Datta and Chakrabarti 2021), but due to higher training and computational time, we do not consider it as the cost-effective mechanism for SA.

### 2.3 Motivation

In recent times, sentiment analysis has received significant attention from researchers to accomplish different applications, therefore a framework to perform scalable and efficient SA is required. We studied various methods to perform SA with the focus on effective feature representation and

machine learning techniques for classification. The feature extraction is a difficult task because of the presence of fake, spam, sarcasm, negation, emoticons, etc., in a large number of online reviews. Machine learning techniques such as supervised and unsupervised methods can be used to classify online reviews into positive or negative sentiments. We believe that the accuracy of machine learning techniques mainly depends on the feature set.

From the above studies, we observed that SA-based methods (Wang et al. 2020; Hao et al. 2020; Zhu et al. 2021; Narsh and Krishna 2021; Singh et al. 2021; Munuswamy et al. 2021; Ayyub et al. 2020; Oyeboode et al. 2020; Iqbal et al. 2019; Khan et al. 2020) are not sufficient to address the challenges related to the exact representation of feelings, emotions, and opinions from online reviews. For that purpose, the ABSA techniques (Schouten et al. 2018; Alqaryouti et al. 2019; Wang 2021; Kumar et al. 2020; Nandal et al. 2020; Li et al. 2020; Prathi et al. 2020; Alamanda 2020; Shams et al. 2020; Bie and Yang 2021) that have been proposed recently are studied but they also have limited scope or limited investigations. Effective feature representation is a challenging task as the use of aspect-related features addresses the ambiguities/sarcasm but fails to handle the negations, emotions, opinions, etc.

Apart from these challenges, some recent ABSA methodologies have not been evaluated through scalable review datasets. Some ABSA methods heavily relied on unsupervised techniques; therefore, they required many manual annotations to data. Some ABSA/SA methods depend only on symbolic feature extraction and hence result in poor accuracy. Among the machine learning methods, the supervised classifiers delivered better scalability and efficiency compared to unsupervised or rule-based techniques.

### 2.4 Contribution

By considering the above challenges of state-of-the-art methods, we propose the novel approach of SA with special reference to CRS. The proposed method consists of pre-processing, feature extraction, and classification phases. The pre-processing uses the common methodology to denoise the input reviews. The feature extraction phase is intelligently designed as a HAS to address the challenges of accurate feature representation of input reviews. The main contributions of the proposed model are:

- As it has been observed that hybrid features set perform efficiently, we divided the feature extraction task into two sub-tasks: Aspect-Related Features (ARF) and Review-Related Features (RRF). Finally, the ARF and RRF numeric features are fused to form the hybrid feature vector.

- In ARF, we propose a novel approach to extract the aspects terms using co-occurrence frequencies and then assign the polarities. The outcome of ARF is the aspect terms with their polarities.
- In RRF, different techniques like Term Frequency-Inverse Document Frequency (TF-IDF), emoticons polarities, and n-gram features are used to represent the end-user emotions, opinions, and feelings accurately.
- For the classification purpose, supervised classifiers such as Naïve Bayes (NB), SVM, and Random Forest (RF) are used. We divided the online review dataset into training and testing datasets for performance analysis.

### 3 Proposed system

From the above recent studies, we found that SA has grown as an active research topic because of several interesting and demanding research problems. Because of its various functional applications, an immense number of start-up businesses are striving for providing SA/RM services. Each organization is interested in knowing how customers view their goods and services. The scientific difficulties and functional elements will keep the SA/RM domain active and vibrant for the coming years. In this paper, we aim to build the hybrid model for accurate SA with the perspective of CRS. Figure 1 shows the functionality of the proposed CRS-HAS model for SA.

As shown in Fig. 1, the first step is online review dataset acquisition. Then the reviews are pre-processed using NLP-based different pre-processing functions like stemming, stop word removal, URL removal, etc. The features are extracted by using the techniques: ARF and RRF. Then the hybrid feature vector is built from both ARF and RRF outcomes. The hybrid feature vector is fed into machine learning classifiers SVM, NB, and RF to classify the input reviews into positive or negative classes. We employed the NLP methods to semantically analyze the text of each review. The input reviews are first semantically pre-processed using NLP (as shown in algorithm 1), and then pre-processed reviews are fed into the hybrid feature extraction utilizing the RRF and ARF in algorithm 2 (which are semantic techniques for text features extraction). Below sub-sections elaborate all these steps in detail.

#### 3.1 Data pre-processing

The pre-processing performs the cleaning of the raw reviews by removing and correcting the complex and unnecessary text. Algorithm 1 shows the functionality of the data pre-processing that starts with tokenization and ends with removing numbers and meaningless words. The tokenization function split the input review into different tokens. Then

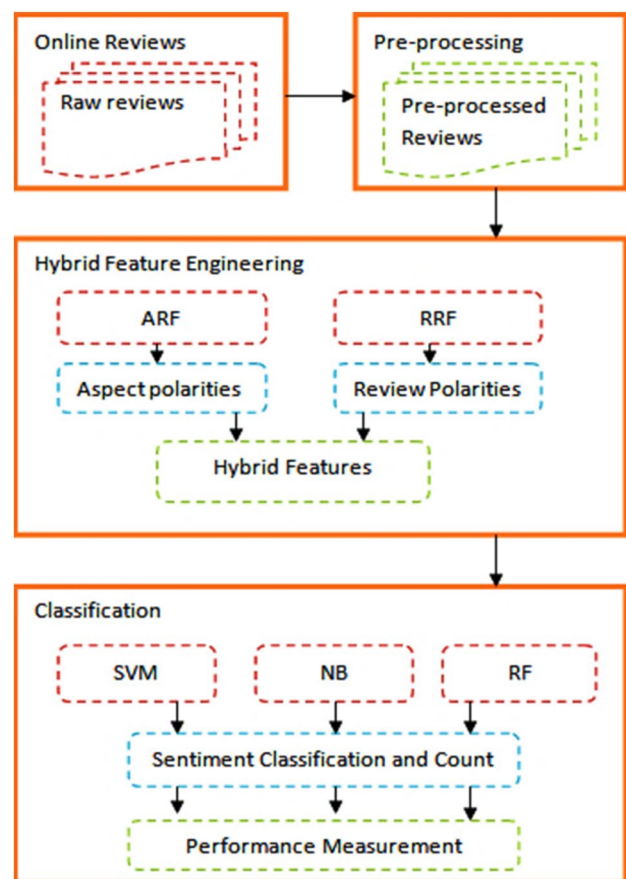


Fig. 1 Proposed framework for hybrid analysis of sentiments

on each token, we apply stemming to reduce the tokens into their singular form (e.g., performing or performed will get converted to perform). Then the stop-words like 'a', 'an', 'I', 'am' etc. are removed to reduce the number of tokens. The special characters (@, #, etc.), dates, meaningless words (a+, B-, etc.), and any URLs are also discovered and removed. Additionally, the algorithm also checks for the words with less than three characters and numbers and removes them. Algorithm 1 ensures the effective reduction in raw reviews' dimensional space. The examples of some reviews before pre-processing and after pre-processing are presented in Table 1.

**Table 1** Illustration of pre-processing algorithm

Before pre-processing	After pre-processing
“Pork shu mai was more than usually greasy and had to share a table with loud and rude family”	“Pork greasy share table loud rude family”
“Food was absolutely amazing”	“Food absolute amaze”
“I’ve asked a cart attendant for a lotus leaf wrapped rice and she replied back rice and just walked away”	“ask cart attendant lotus leaf wrap rice reply rice walk away”
“I was very disappointed with this restaurant”	“disappoint restaurant”

Algorithm 1: Data Pre-processing
<p><b>Input</b> Q: raw review</p> <p><b>Output</b> P: pre-processed review</p>
<ol style="list-style-type: none"> <li>1. Acquisition of test review Q</li> <li>2. <math>T \leftarrow \text{getWordTokenization}(Q)</math></li> <li>3. For <math>i = 1: \text{size}(T)</math></li> <li>4.   <math>S(i) \leftarrow \text{Stemming}(T(i))</math></li> <li>5. End For</li> <li>6. For <math>i = 1: \text{size}(S)</math></li> <li>7.   If <math>(S(i) \neq \text{Stopword})</math></li> <li>8.     If <math>(S(i) \neq \text{URL})</math></li> <li>9.       <math>T1 \leftarrow S(i)</math></li> <li>10.     End If</li> <li>11.   End If</li> <li>12. End For</li> <li>13. For <math>i = 1: \text{size}(T1)</math></li> <li>14.   If <math>(T1(i) \neq \text{specialchar} \parallel T1(i) \neq \text{date} \parallel T1(i) \neq \text{complexchar})</math></li> <li>15.     <math>T2 \leftarrow T1(i)</math></li> <li>16.   End If</li> <li>17. End For</li> <li>18. For <math>i = 1: \text{size}(T2)</math></li> <li>19.   If <math>(T2(i) \neq \text{numbers} \parallel \text{length}(T2(i)) &gt; 2)</math></li> <li>20.     <math>T3 \leftarrow T2(i)</math></li> <li>21.   End If</li> <li>22. End For</li> <li>23. Return <math>(P \leftarrow T3)</math></li> </ol>

### 3.2 Hybrid feature engineering

Several attempts have been made for the feature representation of input reviews. However, accurate, robust, and efficient feature extraction is still a challenging research problem for RM. Therefore, a vital part of any SA system is to build an effective and reliable feature set that delivers high classification accuracy. In this paper, we are proposing an efficient and robust SA framework using the hybrid approach of feature engineering. Therefore, we call this process the HAS. First, the RRF features are extracted by uniquely exploring the different techniques to build the polarities for each term in the pre-processed text including the emoticons and negations. Then the ARF methodology is employed to extract the aspect terms with their polarities. The ARF aims to address and represent sarcastic and ambiguous sentences as well. Finally, each pre-processed review is represented in hybrid form by combining the outcome of ARF and RRF. The HAS model aims at performing the sentiment classification either into neutral, positive, or negative class.

#### 3.2.1 RRF

The RRF is nothing but the sentence-level feature representation approach including emoticons to effectively model opinions, feelings, negations, and emotions from the input reviews. The conventional techniques such as n-gram, TF-IDF, and emoticons-specific polarities are explored in combination to generate the hybrid form of features. TF-IDF is a method involving Bag of Words (BoWs), and n-gram is dependent on the word embeddings. The use of single words for feature extraction leads to several limitations for SA. The negation challenges do not get addressed using a single word feature, and even it leads to misclassification.

To address such challenges, we first performed the n-gram feature extraction from the pre-processed reviews to get a word list. After that, the TF-IDF on the output of n-gram is applied to obtain the TF-IDF of n-gram words. The n-gram plus TF-IDF approach not only reduces the dimensional space but also represents each review more effectively. After that, the emoticons’ specific features are extracted to enhance the SA accuracy further. Finally, RRF is obtained as the combined feature vector of n-gram, TF-IDF, and

emoticons features. The process of RRF is explained below in detail.

Let  $P^i$  be the document containing the pre-processed  $i$ th review. For combined  $n$ -gram and TF-IDF, we first apply the  $n$ -gram technique to pre-processed document.  $N$ -gram is a proximate sequence of  $n$  words from a provided sample of text. If the value of  $n$  in the  $n$ -gram model is 1 ( $n=1$ ) then it is called unigram, if it is 2 ( $n=2$ ) then it is called bigram, if it is 3 ( $n=3$ ) then it is called trigram, four-gram, and so on. For instance, “Good”, “Very Good” are unigram and bigram respectively. The  $n$ -gram technique builds the set of  $n$  consecutive words from the input sentence as:

$$Ngram = getNgram(P^i, n) \tag{1}$$

where  $Ngram$  represents the set of  $n$ -grams computed from the input pre-processed document  $P^i$ . The function  $getNgram()$  takes  $P^i$  and  $n$  as parameters. In this method, we set the value of  $n$  as 2 to balance the efficiency among the negation handling and reliability.

After the outcome of  $n$ -gram is obtained, TF-IDF is applied to get the term frequency-inverse document frequency of such words lists from the training/testing dataset. The TF computes the number of times a word has appeared in the review and IDF computes the number of times a word has appeared in reviews over the total number of reviews.

$$TF - IDF = TF(Ngram^i) \times IDF(Ngram^d) \tag{2}$$

where  $Ngram^i$  represents the list of words for  $i$ th review and  $Ngram^d$  represents the list of words for the entire dataset  $d$ . Suppose that  $Ngram^i$  contains a total of 50 words and the word “bad” appears 4 times, then the outcome of TF is  $\frac{4}{50} = 0.08$ . Similarly, if the document contains a total of 500 words and the word “bad” appears 50 times, then IDF is computed by  $\log\left(\frac{500}{50}\right) = 1$ . Finally, the TF-IDF value for the term “bad” in the  $i$ th review is computed using (2) as  $0.08 * 1 = 0.08$ . The features for the entire document are then extracted into vector  $NT(i)$ :

$$NT(i) = Ngram + TF - IDF \tag{3}$$

After that, the emoticons’ specific features are extracted from the reviews and are stored in the dataset for classification and further processing. Each review may not have the emoticons, therefore, for each review, we initialized the emoticons feature vector (EF) of size  $1 \times 2$  with zero value. Emoticons for each review are counted along with the sentiment label by using discrete probability distribution formula. The positive emoticon is represented by 1 and the negative emoticon is represented by  $-1$ .

For example, if the review contains four emoticons, out of which two are positive and two are negative then the outcome of emoticons features is represented as  $[2, -2]$ .

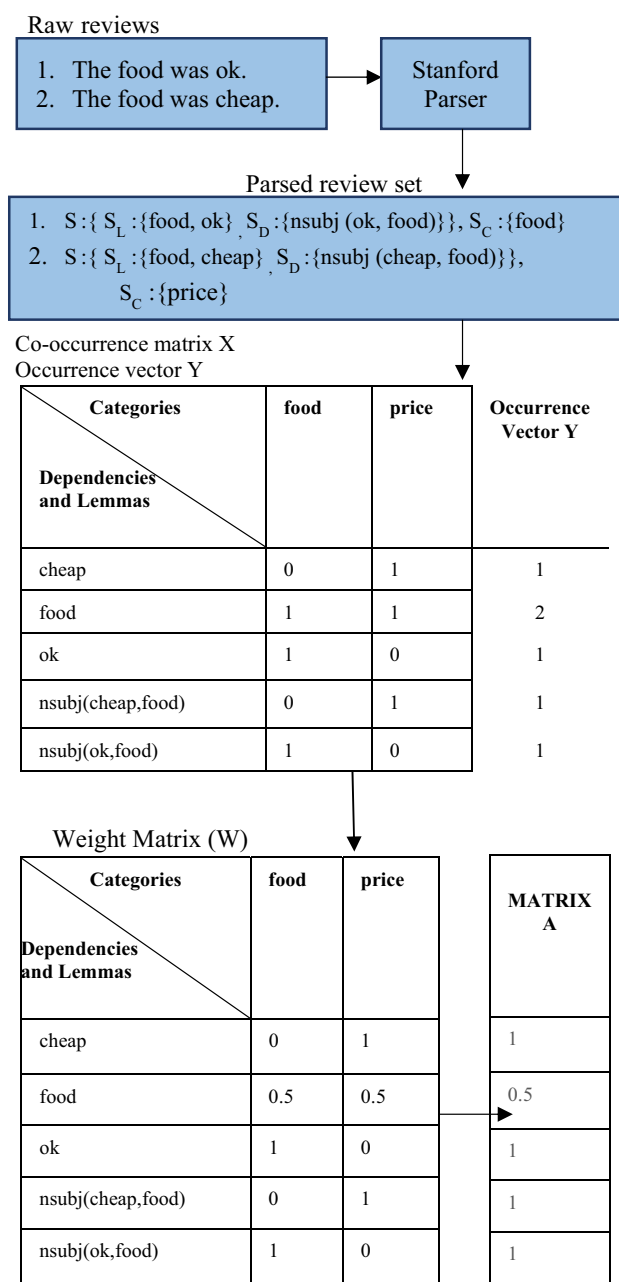


Fig. 2 Illustration of ARF

Missing of positive or negative or both emoticons lead to zero value for emoticons-related features for that review. The emoticons-specific features are then combined with NT features; therefore, the Eq. (3) can be re-written as:

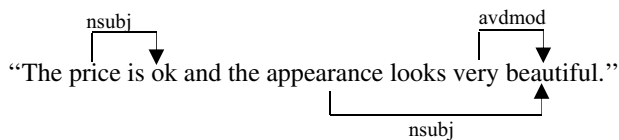
$$NTE(i) = [NT, EF] \tag{4}$$

### 3.2.2 ARF

After extracting the hybrid form of review-specific features, we employ the ARF extraction technique to training and testing dataset which uses the mechanism that is borrowed from Schouten et al. (2018) exclusively for improving the SA performances. The aim is to count: co-occurrence frequencies between lemmas and the annotated categories of a sentence, co-occurrences of lemmas and aspect types, co-occurrences between grammatical dependencies and aspect types. We redesign the approach for ARF in this section by estimating the weight matrix for each category set which is then converted into the aspect features for each pre-processed review in the input dataset. We excluded the mechanism of categories estimation in this work and focused on extracting the aspect terms in each review with their co-occurrence frequencies. It saves the significant processing time of applying the supervised classifier compared to the work proposed by Schouten et al. (2018).

Algorithm 2 shows the functionality of the proposed ARF approach. Let  $Q$  be the training set consisting of  $m$  number of raw online reviews. We aim to extract the categories and then estimate their co-occurrence frequencies against the lemmas and dependencies forms in each input review. The first step of the ARF algorithm is to discover the sets of lemmas, dependencies, and categories for each review. A lemma is the dictionary form of a word and dependencies depict the grammatical relationships that exist between the words of a sentence. In other words, a dependency relation is termed as an irregular binary association of a term named governor or head with another term named dependent or modifier.

An example below shows a sentence containing dependency relations, in which the words ‘price’ (head) and the word ‘ok’ (modifier) make a dependency relation called *nsubj* (also called nominal subject relation) while the word ‘very’ (head) and the word ‘beautiful’ (modifier) form another dependency relation called *avdmod* (also known as adverbial clause modifier).



#### Algorithm 2: Extraction of ARF

Input

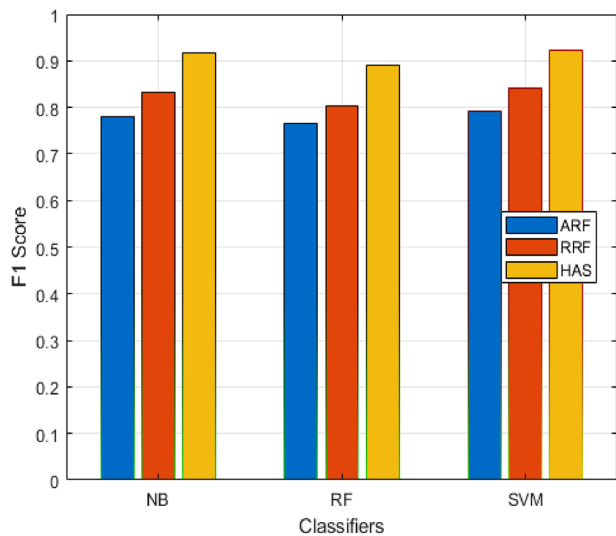
$Q$ : Training set

Output

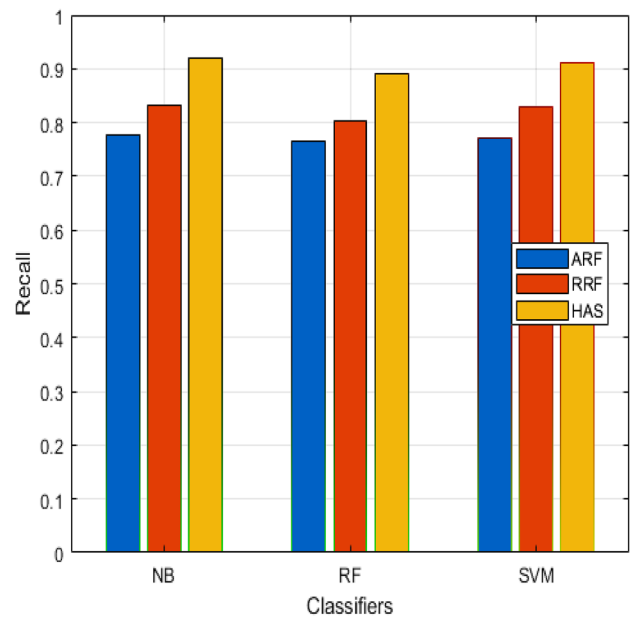
$A$ : Aspect specific features set

1. Initialize  $C, X, Y, W \leftarrow \emptyset$
2. For each review  $i = 1$  to  $\text{length}(Q)$
3.  $[s, s_c] \leftarrow \text{getLemmasDependencies}(Q(i))$
4. For each set  $k = 1$  to  $\text{length}(s)$
5. For each term  $j = 1$  to  $\text{length}(s(k))$
6. Count dependency/lemma occurrence  $j$
7. If  $(s(k, j) \neq Y)$
8.  $Y \leftarrow \text{add}(s(k, j))$
9.  $Y_j \leftarrow Y_j + 1$
10. End
11. For each category  $c = 1$  to  $\text{length}(s_c)$
12. Discover and add unique categories  $c$  in  $C$
13. If  $(s_c(c) \neq C)$
14.  $C \leftarrow \text{add}(s_c(c))$
15. End
16. Count co-occurrence  $(s_c(c), s(k, j))$  in  $X$
17. If  $(s_c(c), s(k, j)) \neq X$
18.  $X \leftarrow \text{add}(s_c(c), s(k, j))$
19.  $X_{c,j} \leftarrow X_{c,j} + 1$
20. End
21. End
22. End
23. End
24. Compute weight matrix for aspect features
25. For each co-occurrence pair  $x = 1$  to  $\text{length}(X)$
26.  $Y_j \leftarrow X(x, :)$
27. If  $(Y_j > 0)$
28.  $W_{x,j} \leftarrow (X_{x,j} / Y_j)$
29.  $A_i \leftarrow \max(W_{x,j})$
30. End
31. End

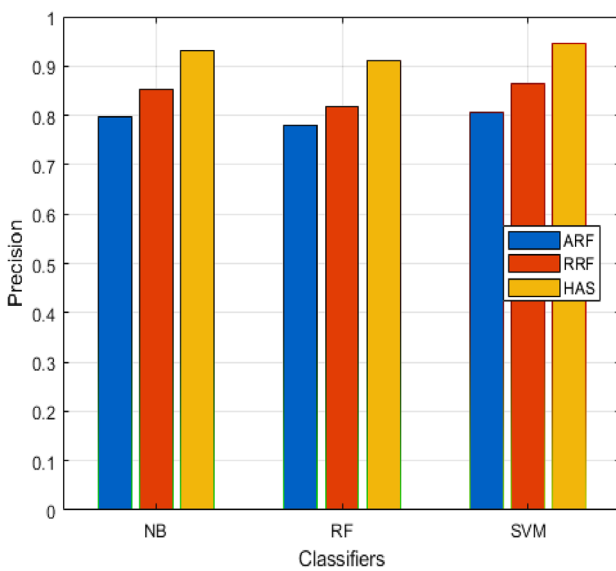




**Fig. 3** F1-score analysis of the proposed feature engineering techniques using different classifiers



**Fig. 5** Recall analysis of the proposed feature engineering techniques using different classifiers



**Fig. 4** Precision analysis of the proposed feature engineering techniques using different classifiers

**Table 2** Performance analysis of different features engineering techniques using different classifiers

	SVM	RF	NB	Average
<b>F1-Score</b>				
ARF	0.781	0.764	0.793	0.7799
RRF	0.832	0.803	0.842	0.8256
HAS	<b>0.922</b>	0.891	0.916	0.9096
<b>Precision</b>				
ARF	0.807	0.781	0.797	0.7950
RRF	0.864	0.819	0.853	0.8453
HAS	<b>0.946</b>	0.912	0.931	0.9296
<b>Recall</b>				
ARF	0.771	0.764	0.778	0.7710
RRF	0.828	0.803	0.833	0.8213
HAS	0.912	0.891	<b>0.919</b>	0.9073

Bold values indicate that among the classifiers SVM, RF, and NB, the F1-score and precision performances of the proposed feature engineering approaches using SVM are higher than the other two classifiers. The recall performance of NB is better than SVM and RF

Algorithm 2 is explained with the help of an example in Fig. 2.

The set  $S$  holds the list of lemmas and dependency forms. The  $S_C$  contains list of aspect categories for input reviews. As NLP step, each review is processed using POS tagger, dependency parser, and lemmatizer function of Stanford CoreNLP framework (Manning 2014). This yields a set of lemmas ( $S_L$ ) and dependency form sets ( $S_D$ ). The training dataset delivers the annotated categories of every sentence  $s$ , which is represented by  $S_C$ .

The next step is to count and add all the unique occurrences of lemmas or dependency forms into vector  $Y$ . Similarly, all aspect categories related to input review are discovered and added into vector  $C$ . After discovering the lemma/form dependency and unique categories, we store the co-occurrence frequency in vector  $X$ . In addition, the occurrence frequencies of all the dependency forms and lemmas of the corresponding review are recorded in vector  $Y$ . These

three vectors (C, X and Y) are generated for the training dataset only.

After receiving occurrence and co-occurrence frequencies in vectors Y and X respectively, we compute the weight matrix for each co-occurrence entry  $X_{x,j}$  with occurrence frequency  $Y_j$  and store it into the vector W. The weight frequency value ( $W_{x,j} \leftarrow (X_{x,j}/Y_j)$ ) for each pair in X is computed only if the corresponding co-occurrence frequency is greater than 0. It overcomes the problem of discovering the optimal threshold for any dataset. From the weight matrix W, we finally estimate the aspect-specific features by taking the maximum co-occurrence value for each pair of W into vector A. In this way, we satisfy our aim of extracting the aspect-related features without high computation requirements using machine learning techniques.

### 3.3 Classifier

As shown in Fig. 1, the hybrid representation of features has been performed by combining RRF and ARF into vector HAS. The name HAS, therefore, justified the hybrid mechanism of investigating the sentiments of end-users. As mentioned earlier, three machine learning techniques: SVM, NB, and RF are used to perform sentiment classification in the proposed work.

**SVM:** SVM is a famous classification method that is used to locate a hyperplane in an N-dimensional space that clearly classifies the data points. There are a lot of feasible hyperplanes that could be chosen to distinct the two classes of data points. Our goal is to discover a plane having the maximum distance between the data points of both the classes. The enhanced margin distance offers some augmenting which helps to classify the future data points with more confidence.

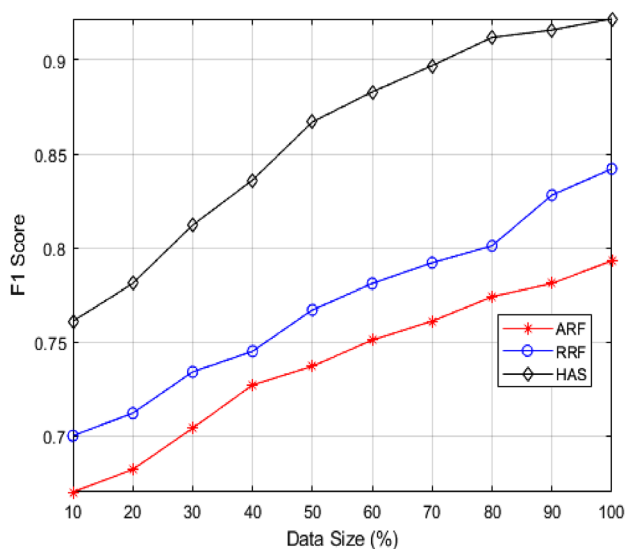


Fig. 6 F1-score analysis with varying data size

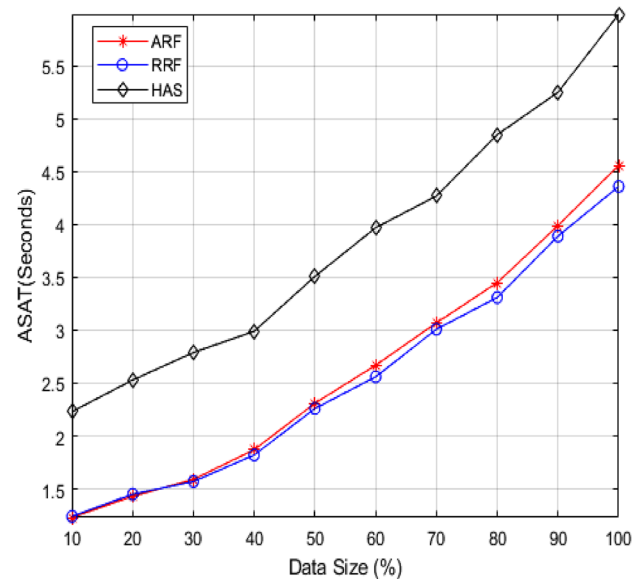


Fig. 7 ASAT analysis with varying data size

**RF:** In this ensemble classifier algorithm, several decision trees are constructed at training time and the class which is the mode of the classes output by each tree is delivered. RF is known for classifying huge amount of data with accuracy as several classifiers are engendered from smaller subsets of the input data and then their distinct results are accumulated based on a polling method to produce the anticipated output of the input dataset. To assess the error rate, we divided the dataset into training and testing parts. We used the training dataset to construct the forest and testing dataset to compute the error rate.

**NB:** We used Multinomial Naive Bayes Classifier as it is effective for classification with discrete features. This probabilistic classifier employs the Bayes probability theorem with the supposition of no relationship between different features for the prediction of unknown classes, such as the probability that a specified record or data point fits in that class. The most probable class is demarcated as the one having the maximum probability.

From the study of the existing literature, we noticed that all these classifiers have several applications in text classification and have shown better performances due to their accuracy and simplicity, which encourages their use in the proposed work. For sentiment classification, the test feature vector along with the training feature vector and associated labels are given as input to each of the above-mentioned classifiers.

## 4 Experimental results

For experimental analysis of the proposed model, we used the MATLAB tool under Windows 10 OS with an I3 processor and 4 GB RAM. We have used three datasets to investigate

**Table 3** Comparative analysis of HAS model using SemEval-2014 restaurant review dataset with state-of-the-art techniques

SA/ABSA technique	Precision	Recall	F1-score
SABSA	0.844	0.831	0.838
SentiVec	0.862	0.842	0.854
TF-IDF + N-gram + SVM	0.851	0.838	0.845
SEML	0.841	0.824	0.833
MTMVN	0.793	0.773	0.785
HAS	<b>0.946</b>	<b>0.912</b>	<b>0.922</b>

Bold values indicate the proposed HAS model shows the improved performances for all the three datasets for the parameters: F1-score, precision, and recall

the performance of the proposed model with state-of-the-art methods. SemEval-2014 restaurant reviews dataset (Maria et al. 2014) is used to evaluate the HAS method. The dataset consists of 3000 training reviews, and 800 test reviews. Each review in this dataset has one or more annotated aspect terms. Sentiment140 dataset (Go et al. 2009) is the Twitter dataset that consists of 1.6 million annotated tweets (0 = negative, 2 = neutral, and 4 = positive). STS-Gold (Saif et al. 2013) dataset consists of a total of 2026 tweets with their IDs and polarities. We have divided SemEval-2014 restaurant

**Table 4** Comparative analysis of HAS model using sentiment140 dataset with state-of-the-art techniques

SA/ABSA technique	Precision	Recall	F1-score
SABSA	0.858	0.839	0.8485
SentiVec	0.877	0.858	0.8675
TF-IDF + N-gram + SVM	0.866	0.846	0.856
SEML	0.854	0.837	0.8455
MTMVN	0.817	0.789	0.803
HAS	<b>0.961</b>	<b>0.938</b>	<b>0.9495</b>

Bold values indicate the proposed HAS model shows the improved performances for all the three datasets for the parameters: F1-score, precision, and recall

**Table 5** Comparative analysis of HAS model using STS-gold dataset with state-of-the-art techniques

SA/ABSA Technique	Precision	Recall	F1-score
SABSA	0.791	0.784	0.7875
SentiVec	0.845	0.837	0.841
TF-IDF + N-gram + SVM	0.831	0.817	0.824
SEML	0.826	0.809	0.8175
MTMVN	0.776	0.758	0.767
HAS	<b>0.927</b>	<b>0.899</b>	<b>0.913</b>

Bold values indicate the proposed HAS model shows the improved performances for all the three datasets for the parameters: F1-score, precision, and recall

reviews dataset, Sentiment140 and STS-Gold datasets into 70% training and 30% testing datasets. We first compared the performance of HAS with the proposed RRF and ARF methods using different classifiers to justify the efficiency of the hybrid feature engineering approach using SemEval-2014 restaurant dataset. Furthermore, we compared the HAS with different state-of-the-art techniques using all three datasets. The performances are compared using the parameters: F1-score, precision, recall, and Average Sentiment Analysis Time (ASAT). The F1-score is computed as follows:

$$F1 - score = \frac{2 \times P \times R}{P + R} \quad (5)$$

where P stands for precision and R stands for Recall which are computed as:

$$P = \frac{TP}{FP + TP} \quad (6)$$

$$R = \frac{TP}{FN + TP} \quad (7)$$

where FP represents False Positive, TP represents True Positive, and FN represents False Negative of sentiment classification. The parameter ASAT is related to the computational time, i.e., the average processing time for sentiment classification. We executed 50 instances of each technique for SA classification to estimate the ASAT parameter.

#### 4.1 Features investigation

This section presents the performance evaluation of RRF, ARF, and HAS using different machine learning classifiers. Figures 3, 4, 5 demonstrate the performances of F1-score, precision, and recall, and their associated readings are shown in Table 2.

From the analysis of F1-score, precision, and recall, it can be observed that for all the three classifiers, the hybrid feature engineering mechanism of the HAS model produced significantly improved results compared to ARF and RRF for sentiment classification. The key reason behind this performance improvement is that the HAS approach builds the feature vector that incorporates the challenges related to sarcasm, ambiguity, negation, emotions, feelings, and opinions about restaurants and their aspects. In RRF method, features are extracted using n-gram + TF-IDF + emoticons approach. It effectively builds the hybrid feature vector that expresses the frequencies of terms that appeared in each review along with emoticon-specific features. On the other side, ARF methodology uses the co-occurrence frequency technique to handle sarcasm and ambiguous terms to further reduce the SA errors. Among the RRF and ARF, the RRF produced better results as it effectively addresses opinions, emotions, and negations.

Among the classifiers SVM, RF, and NB, the F1-score and precision performances of the proposed feature engineering approaches using SVM are higher than the other two classifiers (refer to Table 2). The recall performance of NB is better than SVM and RF. The RF classifier shows the worst performance for each feature extraction technique as it ensembles the performances of different decision tree classifiers for class prediction. Overall SVM classifier produced better SA results than NB and RF as it determines the optimal boundary among the different sentiment classes. Table 2 also shows the average outcome of each feature engineering technique proposed in this paper.

Figures 6 and 7 demonstrate the performances of F1-score and ASAT with the varying training dataset size. The stratified sampling mechanism has been utilized to divide the training dataset into smaller groups of different sizes. As shown in Figs. 6 and 7, for each iteration there is a 10% increase in training data. We investigated the F1-score and ASAT performances of different feature engineering techniques designed in this paper using an SVM classifier. The F1-score performance of the HAS method is better than ARF and RRF techniques. With the increasing data size, the F1-score performance also increases because of the availability of a large amount of labeled data for accurate prediction. The ASAT analysis shows that HAS modeled requires higher computational time as compared to ARF and RRF feature engineering techniques. However, it is acceptable because of the significant performance improvement that the HAS model has shown for sentiment analysis.

## 4.2 State-of-the-art investigations

To analyze the quantitative representation of the HAS model, we examined the proposed HAS method against some of the following state-of-the-art recent methods using the SemEval-2014, Sentiment140, and STS-Gold datasets.

Supervised ABSA (SABSA) (Schouten et al. 2018): This method uses the SemEval-2014 dataset for the evaluation of their proposed supervised machine learning-based aspect category prediction using the co-occurrence technique for SA. The SABSA technique is closely related to our approach of ARF, therefore, we selected it for comparative study with the proposed model.

SentiVec (Zhu et al. 2021): This recently proposed technique is using the hybrid approach of supervised and unsupervised machine learning technique to perform word embedding for SA.

TF-IDF + N-gram + SVM (Ayyub et al. 2020): It is also another recent approach that we selected for the performance analysis. The hybrid feature formation proposed in this paper is closely related to our RRF extraction approach, except for the inclusion of emoticons-specific features in RRF.

SEML (Li et al. 2020): It is an ABSA technique that performs aspect mining and aspect sentiment classification using a semi-supervised classifier. They also used the SemEval-2014 dataset for performance analysis.

MTMVN (Bie and Yang 2021): It is the most recently proposed mechanism for ABSA, therefore it is selected for comparative analysis with the proposed approach. The authors used the deep learning classifier after extracting the aspect-specific features.

Tables 3, 4, and 5 demonstrate the comparative analysis of all the above methods in terms of F1-score, precision, and recall parameters using SemEval-2014, Sentiment140, and STS-Gold datasets. It can be observed that the proposed HAS model shows the improved performances for all the three datasets for the parameters: F1-score, precision, and recall. The reason behind the improved performance is its ability to build the hybrid feature vector using different techniques robustly and efficiently. The ABSA methods: SABSA, SEML, and MTMVN show poor performances compared to the methods SentiVec and TF-IDF + N-gram + SVM due to the lack of negation handling and extraction of review-specific features. The accuracy of sentiment analysis is found to be low for the STS-Gold dataset and is found to be higher for the Sentiment 140 dataset among other datasets. It is due to the reason that the more the number of training reviews more efficient the sentiment analysis would be. The HAS model proposed in this work is focusing on presenting the integrated solution to bring the benefits of both review specific and aspect-specific features for performance improvement.

## 5 Conclusion and future work

This paper addresses the problem of accurate sentiment analysis from the perspective of the CRS model for business owners. We highlighted the challenges of sentiment analysis for the raw online reviews and summarized different research motivations and solutions for them. The HAS model proposed in this paper performs efficient, reliable, and robust sentiment analysis through the hybrid analysis of sentiments. The hybrid feature engineering approach proposed in this paper builds more meaningful features for the input pre-processed reviews. The result analysis shows that the hybrid approach enhances the sentiment analysis performances significantly compared to individual techniques. The performance of HAS model has also been compared with state-of-the-art methods in terms of F1-score, precision, and recall parameters. The precision rate, recall rate and F1-score are found to improve by 10.78%, 9.04%, 9.1% respectively. The various future directions for the HAS model are (1) extension of the HAS model for customer review summarization using the sentiment classification results, (2) investigation of the HAS model performance

using different datasets, and (3) use of deep learning classifiers for performance improvement.

**Funding** No funds, grants, or other support was received.

**Availability of data and material** Data will be made available on reasonable request.

## Declarations

**Conflicts of interest** The authors have no conflicts of interest to declare that are relevant to the content of this article.

**Ethics approval** Not Applicable.

**Consent to participate** Not Applicable.

**Consent for publication** Not Applicable.

## References

- Alamanda MS (2020) Aspect-based sentiment analysis search engine for social media data. *CSI Trans ICT* 8(2):193–197. <https://doi.org/10.1007/s40012-020-00295-3>
- Alqaryouti O, Siyam N, Monem AA, Shaalan K (2019) Aspect-based sentiment analysis using smart government review data. *Appl Comput Inf* 2019:1–12
- Ayyub K, Iqbal S, Munir EU, Nisar MW, Abbasi M (2020) Exploring diverse features for sentiment quantification using machine learning algorithms. *IEEE Access* 8:142819–142831. <https://doi.org/10.1109/access.2020.3011202>
- Bafna K, Toshniwal D (2013) Feature based summarization of customers' reviews of online products. *Procedia Comput Sci* 22:142–151. <https://doi.org/10.1016/j.procs.2013.09.090>
- Baxter GJ, Connolly TM (2014) Using Web 2.0 tools to support the theoretical constructs of organisational learning. In: *International joint conference SOCO'13-CISIS'13-ICEUTE'13*, Springer, vol 239, pp 679–688
- Bie Y, Yang Y (2021) A multitask multiview neural network for end-to-end aspect-based sentiment analysis. *Big Data Min Anal* 4(3):195–207. <https://doi.org/10.26599/bdma.2021.9020003>
- Chakraborty S, Goyal P, Mukherjee A (2020) Aspect-based sentiment analysis of scientific reviews. In: *Proc. of the ACM/IEEE joint conference on digital libraries*, pp 207–216
- Datta S, Chakrabarti S (2021) Aspect based sentiment analysis for demonetization tweets by optimized recurrent neural network using fire fly-oriented multi-verse optimizer. *Sadhana* 46(2):1–23. <https://doi.org/10.1007/s12046-021-01608-1>
- de Kok S, Punt L, van den Puttelaar R, Ranta K, Schouten K, Frasinca F (2018) Review-aggregated aspect-based sentiment analysis with ontology features. *Progr Artif Intell* 7(4):295–306. <https://doi.org/10.1007/s13748-018-0163-7>
- Go A, Bhayani R, Huang L (2009) Twitter sentiment classification using distant supervision. *Processing* 2009:1–6
- Hanni AR, Patil MM, Patil PM (2016) Summarization of customer reviews for a product on a website using natural language processing. In: *2016 International conference on advances in computing, communications and informatics (ICACCI)*, pp 2280–2285
- Hao Y, Mu T, Hong R, Wang M, Liu X, Goulermas JY (2020) Cross-domain sentiment encoding through stochastic word embedding. *IEEE Trans Knowl Data Eng* 32(10):1909–1922. <https://doi.org/10.1109/tkde.2019.2913379>
- Hitkul S, Shahid S, Singhal D, Mahata P, Kumarguru RR, Shah (2020) Aspect-based sentiment analysis of financial headlines and microblogs. In: *Deep learning-based approaches for sentiment analysis*, Springer, pp 111–137
- Hossain MS, Rahman MR, Arefin MS (2020) Aspect based sentiment classification and contradiction analysis of product reviews. In: *Proc of the international conference on computer networks, big data and IoT (ICCBI-2019)*, vol 49, pp 631–644
- Hussein MEDMD (2018) A survey on sentiment analysis challenges. *J King Saud Univ Eng Sci* 30(4):330–338
- Iqbal F et al (2019) A hybrid framework for sentiment analysis using genetic algorithm based feature reduction. *IEEE Access* 7:14637–14652
- Khan A, Gul MA, Zareei M, Biswal RR, Zeb A, Naeem M, Saeed Y, Salim N (2020) Movie review summarization using supervised learning and graph-based ranking algorithm. *Comput Intell Neurosci* 2020:1–14. <https://doi.org/10.1155/2020/7526580>
- Khan J, Jeong BS (2016) Summarizing customer review based on product feature and opinion. In: *2016 international conference on machine learning and cybernetics (ICMLC)*, pp 158–165
- Kumar R, Pannu HS, Malhi AK (2020) Aspect-based sentiment analysis using deep networks and stochastic optimization. *Neural Comput Appl* 32(8):3221–3235. <https://doi.org/10.1007/s00521-019-04105-z>
- Li N, Chow CY, Zhang JD (2020) SEML: a semi-supervised multi-task learning framework for aspect-based sentiment analysis. *IEEE Access* 8:189287–189297
- Ligthart A, Catal C, Tekinerdogan B (2021) Systematic reviews in sentiment analysis: a tertiary study. *Artif Intell Rev* 2021:1–58. <https://doi.org/10.1007/s10462-021-09973-3>
- Liu CL, Hsaio WH, Lee CH, Lu GC, Jou E (2012a) Movie Rating and review summarization in mobile environment. *IEEE Trans Syst Man Cybern Part C Appl Rev* 42(3):397–407. <https://doi.org/10.1109/tsmcc.2011.2136334>
- Liu J, Seneff S, Zue V (2012b) Harvesting and summarizing user-generated content for advanced speech-based HCI. *IEEE J Sel Top Signal Process* 6(8):982–992. <https://doi.org/10.1109/jstsp.2012.2229690>
- Lu Q, Zhu Z, Zhang G, Kang S, Liu P (2021) Aspect-gated graph convolutional networks for aspect-based sentiment analysis. *Appl Intell* 51(7):4408–4419. <https://doi.org/10.1007/s10489-020-02095-3>
- Mahajan HB, Badarla A (2018) Application of internet of things for smart precision farming: solutions and challenges. *Int J Adv Sci Technol* 2018:37–45
- Mahajan HB, Badarla A (2019) Experimental analysis of recent clustering algorithms for wireless sensor network: application of IoT based smart precision farming. *J Adv Res Dyn Control Syst* 11:9. <https://doi.org/10.5373/JARDCS/V11I9/20193162>
- Mahajan HB, Badarla A (2021) Cross-Layer protocol for WSN-Assisted Iot smart farming applications using nature inspired algorithm. *Wirel Pers Commun* 121:3125–3149. <https://doi.org/10.1007/s11277-021-08866-6>
- Mahajan HB, Badarla A, Junnarkar AA (2021) CL-IoT: cross-layer Internet of Things protocol for intelligent manufacturing of smart farming. *J Ambient Intell Humaniz Comput* 12(7):7777–7791. <https://doi.org/10.1007/s12652-020-02502-0>
- Manning CD (2014) The Stanford Core NLP natural language processing toolkit. In: *Proc 52nd annu meeting assoc*, pp 55–60
- Maria P, Dimitrios G, John P, Harris P, Ion A, Suresh M (2014) SemEval-2014 Task 4: Aspect based sentiment analysis. In: *Proc. 8th Int. workshop semantic eval. (SemEval)*, pp 27–35

- Moussa ME, Mohamed EH, Haggag MH (2018) A survey on opinion summarization techniques for social media. *Future Comput Inf J* 3(1):82–109. <https://doi.org/10.1016/j.fcij.2017.12.002>
- Munuswamy S, Saranya MS, Ganapathy S, Muthurajkumar S, Kannan A (2021) Sentiment analysis techniques for social media-based recommendation systems. *Natl Acad Sci Lett* 44(3):281–287. <https://doi.org/10.1007/s40009-020-01007-w>
- Nandal N, Tanwar R, Pruthi J (2020) Machine learning based aspect level sentiment analysis for Amazon products. *Spat Inf Res* 28(5):601–607. <https://doi.org/10.1007/s41324-020-00320-2>
- Naresh A, Krishna PV (2021) An efficient approach for sentiment analysis using machine learning algorithm. *Evol Intel* 14(2):725–731. <https://doi.org/10.1007/s12065-020-00429-1>
- Oyebode O, Alqahtani F, Orji R (2020) Using machine learning and thematic analysis methods to evaluate mental health apps based on user reviews. *IEEE Access* 8:111141–111158. <https://doi.org/10.1109/access.2020.3002176>
- Pannala NU, Nawarathna CP, Jayakody JTK, Rupasinghe L, Krishnadeva K (2016) Supervised learning based approach to aspect based sentiment analysis. In: 2016 IEEE international conference on computer and information technology (CIT), pp 662–666
- Parlar T, Özel SA, Song F (2018) QER: a new feature selection method for sentiment analysis. *HCIS* 8(1):1–19. <https://doi.org/10.1186/s13673-018-0135-8>
- Prathi JK, Raparathi PK, Gopalachari MV (2020) Real-time aspect-based sentiment analysis on consumer reviews. *Data Eng Commun Technol Springer* 1079:801–810
- Ramírez-Tinoco FJ, Alor-Hernández G, Sánchez-Cervantes JL, Olivares-Zepahua BA, Rodríguez-Mazahua L (2017) A brief review on the use of sentiment analysis approaches in social networks. In: International conference on software process improvement, pp 263–273
- Saif H, Fernández M, He Y, Alani H (2013) Evaluation datasets for twitter sentiment analysis: a survey and a new dataset, the STS-Gold. *ESSEM@AI\*IA*
- Schouten K, Frasincar F (2016) Survey on aspect-level sentiment analysis. *IEEE Trans Knowl Data Eng* 28(3):813–830. <https://doi.org/10.1109/tkde.2015.2485209>
- Schouten K, van der Weijde O, Frasincar F, Dekker R (2018) Supervised and unsupervised aspect category detection for sentiment analysis with co-occurrence data. *IEEE Trans Cybern* 48(4):1263–1275. <https://doi.org/10.1109/tycb.2017.2688801>
- Shahana PH, Omman B (2015) Evaluation of features on sentimental analysis. *Procedia Comput Sci* 46:1585–1592. <https://doi.org/10.1016/j.procs.2015.02.088>
- Shams M, Khoshavi N, Baraani-Dastjerdi A (2020) LISA: language-independent method for aspect-based sentiment analysis. *IEEE Access* 8:31034–31044. <https://doi.org/10.1109/access.2020.2973587>
- Singh J, Singh G, Singh R (2016) A review of sentiment analysis techniques for opinionated web text. *CSIT* 4:241–247
- Singh M, Jakhar AK, Pandey S (2021) Sentiment analysis on the impact of coronavirus in social life using the BERT model. *Soc Netw Anal Min* 11(1):11–33. <https://doi.org/10.1007/s13278-021-00737-z>
- Trilla A, Alias F (2013) Sentence-based sentiment analysis for expressive text-to-speech. *IEEE Trans Audio Speech Lang Process* 21(2):223–233. <https://doi.org/10.1109/tasl.2012.2217129>
- Vila J, Ribeiro-Soriano D (2014) An overview of Web 2.0 social capital: a cross-cultural approach. *Serv Business* 8(3):399–404. <https://doi.org/10.1007/s11628-014-0245-y>
- Wang D (2021) Coarse alignment of topic and sentiment: a unified model for cross-lingual sentiment classification. *IEEE Trans Neural Netw Learn Syst* 32:736–747
- Wang L, Niu J, Yu S (2020) SentiDiff: combining textual information and sentiment diffusion patterns for twitter sentiment analysis. *IEEE Trans Knowl Data Eng* 32(10):2026–2039. <https://doi.org/10.1109/tkde.2019.2913641>
- Zhu L, Li W, Shi Y, Guo K (2021) SentiVec: learning sentiment-context vector via kernel optimization function for sentiment analysis. *IEEE Trans Neural Netw Learn Syst* 32(6):2561–2572. <https://doi.org/10.1109/tnnls.2020.3006531>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.