Journal of
Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7652270

# A HYBRID CNN BASED MULTIPLE WORD EMBEDDING METHODS FOR ASPECT SENTIMENT CLASSIFICATION

**Monika Agrawal[1] and Nageswara Rao Moparthi[2]**

[1]Research Scholar, Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India. monika.agrawal1986@gmail.com

ORCID: 0000-0002-9705-6849

[2] Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India.

mnrphd@gmail.com

ORCID:0000-0001-6406-4554

**Abstract**

As the size of the sentiment data increases, it is difficult to predict the word polarity of an aspect term in a sentence in real-time applications. Most of the conventional models have a dependency of one or more aspect terms for classification problem. However, these models require high computational time for word embedding process and aspect level sentiment classification. Also, these models have high true negative rate and misclassification rate on large aspect databases. In this work, a hybrid multiple word embedding methods and classification approach are implemented in the CNN framework on the large databases. Experimental results show that the proposed aspect level word embedding based classification approach has better efficiency in terms of true positive rate , runtime and accuracy than the conventional aspect level classification approaches.

**Keywords: Aspect sentiment, filtering, classification , polarity.**

## 1.Introduction

Traditional classification of topical text attempts to classify a document by subject and often relies upon word frequencies for decision making. For example, if there is a high frequency of a document with certain key words such as "zoom lens," it is more likely that it concerns the camera. Since sentiment mining deals with feelings concerning topics and not subjects themselves, the approaches to classification are necessarily different from those for the topical classification of text. Early work has focused on using the important indicators of feeling adjectives like 'good' and 'bad,' as well as adverbs like 'terrifically' and 'hatefully.' Later research suggests, however, that other parts of speech, like verb and even substantives, might also provide valuable indicators of feeling. The second difference between the classification of feelings and the topical text is how these indicators are combined to deduce the category of polarity (positive or negative) of a document. The use of indicator frequencies (as in the class of topical text) had a negative impact. They therefore proposed a presence-based scheme (whether an indicator is included in the document or not). In addition, the present approach appears to be reaffirming the subjective indicators less frequently than current indicators.

Third, feeling and subjectivity rely heavily on domain. The reason is that we tend to describe a wide range of things using the same words. For instance, a film can be defined as 'big,' thus a product can be described. There are, however, examples of the same word having different significances in different fields. The word "unpredictable" is one such example. Since user generated documents are written in a free and informal way and only with a liberal use of correct orthography or grammar, it is ultimately difficult to extract information, represent and finally make predictions based on it. Human behaviour is influenced by the various opinions generated in society. Public opinion influences our decisions most often. Business and establishments always need to collect the opinion of the society which they try to obtain using customer feedback forms and questionnaires or surveys which help them to be aware of shortcomings if any, and to use suggestions to improve quality. It works in the same way for customers as well as the opinions of other customers about a particular product can come in handy when deciding to buy a product. Before the introduction of Sentiment Analysis, it was a herculean task for the companies to collect feedback from the customers. They had to physically approach them, individually for this. With the blistering advancement of online networking groups, the people as well as organizations make use of this platform to understand the public sentiment. If a person needs to purchase a produce, he/she is not limited anymore asking his friends or colleagues for their opinion, exhaustive reviews and feedbacks are available on the internet about the product. Organizations no longer need to conduct surveys since an abundance of such public opinions regarding a product is available on the Web. However, finding and monitoring such opinions present in the Internet and filtering the required information is a formidable task because of the huge amount of data that is available. It is not an easy task for an average reader to identify relevant sites and extract the opinion in them. In such difficult situations, Sentimental Analysis can play a big role in helping the user. A unique feature of the property is called ‘Aspect category’ of a property. ‘Aspect expression’ is a real term or expression that comes in the text representing an aspect category. The method of aligning aspect words into aspect categories is called ‘Aspect categorization’. Aspect Expressions are generally noun and noun phrases but may also be verbs, verb phrases, adjectives and adverbs. Aspect Expressions which are nouns and noun phrases are called ‘Explicit Aspect Expressions’. In the sentence, “The sound clarity of this mobile phone is great” the noun phrase ‘sound clarity’ is an example for ‘Explicit Aspect Expression’. The ‘Aspect Expressions’ which are not nouns or noun phrases are known as ‘Implicit Aspect Expressions’. Sentiment Classification is framed as a dual–type taxonomical issue, positive and negative. Product reviews or movie reviews are usually used as the data to be tested. But in this study, a literary text has been used as the testing data and the class classifications are the different human emotions present in the literary texts instead of just the two class classifications - positive and negative. In the two-class classification, these two classes are determined using the ratings from 1 to 5. For example, if a review has a rating of 4 or 5, then it will be considered as positive review whereas if it is 1 or 2, then it is considered negative and 3 star reviews are considered as neutral reviews. In sentiment classification, sentimental words such as good, excellent, amazing, poor, bad and so on indicate various sentiments. Based on the number of these words, the sentiment is classified. Sentiment categorization of documents expressing the user’s feelings and views in various languages is called Cross-language sentiment classification. Large scale of research to build sentiment analysis systems has been done in English than other languages. Hence, the first motivation is to leverage the emotion detecting system in English language to construct good sentiment analysis system in different languages outside English. Many scholars have researched on this issue. Majority of the existing studies concentrates on emotion categorization at the text level and individuality and emotion categorization at the statement level. Aspect level classification is carried out in limited scale.

For topic level assignments, they used a pattern-based approach. They combined sentimental analysis with language translation, according to which a system was put forward to extract sentimental units from the text. You used the machine translating engine based on the transfer, and replaced the two-language patterns and patterns with feelings and lexicons of polarity. Using the machine transfer engine, a precision score of as high as 100% was achieved. The feeling units created by the machine translation engine were less informative than the naive structures of predicate argument. The three common types of review formats were examined. 1. Pro and Compatibility 2. Detailed review of the pros and cons 3. Free size. They were able to extract the product feature using the association rules mining if the review is either format 2 or format 3. If your review is in pros and cons format, you use a supervised mining method to extract the product feature. With WordNet, the polarity of the identified words was determined. The orientation of the sentence level was determined on the basis of the dominant direction. The method for sentence-level analysis has been developed. The first decisions were whether an expression was neutral or polar and then the polarity of the polar expressions was disambiguated. They show the concept that positive and negative words are prior polarity and contextual polarity. There is a positive prior polarity, for instance, in the word "motion," but if "no" follows, the negation occurs and a negative contextual Polarity of the word "motivation." They use the 5000-round algorithm "BoosTexter AdaBoost" for learning machines to dispel the contextual polarity. They suggested a propagation algorithm to detect opinions and related objectives. The algorithm begins with a seed word and identifies the related target, identifying other available opinion words and targets until new views and objectives can no longer be added. The Stanford POS tagging tool was used for the purposes of speech-based (POS) markup and the "Minipar" sentence parser was used to parse the phrase sentence. They saw words as adjectives and objectives as substantives or nouns.They proposed a sentimentally sensitive thesaurus for the expansion of the vector function in order to develop a binary class. They proposed a sentiment classification method if labelled data are not available for a destination domain, but certain labelled data is available for several other domains designated as source domains. Through effectively learning from several source domain areas, cross-domain feeling classification was achieved. In their research, they used unigrams and bigrams as lexical elements. POS tagging and lemmatization were used in order to recognize lexical components in source and objective domains from both sources and unlabelled domain reviews. After the lexical element is identified, the lexical elements co-occurring in the same review phrase are selected as features and the co-occurring elements are added to the feature list from other source dominants wherever the lexical element occurs. This can represent a particular lexical element as a vector with all possible characteristics that occur in conjunction with it. It was established that the extension of feature vectors using the sensitive sentimental thesaurus bridges a gap in the cross-domain sentiment classification between source and target areas. The SLDA was then expanded to include the Unification Model of Aspect and Sentiment (ASUM). This ASUM has the advantage that no labelled reviews are required. The use of this method is better than basic classifiers for reviews concerning electronic goods and restaurants and ASUM's classification performance. For each topical term and its context words, they created the highest spanning trees (MSTs). The MST knots are opinions or expressions and the edges between the knots are associated. The mutual information (CPI) was used for measuring the weight of the connection of a pair within the MST. The classification of sentence levels is based on the fact that if an MST is produced by more positive than negative polarity, the chances are greater and vice versa. Finally, the document level rating is performed using a log ratio decision function by comparing the MSTs rates of sensitive sentences. These models have been designed to develop optimal approaches for profound models without direct monitoring. At any time, the field of sentiment analysis is usually linked to a solitary area, e.g. film review or review [Pang and Lee, (2008)]. More recently, a lot has been done to improve techniques for sentimental analysis in the different areas of films and product review, prediction of election results, outbreak of disease, stock market, etc. The excessively extensive e-commerce presence of these customers helps in enhancing the experiences of the customer with effective examination and use of these valuable data. On the other hand, it can help buyers choose what to buy by extracting feelings from the contents and their points of view. Recurrent Neural Networks (RNN) were used for word analytics. One of the restrictions of CNN is that long-distance dependence cannot be captured. This limitation may be met by RNN based LSTM network, which offers sequential text-by-sentence modelling for these deep networks. Authors has been using multiple data types but there has been limited effort in the dimensional sentiment analysis[1]. During the Web 2.0 phase, the amount of unstructured customer data generated on e-commerce websites increased significantly. The opinion of customers is a valuable and important type of information that the e-commerce industry should not neglect or overlook. From the customer 's perspective, it is a common behaviour on the Internet to look at the opinions of other users before buying a product. According to [2][3], several surveys agencies claimed that 70-90% of customers who acquire decisions have access to thousands of opinions through online review of this e-commerce website. This feature shows that the proven sales drivers are customer reviews. The basic intention of customers is to find or buy best at the cheapest price, that products that meet their requirements in the price range come from a neutral nature by analysing other opinions which will definitely benefit them. These views are the voice of ordinary consumers, which differs considerably from advertising. From these viewpoints, the positive and negative aspects of the product can be easily understood. From the e-commerce perspective, its business action plan can improve significantly to increase profitability through consumer reviews / feedback. These reviews improve customer satisfaction and business intelligence. Products with millions of opinions are frequently found, so an analyse them all could be difficult for a customer. The customer is assisted in making decisions through sentiment analysis tools and the social media analyst platform. The analysis of sentiment is the computing study of the opinions, feelings, feelings, and attitudes of individuals. It focuses on the development of automated systems to analyse texts in the natural language to determine their feelings. Subjectivity, opinion, impact, attitude, orientation, feelings, emotions, and tone in the text are frequently used in a wide sense [4]. In sentiment analysis, most of the current work is focused on the task to determine the presence of feelings and their value, i.e. the classification of feelings in a positive or negative way. Existing work on the analysis of feelings can be classified as the problem of the classification of texts in different forms, such as document sentiment analysis, sentence-level sentiment analyses, aspect-based sentiment analysis and sentiment lexicon acquisitions. In general, three processing phases can be defined for the sentimental aspect-level analysis: detection, classification and aggregation[5]. While not every approach takes these three steps, it is important to examine the facets of emotions in action or in this exact order. The first step is to define the pair of feelings in the letter. The next move is to identify the goal-feeling pairs. The articulated opinion is categorized by predefined values, such as positive and negative values. Often the goal is often identified by a predefined number of aspects. Finally, the sensing values are merged to provide a succinct description of each dimension. The final design is based on the particular needs and specifications of the company[6].

As described in[7], the study of emotions was mainly analysed at three stages. Feelings are graded either at the level of the text, the level of the sentence or the level of the person or the element. Focusing on the first point means that the whole text is a single subject. Of example, this is not the case in other cases. Similarly, a second level emphasis is believed to contain only one subject in one paragraph. The same term also includes referencing different individuals or contrasting those emotions of thought. Calculated sentiment values are not related to objects (i.e. persons or facets of persons) mentioned in the text at both the level of the paper and the level of the paragraph. Likewise, emotions can also be measured via a full corpus of any random text (e.g. a corpus of microblog entries in which any entry is called a document). In the other hand, the aspect-level study of emotions aims at identifying sentiment-target pairs in a given text (i.e. it may vary from sentences or small texts to full businesses with multiple documents). The overall feeling would usually refer to the individual in the sense of the sentiment analysis, while the sentiment analysis in the facets of the individual under consideration should be correlated with the emotions. This allows for a more detailed study that makes use of more textual research information. This research would therefore concentrate on review at the level of the dimension and its numerous undertakings. There are not many supervised machine-learning approaches that are strictly machine-learning approaches for aspect detection. As the strength of supervised solutions rests in the features used, the features often consist of a particular system (e.g. frequency-based methods) for generating features that are more popular than mere bags of words or pieces of expression. Aspect identification in[8] is achieved in the form of a mark problem solved by the use of a linear sequence, a regular CRF, to process a whole set of terms (e.g. a sentence). This also takes into account the meaning of a term as the mark is given. Several characteristics are used to decide the correct mark of a word, including the real name, the portion of the vocabulary tag, whether there is a clear relationship of dependency between that name and the expression of feeling, whether that word is in the concrete phrase nearest to the expression of feeling, and whether that word is in the expression of feeling. In order to train the algorithm, the ground reality is used from a portion of the data collection used[9]. The following data sets span four domains: video, online resources, vehicles and cameras. There are no variations to each classification system and this sample does not vary from the classification system used. In [10] the WordNet synonym / antonym graph is used to encourage the recognized meaning of a few terms of the root. Just adjectives are used here as expressions of emotion. A emotion class from the produced feeling dictionary (i.e. positive or negative) is assigned to each sentence's adjective. If a word negation occurs within five words of the word thought, the polarity is inverted. The emotion class shall then be decided by a plurality vote for every paragraph.

In the section 2, different types of text pre-processing models, classification models and word embedding models are discussed along with the limitations. In the section 3, proposed framework is implemented on the training data. In this section, hybrid word embedding model, classification model are discussed using the CNN framework. In the section 4, experimental results are evaluated on the training data and statistical results are analysed. Finally, in the section 5, conclusion and future scope are presented.

## 2.Related works

Various methods have been used for abstraction, i.e. frequency [12], lexicon [13]. Initially, Hu & Liu researched the extraction of an aspect. They extracted aspects using the association rule for mining in order to identify the frequent element set[14]. The idea is to extract the nouns / noun phrases that are widely cited as a candidate in the data review. This method is quick, effective and quiet. However, while this approach is easy and quietly efficient, it has not been able to distinguish subtle features and low-frequency aspects. However, this approach has been established in several studies [15-17]. They used the PMI-score to test the extracted word. Using another frequency-based approach, they used the TFIDF scheme. [18] improved the frequency-based approach with an added pattern law to reduce noise. The downside of the frequency-based process has been strengthened by a lexicon-based approach. [19] introduced a comprehensive lexicon-based approach to resolving the previous problem and improved the frequency-based approach . Instead of looking at the present term on its own, this approach takes advantage of contextual knowledge and facts of other phrases and other studies, as well as certain linguistic patterns of natural language languages, to infer the direction of terms of opinion. No prior knowledge of a domain or user feedback is needed. This method has been shown to be capable of removing an element that has a low amount, since the hunt is based on a term that is nearest to the terms of opinion. This method has been extensively developed in subsequent studies[20]. The lexicon-based approach relies on the lexicon used for the view. Most work therefore introduces a semantic-based approach that is simpler and more accurate. Semantic-based approach used dependency relationships and rule-based patterns to derive the list of candidates for that dimension. Turney initially developed a semantic-based approach using the POStag pattern.[21] noted that sentiment can be measured and categorized either by machine learning techniques or by lexicon-based techniques. Tweeter is considered to be one of the most popular social networking facilities where millions of users share their ideas and opinions on multiple fields. Examples of such areas are: politics, goods, celebrities, etc. Sentiment analysis is used to interpret the resources of sentimental change in public attitudes, mining, and product reviews to address the polarity shift problem. The writers used different classification structures such as the Naïve Bayes classifier, the help vector machine, the decision tree, and others. [22] noted that online content in Arabic is limited and that the accuracy of existing methods is lower. The Arab opinion analysis is based on two main issues: unique Arabic issues and general linguistic issues. Arabic basic difficulties are exacerbated by Arab morphological uncertainty, limited resources, and dialects. General linguistic problems include polarity fuzziness, spam, quality review, and others. [23] stated that there are a variety of approaches in the text to detect feelings. Lexical resources; such as the dictionary of terms of opinion; are essential. SentiWordNet is a tool that provides opinion information on words derived from the WordNet database. Using semi-supervised learning methods, users can classify the reviews of products as positive or negative using SentiWordNet lexicon. [24] introduced a lexicon-based sentiment analysis algorithm and concentrated on an interpretation of the meaning of the tweeter in real-time. This is important to measure the strength of the emotion rather than the positive / negative mark. The authors stated that sentiment analysis was extended to specific, non-necessary realms of public opinion tracking and forecasting. [25] tested the impact of standardization, halting, and stop-word elimination on the efficiency of the Arabic sentiment analysis method. The program used Tweeter's Arabic tweets. The mood of the crawled tweets was evaluated in order to understand the public's attitude. The authors stated that the key task of sentiment analysis is to evaluate the writer's attitude to certain subjects. [26] proposed an introduction to the definition of emotion polarity in twitter posts. The method is based on the abstraction of the weighted node vector from the WordNet list. The weights are used in SentiWordNet to measure the final polarity calculation.

The main problem of these conventional models include

a) Difficult to handle large datasets.
b) Difficult to process high dimensional feature sets.
c) These models are independent of contextual key feature relationships.

## 3.Proposed Model

The classification strategy based on supervised learning involves the manual preparation of training classes at a much lower cost. An unsupervised learning approach, on the other hand, has to create its own training data, which greatly increases the cost. However, it is difficult for supervised learning techniques to classify texts with large-scale data due to the challenges of creating precise training data sets that require automatic text classification techniques. The process of Automatic Text Classification extracts, categorises and analyses vast data and represents specific data across multiple domains. The information accessed in today's Big Data era is not organised and has changed the classification strategy, methodology and technology involved. The Text Classification Methodology includes different stages of Feature Generation and Feature Selection. The documents are associated with relevant concepts representing the document 's data in the generation of features. First features are extracted in this process and the characteristics are reduced secondly. A set of valid features that undergo feature selection are generated after feature reduction, where the most distinctive characteristics representing the maximum number of documents are selected. Next in the process, the documents set are categorised by a classifier based on the chosen features. A Performance Analysis of the classification process is carried out to determine the efficiency of the applied methodologies. The processes for selecting features are based on three methods: wrapper, embedded and filtering methods. The Filter Method is independent of the learning process and relies on the classification training data. In this process, with the training data, the terms are independently evaluated and each term is given a score or rank. At the feature extraction, reduction and feature selection stage, the philtre methods can be applied. The techniques use a threshold value in which a certain degree of term reduction is achieved from the enormous number of terms generated. The function selection process selects the best features to represent the documents on the basis of the ranking or scores given to the features. Due to its ease of implementation, its ability to handle large numbers of texts and features and its scalability with less complexity, philtre techniques are widely used in the text classification process. Multi-Field Records, Word Merging, N-grams, Word Phrases and Other Properties are the term generator method used. Using the above methods, the generation of features is automated with techniques such as philtre and wrapper metrics and algorithms such as CITRE, GALA, ID2-of-3, M-of-N concepts, FICUS algorithm, etc. Removing rare and common words, stemming the merged words, etc. is the term reduction process. Stop Word Removal, Tokenization, and Stemming are the approaches used. By using algorithms such as the Porter algorithm with the above methods and techniques such as philtre and wrapper and metrics such as IG, CHI2, Odds Ratio, etc., dimensionality reduction is automated. Feature Selection is a process in which, according to the training examples provided, the selection of a feature subset or features that can best categorise the documents. The FS process classifies documents according to the label criteria required, such as (1) multi-labelling, (2) single-labelling and (3 ) non-labelling. Feature extraction performs an analysis of the documents and extracts document terms that mostly represent a document based on the training data provided. In relation to the overall classification strategy, the terms extracted that could be dominant predictive features in the classification features are obtained. In the process of extraction of features, we first determine what kind of words can be considered as a term. Based on an external knowledge repository, feature extraction associates the document with relevant concepts or terms. In this process, by studying the documents, the feature generator, similar to any text classifier, learns the classification model. The terms chosen are the most representative and most relevant words of a document that are taken as concepts that demonstrate the texts' properties. These concepts best demonstrate the content of the document, which may represent multiple categories across domains. The automatic feature or term extraction algorithms used are FRINGE, CITRE, GALA, ID2-of-3, FICUS, etc. Machine learning algorithms such as NB, centroid-based methods, etc. are used with standard metrics such as TF-IDF, IG, identity words, etc. in extracting features and handling a large number of concepts where a large number of texts have to be categorised. The large number of features generated by feature extraction are feature reduction in text classification to minimise the task. Eliminating hundreds of thousands of terms using the technique such as Stop word removal, Stemming and Statistical filtering, etc., is used to decrease the high dimensionality. "Stop word removal removes words based on a stop word list with little or no meaning, e.g. words such as," a," "the," "but "etc. Stemming is the elimination of words with a common root for a single function, such as words such as books, books, books, bookshelf, etc. Stemming algorithms can be used to decrease dimensionality, such as the Porter algorithm. Using statistics, statistical filtering finds words with greater implications. The methods used are MI, IG, CHI2 and so on. A score is given to a term that occurs frequently and infrequently in terms of a category and is based on conformity and non-conformity, where a high score is given to terms with most document representation.

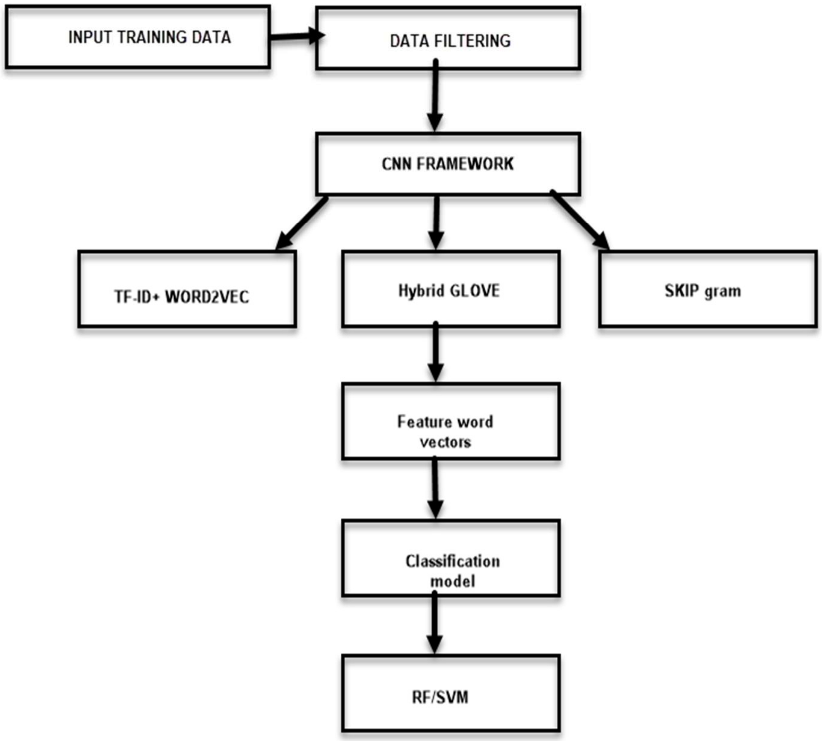



**Figure 1: Proposed CNN based aspect sentiment classification Framework**

Figure 1, describes the flowchart of the proposed model for aspect sentiment classification. A three-phase framework is described in the figure1. In the phase one, input data are pre-processed using the Stanford NLP library. In the text pre-processing, data tokenization, stemming and stop words removal are executed on the input training data. Filtered data of the phase one is given to the second phase for essential features extraction. In the second phase, a hybrid text embedded model is designed and implemented on the filtered data for aspect features identification. In this framework, different hybrid word embedding methods such as TF-ID word2vec , hybrid GLOVE and hybrid SKIP gram are enhanced to improve the feature selection process in the CNN framework. The features that are marked in the second phase are given to third phase for text classification process. In the third phase, a hybrid text classification model is implemented on the high dimensional feature space.

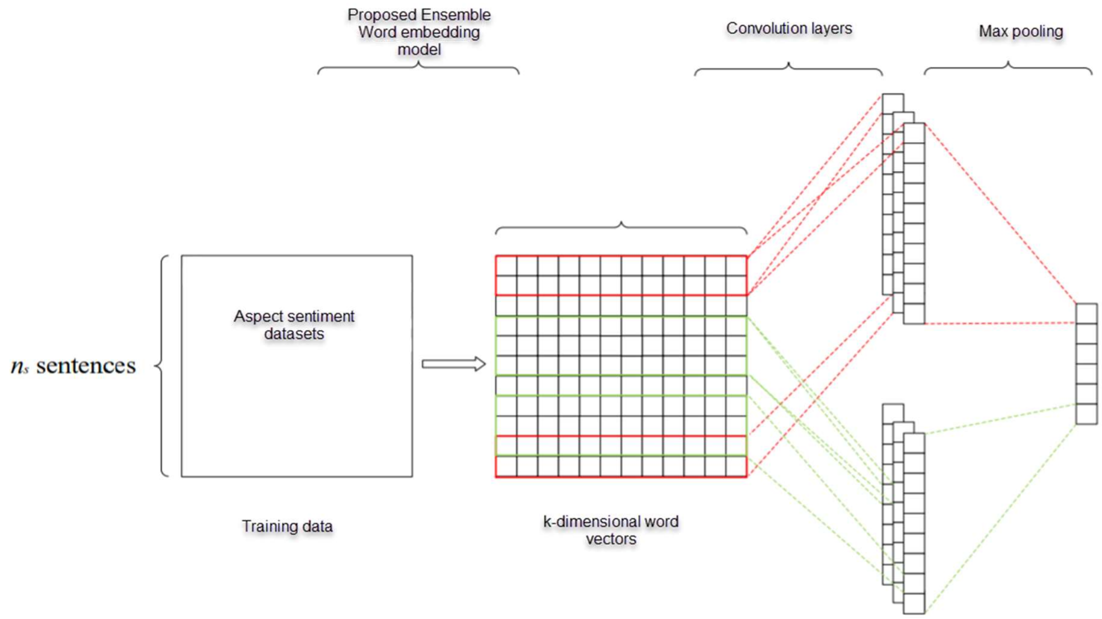



Figure 2: Proposed ensemble word embedding model in the CNN network .

Figure 2, illustrates the working procedure of the proposed ensemble word embedding model in the CNN network. In this framework, initially all the aspect related sentences are taken as input data for sentiment classification. This input data is given to ensemble word embedding model in order to filter the essential k dimensional word vectors. These k-dimensional word vectors are given to 'N' convolution layers to find the essential key features for MAX pooling operation. Finally, the vectors in the MAX pooling operation are given to proposed classification model for aspect sentiment classification.

**Proposed Ensemble Word Embedded Measures:**

**a) Hybrid TF-ID based Word2Vec Formula**

In the hybrid TF-ID based Word2Vec embedding measure, different aspect sentences and its related feature attributes are taken as input to compute the probability score. Here, the computed probabilistic score is used to find the selection of key feature aspect sentence and its related attributes in the CNN network. Let $w_i$ be the word in the aspect sentence and its related attributes are named as A.

1. The log estimated probability to find the key aspect word in the given dataset is computed as:

$$P_{\text{HTF-ID}} = \frac{1}{N}\sum_{i,j} tf_{i,j}.\log p(w_i \mid A_{j+i}) \quad \text{---(1)}$$

Where tf is the term frequency and p is the conditional probability of occurrence of word in the given aspect related features as shown in eq (1).

2. The hybrid word2vec measure of the entire aspect data is given in eq (2)

$$\text{HWord2vec}_{i,j} = \frac{|n|_{ij}}{prob(n_{ij}/d_j)} \times \log_2 \frac{|D|}{1+|prob\{\{w_i \in d_j\}/A_w\}|} \quad \text{--(2)}$$

3. The weighted hybrid word2vec of the aspect term is given in eq(3)

$$WE(HWord2vec_{i,j}, P_{HTF\text{-}ID}) = maxwordvote(\frac{|n|_{ij}}{prob(n_{ij}/d_j)} \times \log_2 \frac{|D|}{1+|prob\{\{w_i \in d_j\}/A_w\}|}, \frac{1}{N}\sum_{i,j} tf_{i,j}.\log p(w_i \mid A_{j+i})\} \text{ --(3)}$$

b) **Proposed Glove Word embedding method**

In the glove word embedding model, initially all the aspect sentences are tokenized to find the glove main and its contextual words. Glove method initially partition the aspect terms into main words and its contextual words based on the initialization parameters. The mathematical objective function used to find the contextual words to the main vectors are given as.

1. Define soft constraints for each word pair is computed using the eq(4)

$$C = CostFunction = b_i w_i^T w_j + b_j w_i^T w_j + \theta - (\log(X_{ij}) / \max\{\|w_i\|, \|w_j\|\} \text{ --(4)}$$

$$\eta = weight = f(X_{ij}) = \begin{cases} (\frac{X_{ij}}{x_{max}})^{\alpha} & \text{if } X_{ij} < XMAX \\ 1 & \text{otherwise} \end{cases}$$

where $\theta, \alpha$ are the scaling factors of the main and contextual word vectors .

2. The hybrid Glove vector model to define the cost function is given in eq(5)

$$J = \sum_{i=1}^{V}\sum_{j=1}^{V} \eta.(b_i w_i^T w_j + b_j w_i^T w_j + \theta - (\log(X_{ij})/\max\{\|w_i\|, \|w_j\|\}))^2 \text{ --5)}$$

**Hybrid Random forest model**

In this model, a hybrid the feature word selection measure is proposed in order to improve the classification accuracy in the traditional random forest approach. In this model, a new probabilistic based measure is proposed on the aspect sentiment words in order to select the feature for the classification problem. The optimized feature selection measure is given in eq(6)

$$PHD(W_i) = \max\{\log(P(A_+ \mid w_i), \log(P(A_- \mid w_i)\}.\sqrt{\sum_{i \in V_{X_j}} (\sqrt{Prob(A_+ \mid W_i)} - \sqrt{Prob(A_- \mid W_i)})^2} \text{ ----(6)}$$

## 4.Experimental results

Experimental results are simulated in java environment with third party libraries. In this work, java based deep learning framework is used to test the ensemble word embedding models. In this section, different statistical evaluation metrics such as features rank, features runtime, model accuracy and recall are tested on the training input aspect sentiment dataset. In the experimental results, aspect sentiment dataset is used to find the aspect sentiment classification for prediction process.

Table 1: Generated decision pattern for the given input aspect sentiment data.

| |
|---|
| example_id = 14_0 : 1 (0/0) [0/0] |
| example_id = 3302_0 : 1 (0/0) [0/0] |
| example_id = 3302_1 : 1 (0/0) [0/0] |
| example_id = 160_0 : 1 (0/0) [0/0] |
| example_id = 160_1 : 1 (0/0) [0/0] |
| example_id = 2427_0 : 1 (0/0) [0/0] |
| example_id = 2427_1 : 1 (0/0) [0/0] |
| example_id = 2427_2 : 1 (0/0) [0/0] |
| example_id = 2427_3 : 1 (0/0) [0/0] |
| example_id = 2404_0 : 1 (0/0) [0/0] |
| example_id = 116_0 : 1 (0/0) [0/0] |
| example_id = 116_1 : 1 (0/0) [0/0] |
| example_id = 116_2 : 1 (0/0) [0/0] |
| example_id = 2008_0 : 1 (0/0) [0/0] |
| example_id = 2008_1 : 1 (0/0) [0/0] |
| example_id = 2206_0 : 1 (0/0) [0/0] |
| example_id = 1010_0 : 1 (0/0) [0/0] |
| example_id = 89_0 : 1 (0/0) [0/0] |

Table 1, describes the decision patterns generated on the given input aspect training data. Table 1, illustrates the decision patterns of the input aspect training data using the proposed classification model. These patterns are generated based on the proposed feature ranking measure on the input data.

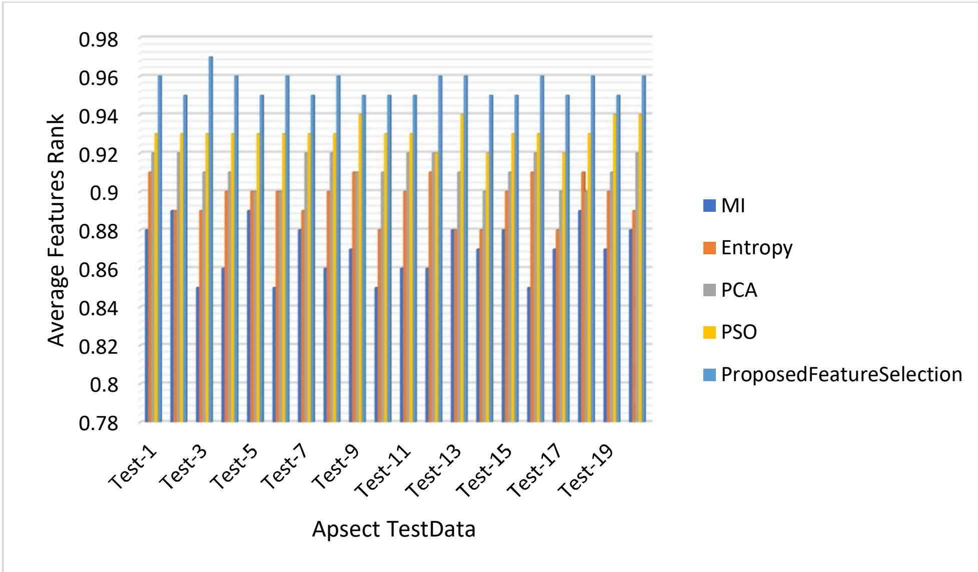



Figure 3: Comparative analysis of proposed classification model to the conventional model on the given input training dataset.

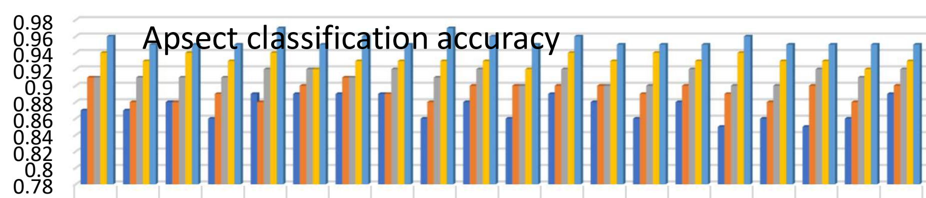
Figure 3, describes the average feature ranking of the proposed ensemble word embedding model to the conventional models on the given aspect training dataset. From the figure, it is noted that the proposed ensemble word embedding model has better average ranking than the conventional ranking measures.

Table 2: Comparative analysis of proposed ensemble ranking measure runtime to the conventional measures.

| AspectData | MI | Entropy | PCA | PSO | **ProposedFeatureSelection** |
|---|---|---|---|---|---|
| TestFS-1 | 5477 | 5006 | 4765 | 5739 | **4325** |
| TestFS-2 | 5288 | 4993 | 4747 | 5337 | **4205** |
| TestFS-3 | 4973 | 5362 | 5590 | 5302 | **4332** |
| TestFS-4 | 5434 | 5575 | 5579 | 5642 | **4019** |
| TestFS-5 | 5116 | 5242 | 5638 | 4924 | **4104** |
| TestFS-6 | 5772 | 4672 | 5583 | 5653 | **4306** |
| TestFS-7 | 5194 | 5800 | 5502 | 4654 | **4334** |
| TestFS-8 | 4801 | 5245 | 5820 | 5837 | **4075** |
| TestFS-9 | 5540 | 5435 | 5506 | 5497 | **3971** |
| TestFS-10 | 5008 | 4693 | 5851 | 5359 | **4244** |
| TestFS-11 | 5504 | 5568 | 5358 | 4901 | **4120** |

| | | | | | |
|---|---|---|---|---|---|
| TestFS-12 | 4952 | 5106 | 5487 | 5464 | **4207** |
| TestFS-13 | 4867 | 5766 | 4967 | 5608 | **4105** |
| TestFS-14 | 4910 | 5819 | 4732 | 5512 | **4195** |
| TestFS-15 | 5666 | 5818 | 5669 | 4956 | **4216** |
| TestFS-16 | 4871 | 5815 | 5430 | 5231 | **3963** |
| TestFS-17 | 5222 | 5549 | 5406 | 5502 | **4139** |
| TestFS-18 | 4672 | 5372 | 4749 | 5778 | **4353** |
| TestFS-19 | 5374 | 5645 | 4838 | 5450 | **3969** |
| TestFS-20 | 5480 | 4729 | 5362 | 4707 | **4005** |

Table 2, describes the performanc of the average runtime of the apsect dataset using the proposed model to the conventional approaches. From the table, it is noted that the average runtime of the proposed ensemble feature rankign measre has better efficiency than the conventional models using CNN.




| | TestAD-1 | TestAD-2 | TestAD-3 | TestAD-4 | TestAD-5 | TestAD-6 | TestAD-7 | TestAD-8 | TestAD-9 | TestAD-10 | TestAD-11 | TestAD-12 | TestAD-13 | TestAD-14 | TestAD-15 | TestAD-16 | TestAD-17 | TestAD-18 | TestAD-19 | TestAD-20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| MI | 0.87 | 0.87 | 0.88 | 0.86 | 0.89 | 0.89 | 0.89 | 0.89 | 0.86 | 0.88 | 0.86 | 0.89 | 0.88 | 0.86 | 0.88 | 0.85 | 0.86 | 0.85 | 0.86 | 0.89 |
| Entropy | 0.91 | 0.88 | 0.88 | 0.89 | 0.88 | 0.9 | 0.91 | 0.89 | 0.88 | 0.9 | 0.9 | 0.9 | 0.9 | 0.89 | 0.9 | 0.89 | 0.88 | 0.9 | 0.88 | 0.9 |
| PCA | 0.91 | 0.91 | 0.91 | 0.91 | 0.92 | 0.92 | 0.91 | 0.92 | 0.91 | 0.92 | 0.9 | 0.92 | 0.9 | 0.9 | 0.92 | 0.9 | 0.9 | 0.92 | 0.91 | 0.92 |
| PSO | 0.94 | 0.93 | 0.94 | 0.93 | 0.94 | 0.92 | 0.93 | 0.93 | 0.93 | 0.93 | 0.92 | 0.94 | 0.93 | 0.94 | 0.93 | 0.94 | 0.93 | 0.93 | 0.92 | 0.93 |
| ProposedFeatureSelection | 0.96 | 0.95 | 0.95 | 0.95 | 0.97 | 0.95 | 0.96 | 0.95 | 0.97 | 0.96 | 0.95 | 0.96 | 0.95 | 0.95 | 0.95 | 0.96 | 0.95 | 0.95 | 0.95 | 0.95 |

Figure 4: Performance of average ensemble feature ranking based classification accuracy on the aspect sentiment data.

Figure 4, describes the performanc of the average accuracy of the aspect feature space using the proposed model to the conventional approaches. From the figure , it is noted that the proposed average accuracy has better efficiency than the conventional models on aspect sentiment data.

Table 3: Performance of average ensemble feature ranking based classification recall on the aspect sentiment data.

| TestData | MI+LR | Entropy+SVM | PCA+RF | PSO+ANN | **ProposedClassifier** |
|---|---|---|---|---|---|
| TestAD-1 | 0.87 | 0.91 | 0.91 | 0.94 | **0.96** |
| TestAD-2 | 0.87 | 0.88 | 0.91 | 0.93 | **0.95** |

| | | | | | |
|---|---|---|---|---|---|
| TestAD-3 | 0.88 | 0.88 | 0.91 | 0.94 | **0.95** |
| TestAD-4 | 0.86 | 0.89 | 0.91 | 0.93 | **0.95** |
| TestAD-5 | 0.89 | 0.88 | 0.92 | 0.94 | **0.97** |
| TestAD-6 | 0.89 | 0.9 | 0.92 | 0.92 | **0.95** |
| TestAD-7 | 0.89 | 0.91 | 0.91 | 0.93 | **0.96** |
| TestAD-8 | 0.89 | 0.89 | 0.92 | 0.93 | **0.95** |
| TestAD-9 | 0.86 | 0.88 | 0.91 | 0.93 | **0.97** |
| TestAD-10 | 0.88 | 0.9 | 0.92 | 0.93 | **0.96** |
| TestAD-11 | 0.86 | 0.9 | 0.9 | 0.92 | **0.95** |
| TestAD-12 | 0.89 | 0.9 | 0.92 | 0.94 | **0.96** |
| TestAD-13 | 0.88 | 0.9 | 0.9 | 0.93 | **0.95** |
| TestAD-14 | 0.86 | 0.89 | 0.9 | 0.94 | **0.95** |
| TestAD-15 | 0.88 | 0.9 | 0.92 | 0.93 | **0.95** |
| TestAD-16 | 0.85 | 0.89 | 0.9 | 0.94 | **0.96** |
| TestAD-17 | 0.86 | 0.88 | 0.9 | 0.93 | **0.95** |
| TestAD-18 | 0.85 | 0.9 | 0.92 | 0.93 | **0.95** |
| TestAD-19 | 0.86 | 0.88 | 0.91 | 0.92 | **0.95** |
| TestAD-20 | 0.89 | 0.9 | 0.92 | 0.93 | **0.95** |

Table 3, describes the performanc of the average recall of the aspect feature space using the proposed model to the conventional approaches. From the figure , it is noted that the proposed average recall has better efficiency than the conventional models on aspect sentiment data.

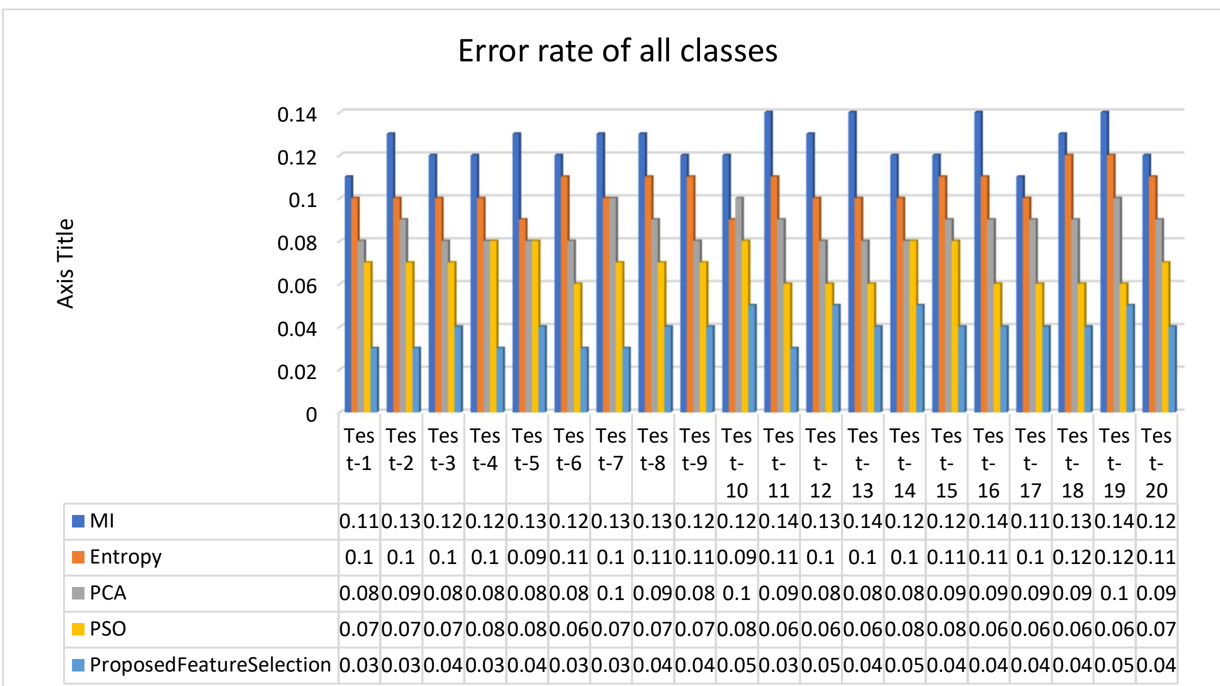



Figure 5 , describes the performanc of the average error rate of the aspect feature space using the proposed model to the conventional approaches. From the figure , it is noted that the proposed average error has better efficiency than the conventional models on aspect sentiment data.

## 5.Conclusion

In this article, a hybrid ensemble feature rankign based classification model is proposed on the large aspect databases. In this work, an advanced multiple word embedding methods are implemented to improve the essential feature extraction problem in the aspect level sentiment process. These multiple word embedding methods are applied on the sentiment databases in the CNN framework. A hybrid classification approach is integrated in the CNN to predict the aspect level classification problem on the databases. Experimental results show that the proposed aspect level CNN based classification model has better performance than the existing aspect sentiment models on large databases. In the experimental results, proposed model has nearly 3% better efficiency in terms of accuracy and runtime(ms) .This work can be further extended to implement a new feature selection model for extracting essential features for efficient cross domain aspect sentiment classification.

## References

[1]M. S. Akhtar, T. Garg, and A. Ekbal, “Multi-task learning for aspect term extraction and aspect sentiment classification,” Neurocomputing, vol. 398, pp. 247–256, Jul. 2020, doi: 10.1016/j.neucom.2020.02.093.

[2]F. Chen, Z. Yuan, and Y. Huang, “Multi-source data fusion for aspect-level sentiment classification,” Knowledge-Based Systems, vol. 187, p. 104831, Jan. 2020, doi: 10.1016/j.knosys.2019.07.002.

[3]Y. Du, M. He, L. Wang, and H. Zhang, “Wasserstein based transfer network for cross-domain sentiment classification,” Knowledge-Based Systems, vol. 204, p. 106162, Sep. 2020, doi: 10.1016/j.knosys.2020.106162.

[4]J. Duan, B. Luo, and J. Zeng, “Semi-supervised learning with generative model for sentiment classification of stock messages,” Expert Systems with Applications, vol. 158, p. 113540, Nov. 2020, doi: 10.1016/j.eswa.2020.113540.

[5]X. Fu et al., “Semi-supervised Aspect-level Sentiment Classification Model based on Variational Autoencoder,” Knowledge-Based Systems, vol. 171, pp. 81–92, May 2019, doi: 10.1016/j.knosys.2019.02.008.

[6]J. A. García-Díaz, M. Cánovas-García, and R. Valencia-García, “Ontology-driven aspect-based sentiment analysis classification: An infodemiological case study regarding infectious diseases in Latin America,” Future Generation Computer Systems, vol. 112, pp. 641–657, Nov. 2020, doi: 10.1016/j.future.2020.06.019.

[7]F. Liu, L. Zheng, and J. Zheng, “HieNN-DWE: A hierarchical neural network with dynamic word embeddings for document level sentiment classification,” Neurocomputing, vol. 403, pp. 21–32, Aug. 2020, doi: 10.1016/j.neucom.2020.04.084.

[8]S. Liu, K. Lee, and I. Lee, “Document-level multi-topic sentiment classification of Email data with BiLSTM and data augmentation,” Knowledge-Based Systems, vol. 197, p. 105918, Jun. 2020, doi: 10.1016/j.knosys.2020.105918.

[9]R. Ma, K. Wang, T. Qiu, A. K. Sangaiah, D. Lin, and H. B. Liaqat, “Feature-based Compositing Memory Networks for Aspect-based Sentiment Classification in Social Internet of Things,” Future Generation Computer Systems, vol. 92, pp. 879–888, Mar. 2019, doi: 10.1016/j.future.2017.11.036.

[10]D. Meškelė and F. Frasincar, “ALDONAr: A hybrid solution for sentence-level aspect-based sentiment analysis using a lexicalized domain ontology and a regularized neural attention model,” Information Processing & Management, vol. 57, no. 3, p. 102211, May 2020, doi: 10.1016/j.ipm.2020.102211.

[11]M. E. Mowlaei, M. Saniee Abadeh, and H. Keshavarz, “Aspect-based sentiment analysis using adaptive aspect-based lexicons,” Expert Systems with Applications, vol. 148, p. 113234, Jun. 2020, doi: 10.1016/j.eswa.2020.113234.

[12]H.-T. Nguyen and L.-M. Nguyen, “ILWAANet: An Interactive Lexicon-Aware Word-Aspect Attention Network for aspect-level sentiment classification on social networking,” Expert Systems with Applications, vol. 146, p. 113065, May 2020, doi: 10.1016/j.eswa.2019.113065.

[13]H. Park, M. Song, and K.-S. Shin, “Deep learning models and datasets for aspect term sentiment classification: Implementing holistic recurrent attention on target-dependent memories,” Knowledge-Based Systems, vol. 187, p. 104825, Jan. 2020, doi: 10.1016/j.knosys.2019.06.033.

[14]D.-H. Pham and A.-C. Le, “Exploiting multiple word embeddings and one-hot character vectors for aspect-based sentiment analysis,” International Journal of Approximate Reasoning, vol. 103, pp. 1–10, Dec. 2018, doi: 10.1016/j.ijar.2018.08.003.

[15]K. Shuang, M. Gu, R. Li, J. Loo, and S. Su, “Interactive POS-Aware Network for Aspect-Level Sentiment Classification,” Neurocomputing, Sep. 2020, doi: 10.1016/j.neucom.2020.08.013.

[16]K. Shuang, X. Ren, Q. Yang, R. Li, and J. Loo, “AELA-DLSTMs: Attention-Enabled and Location-Aware Double LSTMs for aspect-level sentiment classification,” Neurocomputing, vol. 334, pp. 25–34, Mar. 2019, doi: 10.1016/j.neucom.2018.11.084.

[17]J. Wen, G. Zhang, H. Zhang, W. Yin, and J. Ma, “Speculative text mining for document-level sentiment classification,” Neurocomputing, vol. 412, pp. 52–62, Oct. 2020, doi: 10.1016/j.neucom.2020.06.024.

[18]Q. Xu, L. Zhu, T. Dai, and C. Yan, "Aspect-based sentiment classification with multi-attention network," Neurocomputing, vol. 388, pp. 135–143, May 2020, doi: 10.1016/j.neucom.2020.01.024.

[19]M. Yang, Q. Jiang, Y. Shen, Q. Wu, Z. Zhao, and W. Zhou, "Hierarchical human-like strategy for aspect-level sentiment classification with sentiment linguistic knowledge and reinforcement learning," Neural Networks, vol. 117, pp. 240–248, Sep. 2019, doi: 10.1016/j.neunet.2019.05.021.

[20]X. Ye, H. Dai, L. Dong, and X. Wang, "Multi-view ensemble learning method for microblog sentiment classification," Expert Systems with Applications, p. 113987, Sep. 2020, doi: 10.1016/j.eswa.2020.113987.

[21]J. Zhang, M. Chen, H. Sun, D. Li, and Z. Wang, "Object semantics sentiment correlation analysis enhanced image sentiment classification," Knowledge-Based Systems, vol. 191, p. 105245, Mar. 2020, doi: 10.1016/j.knosys.2019.105245.

[22]M. Zhang, V. Palade, Y. Wang, and Z. Ji, "Attention-based Word Embeddings using Artificial Bee Colony Algorithm for Aspect-level Sentiment Classification," Information Sciences, Sep. 2020, doi: 10.1016/j.ins.2020.09.038.

[23]P. Zhao, L. Hou, and O. Wu, "Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification," Knowledge-Based Systems, vol. 193, p. 105443, Apr. 2020, doi: 10.1016/j.knosys.2019.105443.

[24]J. Zhou, Q. Chen, J. X. Huang, Q. V. Hu, and L. He, "Position-aware hierarchical transfer model for aspect-level sentiment classification," Information Sciences, vol. 513, pp. 1–16, Mar. 2020, doi: 10.1016/j.ins.2019.11.048.

[25]J. Zhou, J. X. Huang, Q. V. Hu, and L. He, "SK-GCN: Modeling Syntax and Knowledge via Graph Convolutional Network for aspect-level sentiment classification," Knowledge-Based Systems, vol. 205, p. 106292, Oct. 2020, doi: 10.1016/j.knosys.2020.106292.