**ORIGINAL ARTICLE**

# Advancing aspect-based sentiment analysis with a novel architecture combining deep learning models CNN and bi-RNN with the machine learning model SVM

**Sarsabene Hammi[1] · Souha Mezghani Hammami[1] · Lamia Hadrich Belguith[1]**

## Abstract

Over the last decades, the aspect-based sentiment analysis (ABSA) task has been given great attention and has been deeply studied by the scientific community. It was first introduced in 2002 to extract the users' fine-grained sentiments from textual data by focusing on aspect terms. In this paper, we propose a machine learning-based architecture called CBRS (CNN-Bi-RNN-SVM) to enhance the ABSA of smartphone reviews. This architecture combines two deep learning models [convolutional neural network (CNN) and bidirectional recurrent neural network (Bi-RNN)] with the classical machine learning model support vector machine (SVM). The CNN and the Bi-RNN models are used to capture both local features and contextual information. The SVM model is applied to classify the sentiments, expressed towards aspect terms, as positive or negative. To evaluate the performance of the developed architecture, 8,000 French smartphone reviews, extracted from the Amazon website, are annotated to create a dataset including 15,411 positive aspects and 14,627 negative aspects. The obtained findings corroborated the efficiency of the designed architecture by achieving an F-measure value of 94.05%, for the smartphone dataset, and 85.70% for the SemEval-2016 restaurant dataset. A comparative study demonstrates that the overall performance of our proposed architecture outperformed that of the existing ABSA models.

**Keywords** Aspect-based sentiment analysis · Bi-RNN · CNN · SVM

## 1 Introduction

The content of the web pages is a valuable source of data used and examined in numerous domains including commerce, sociology, finance, etc. For instance, in commercial domain, exploring the customers' opinions about the provided services and products enables companies to optimize their strategies, enhance engagement, and lure new customers (Nasr et al. 2014; Grewal et al. 2017). Recently, with the growing demand to examine this digital content, researchers (Pang et al. 2002) introduced, in 2022, a new task known as sentiment analysis (SA) to identify the sentiments expressed by customers towards someone or something. Moreover, in their previous studies, (Jagtap and Pawar 2013; Medhat et al. 2014) and (Ray and Chakrabarti 2022) distinguished three levels of SA, namely document-level sentiment analysis, sentence-level sentiment analysis and aspect-level sentiment analysis. Indeed, the first and second levels focus on classifying texts as positive, negative or neutral without considering the fine-grained sentiment (Ray and Chakrabarti 2022). However, in the third level, deeper analysis is performed to determine the user's sentiments expressed towards specific aspects of products or service (Ray and Chakrabarti 2022). Besides, applying aspect-based sentiment analysis (ABSA), companies can gain valuable insights about the characteristics of specific products and the customers' viewpoints about them to enhance their quality. For instance, in the comment written below, the positive sentiment of customers regarding the aspects "***design***" and "***camera***" and his/her negative sentiment about the aspects "***battery life***" and "***performance***" can be detected using ABSA.

✉ Sarsabene Hammi
   hsarsabene@gmail.com

   Souha Mezghani Hammami
   souha.hammami@ihecs.usf.tn

   Lamia Hadrich Belguith
   lamia.belguith@fsegs.usf.tn

[1]   University of Sfax, Sfax, Tunisia

*This mobile phone shines in **design** and **camera performance** while it has minor drawbacks in **battery life** and occasional **performance** hiccups.*

Despite the fact that the ABSA task was investigated and employed in numerous research works (Hu and Liu 2004; Hamdan et al. 2015; Ruder et al. 2016), it is obvious that a special interest was given to texts written in the English language and, therefore, there remains a conspicuous research gap concerning other languages including French. In the current study, a machine learning-based method is introduced to identify, from text written in the French language, the sentiments towards aspects. In the conducted experiments, the dataset was pre-processed by treatment of emoticons, removing repeated letters, deleting stop words and stemming. Subsequently, a new architecture called CBRS was designed. It combines deep learning models (CNN and Bi-RNN) with the classical machine learning model SVM. To build up this architecture, the Word2Vec model was initially employed to generate word embedding vectors capturing the semantic information of words. Afterwards, the CNN model processed the generated vector representations, using convolutional filters. Thus, novel vectors that learn the local patterns and features were produced. They were, afterwards, treated by the Bi-RNN model to analyse the sequential dependencies among words and produce new vectors capturing the contextual information and local features. Finally, the SVM model was applied to process the vectors constructed by the Bi-RNN model and classify each aspect term into positive or negative. To evaluate the performance of the introduced architecture, a dataset, with a size comparable to that used in

the state-of-the-art (Yan et al. 2015; Alqaryouti et al. 2020; Hammi et al. 2022), was annotated. It included 8,000 French smartphone reviews collected from the Amazon website.

The remainder of this paper is organized as follows: Section 2 provides an overview about studies focusing on ABSA task. Section 3 describes the key components of the developed architecture. Section 4 presents the introduced ABSA method based on machine learning. Section 5 depicts the data-gathering process and the methodology used to annotate the examined dataset. In Section 6, the different steps of the conducted experiments are illustrated and the performance of the suggested method is assessed when applied on the smartphone dataset and the SemEval-2016 restaurant dataset. Section 7 concludes the paper by summarizing the key findings and outlining future research directions.

## 2 Related works

This section presents some studies focusing on the ABSA task. In early research works (Hu and Liu 2004; Ray and Chakrabarti 2022; Brauwers and Frasincar 2022), the authors identified three ABSA approaches: rule-based approach (Hu and Liu 2004; Moghaddam and Ester 2010; Piryani et al. 2017; Banjar et al. 2021), machine learning-based approach (Mubarok et al. 2017; Wang et al. 2018; Liang et al. 2022; Zhao et al. 2023) and hybrid approach (Vanaja and Belwal 2018; Al-Smadi et al. 2019; Ray and Chakrabarti 2022). Table 1 shows a comprehensive summary of the research works discussed in this section.

**Table 1** Overview about the studies focusing on the ABSA task

| Study | Language | Domain | Method | Results |
|---|---|---|---|---|
| Hu and Liu (2004) | English | Products | Rules | F-measure: 75.80% |
| Moghaddam and Ester (2010) | English | Products | Rules | Accuracy: 81.38% |
| Piryani et al. (2017) | English | Products | Rules | F-measure: 86.42% |
| Banjar et al. (2021) | English | Books | Rules | Accuracy: 87.00% |
| Mubarok et al. (2017) | English | Products | NB | F-measure: 75.00% |
| Wang et al. (2018) | English | Restaurants | Bi-LSTM | F-measure: 68.50% |
|  |  | Laptops |  | F-measure: 66.70% |
| Liang et al. (2022) | English | Restaurants | CNN | F-measure: 62.37% |
|  |  | Laptops |  | F-measure: 63.04% |
| Zhao et al. (2023) | English | Restaurants | DGCN | F-measure:70.44% |
|  |  | Laptops |  | F-measure:55.16% |
| Vanaja and Belwal (2018) | English | Hotels | Linguistic features SVM | F-measure: 84.10% |
| Al-Smadi et al. (2019) | Arabic | Hotels | Linguistic features SVM | Accuracy: 95.40% |
| Ray and Chakrabarti ((2022) | English | Products | Rules | F-measure: 82.02% |
|  |  | Restaurants | CNN | F-measure: 83.34% |
|  |  | Movies |  | F-measure: 80.34% |

## 2.1 Rule-based approach

The rule-based approach is widely applied in Natural Language Processing (NLP) tasks (Yao et al. 2019; Aubaid and Mishra 2020), including text classification, entity-named recognition and sentiment analysis. It uses a set of pre-defined rules constructed by considering the opinion lexicons and various linguistic features such as part-of-speech (POS) tags, syntactic dependencies and lexical cues. In this approach, the POS tags are examined to identify the grammatical category of each word (Straka and Straková 2017). On the other hand, the syntactic dependencies are considered to identify the different dependencies and grammatical relations between words in a sentence (Potisuk 2010). The lexical cues are utilized to specify the sentiments. These rules are essentially applied to capture specific linguistic patterns used to identify the sentiments expressed about aspects within the studied text. The rule-based approaches offer several advantages including interpretability and explicit control of the classification process. In fact, they allow researchers to define rules by investigating their expertise and domain knowledge, enabling the aspect-level sentiment classification in the text. However, they present certain deficiencies in handling complex linguistic variations and contextual nuances. In other words, these approaches are not efficiently utilized to capture subtle linguistic cues or accommodate evolving language usage. Such approaches were applied in many research works like that of (Hu and Liu 2004) who classified the sentiments expressed towards aspect terms using a rule-based method. The authors created two sets of opinion words: positive seed words set (e.g. amazing, wonderful, etc.) and negative seed words set (e.g. terrible, unpleasant, etc.). Subsequently, they examined each opinion word mentioned in the dataset by employing Word-Net, which is a lexical database, to determine if a synonym of that word existed in one of the seed word sets. If it was the case, the opinion word would be added to the appropriate seed word set. Afterwards, based on seed word sets, positive sentiments and negative sentiments were assigned to the aspects. If the aspect term was not associated with an opinion word, a neutral sentiment would be attributed to it. The experimental results revealed the effectiveness of the proposed ABSA method, achieving an F-measure value of 75.80%. In a related study, (Moghaddam and Ester 2010) improved the process of classifying the sentiments expressed towards aspects by increasing the number of sentiment classes to five: positive, very positive, neutral, negative and very negative. They started by compiling a list of opinion words consisting of adjectives collected from Epinions. com. Then, the authors validated each recognized aspect term to check its presence adjacent to any of the listed opinion words. If such proximity existed, a sentiment polarity would be assigned to the aspect term. Moreover, (Piryani

et al. 2017) introduced an enhanced rule-based method to identify the sentiments towards aspects. First, the researchers pre-processed the movie's reviews by correcting the spelling, removing punctuation marks, etc. Subsequently, they constructed a series of linguistic patterns based on the lexical dictionaries SentiWordNet and Generic Lexicon to identify the sentiment. (Banjar et al. 2021) combined linguistic patterns with the lexical resource SentiWordNet to perform the ABSA task. They initially deleted unnecessary elements from the sentences (e.g. hyperlinks hashtags, non-English words, etc). Then, the authors removed objective statements that convey factual data without expressing personal opinions, sentiments or biases. These statements were identified employing the lexical resource SentiWordNet. Subsequently, (Banjar et al. 2021) created a set of syntactic patterns to detect the aspect terms and their related opinion words. Finally, to determine the sentiment polarity of each aspect term, the authors applied the sentiment score calculation function and SentiWordNet on each identified aspect–opinion pair.

## 2.2 Machine learning-based approach

The machine learning-based approach is another approach intensively used in ABSA task (Wang et al. 2018; Liang et al. 2022). Unlike the rule-based approach that relies on pre-defined rules, it leveraged machine learning models to automatically learn the patterns and relationships from data. In the context of aspect-level sentiment classification, the machine learning-based approach trains first the models on the annotated training dataset. Then it uses them to identify the sentiments assigned to the aspect terms existing in the test dataset. In fact, these trained models were designed using various classical machine learning algorithms, such as support vector machines (SVM), Naïve Bayes (NB) and conditional random fields (CRF), or more advanced algorithms called deep learning algorithms like recurrent neural networks (RNNs), convolutional neural networks (CNN), etc. The main advantage of the machine learning-based approach resides in its ability to adapt and generalize new data as the used models learn from examples rather than from handcrafted rules. These models can capture complex relationships and patterns that are difficult to define using a rule-based approach. Despite their efficiency in the ABSA, the machine learning-based approaches show certain limitations. For instance, they can be applied only to substantial amount of annotated training data, which makes their use time-consuming and costly. Below, some previous works relying on machine learning models for the aspect-level sentiment classification task are presented. For example, we mention the work of (Mubarok et al. 2017) who adopted the NB algorithm in ABSA of product reviews. The proposed method consists

of the following steps. First, the reviews were pre-processed to remove noise and irrelevant information. Afterwards, the chi-square (Chi2) method was utilized to select pertinent features to the classification. Subsequently, these features were fed into the NB classifier to determine the sentiments expressed about aspects. Notably, the developed method provided encouraging results by achieving an F-measure of 75.00%. (Wang et al. 2018), on the other side, adopted the Bi-LSTM (bidirectional long short-term memory) model to identify the user's sentiments towards aspects. They exploited both word-level and clause-level information to improve the performance of the used sentiment classification method. First, they enhanced the Bi-LSTM layer by integrating aspect-augmented embedding vectors built by concatenating word vectors and aspect vectors. The Bi-LSTM layer leveraged these produced vectors to generate new vectors containing contextual information about opinion words and aspects within the text. These contextual vectors were, then, integrated into the word-level attention network to specify the most important words based on which the sentiments would be determined. A clause-level attention network was also utilized to extract significant clauses from the dataset by considering contextual and aspect vectors. In the final step, the vectors obtained from the word and clause-level attention networks were integrated into the Bi-LSTM model to predict the sentiments about aspect terms within the dataset. The experimental results revealed that the proposed method outperformed the existing methods by providing the highest F-measure value (66.70% for the laptop dataset). Besides, (Liang et al. 2022) designed a novel method of ABSA relying on the dependency domain knowledge and the CNN algorithm, which allowed better identification of the sentiments towards aspects. (Zhao et al. 2023) introduced a dependency-enhanced graph convolutional network (DGCN) model to perform the ABSA task. This model leveraged both the syntactic dependency structure and the semantic information present in the text. (Zhao et al. 2023) first constructed a dependency graph representation of the input sentences. In this representation, words are nodes while edges correspond to the syntactic relationships between them. Subsequently, the authors enhanced the graph by incorporating semantic information such as word embeddings. Afterwards, the dependency graph convolutional networks (DGCNs) model exploited this enhanced graph to learn representations that consider both the word's sentiment and its relationship with other words in the graph. Finally, the sentiment classifier used the learned representations to predict the sentiment towards each aspect. To assess the performance of the proposed method of ABSA, (Zhao et al. 2023) conducted experiments on benchmark datasets. The obtained results demonstrated that the DGCN model outperformed several

state-of-the-art methods by showing the highest effectiveness in capturing both syntactic and semantic information considered in the sentiment analysis task.

## 2.3 Hybrid approach

The hybrid approach combines the advantages of both rule-based and machine learning-based approaches to promote the accuracy and effectiveness of aspect-level sentiment classification (Villena et al. 2011). It enhances the precision and interpretability of the rule-based approach and improves the generalization and adaptability of the machine learning-based approach. This approach was adopted in several studies such as (Vanaja and Belwal 2018) where the researchers combined the association rules with the SVM algorithm to ameliorate the accuracy of the ABSA method. The authors started by extracting pertinent features that would be used to classify the users' sentiments. These features were identified by applying the association rules algorithm called "Apriori" to extract the frequent aspects in the dataset. Subsequently, they were integrated into the SVM algorithm to classify the sentiment about each aspect term into positive or negative. The utilized algorithm showed high effectiveness in the classification tasks and was utilized to enhance the accuracy of sentiment classification. (Al-Smadi et al. 2019) integrated the morphological, syntactic and semantic features into the SVM algorithm to improve the ABSA of Arabic hotel reviews. The authors first pre-treated the studied dataset using the AraNLP framework (Althobaiti et al. 2014). Afterwards, they created a set of rules applied to extract morphological, syntactic and semantic features from comments. In fact, the morphological features (words' root forms, prefixes and suffixes) were analysed to capture the sentiment-bearing characteristics. However, the syntactic features were determined by extracting dependency relations and grammatical patterns between words to get insights about the expressed sentiment. On the other hand, to identify the semantic features, word embeddings and sentiment lexicons were employed to capture the contextual meaning of words. These features were, then incorporated, into the SVM algorithm to specify the sentiments expressed towards aspect terms. To evaluate the performance of the proposed method, the authors carried out experiments on the Arabic hotel dataset. They also compared the performance of the developed method with that of the existing state-of-the-art methods. The obtained results revealed that the integration of the morphological, syntactic and semantic features into the SVM model significantly improved the accuracy and effectiveness of ABSA in the Arabic hotel reviews. Moreover, (Ray and Chakrabarti 2022) developed a hybrid method to obtain the better results in ABSA task. In their study, the authors initially constructed a set of linguistic and syntactic rules, by taking into consideration intensifiers, negation

words and contextual information, to annotate the learning dataset. The utilized rules were constructed based on the lexical resource SentiWordNet, patterns and sentiment scoring function. Subsequently, (Ray and Chakrabarti 2022) applied a CNN model on the annotated dataset to identify the sentiment associated with each aspect term. This model was chosen as it allows understanding complex sentiment patterns and representations in textual data. To evaluate the goodness of fit of their method, the researchers performed experiments on benchmark datasets used in ABSA. The obtained findings proved that the proposed hybrid method achieved the best performance, compared to the existing methods, by providing interpretability through incorporating domain-specific rules.

# 3 Preliminaries

This section presents a brief overview about the deep learning models, namely convolutional neural network (CNN) and bidirectional recurrent neural network (Bi-RNN), together with the classical machine learning model, support vector machine (SVM), used in the proposed CBRS architecture.

## 3.1 Convolutional neural network

The convolution neural network (CNN) was initially designed by (Fukushima 1980) to perform tasks associated with non-structured data such as images and forms. Lately, CNNs have gained significant popularity within the aspect-level sentiment classification tasks (Liang et al. 2022; Ray and Chakrabarti 2022). The basic target of employing CNN in this research was to reduce the size of input data and focus specifically on the essential features in the ABSA task. Several researchers, such as (Kumar and Sundaram 2022) and (AlAjlan and Saudagar 2021), have consistently highlighted the effectiveness of the CNN compared to other models at the level of selecting relevant local features. As

shown in Figure 1, the architecture of the CNN model is made up of three main layers: convolutional layer(s), pooling layer(s) and a flatten layer. More precisely, the convolutional layer plays a fundamental role in CNN by detecting meaningful features and local patterns from the input data using a set of filters. The output of each convolution layer is called a feature map(s). On the other hand, the pooling layer is employed to reduce the dimensionality of the feature maps produced by the convolutional layers. In fact, it decreases the computational complexity and enhances the model's robustness by taking the maximum value within a fixed-size window. However, the flatten layer reshapes the multi-dimensional vectors generated by the pooling layer into one-dimensional vectors.

## 3.2 Bidirectional recurrent neural network

The recurrent neural network (RNN) model is an artificial neural network introduced by (Graves 2013) to treat sequential data. It features a straightforward architecture and exhibits high efficiency in ABSA tasks (Graves 2013). Unlike the traditional feedforward neural networks that treat each input independently, RNNs incorporate internal memory that enables them to extract contextual information from previous units in the sequence of data. By incorporating the context, they generate outputs influenced by the entire input history, making them effectively used in the analysis of sequential tasks.

The classical RNN model comprises recurrent units which are often referred to as cells. The latter constitute the fundamental building blocks of the RNN model, allowing the model to process sequential data efficiently and uncover meaningful patterns and relationships.

Moreover, the RNN model makes use of the forward propagation mechanism to make predictions on the sequential data. It processes the inputs, sequentially from left to right. In the present study, an enhanced variant of the model, called bidirectional RNN (Bi-RNN), is utilized. It incorporates information from both preceding units (forward



**Fig. 1** The architecture of the CNN model (Hidaka and Kurita 2017)
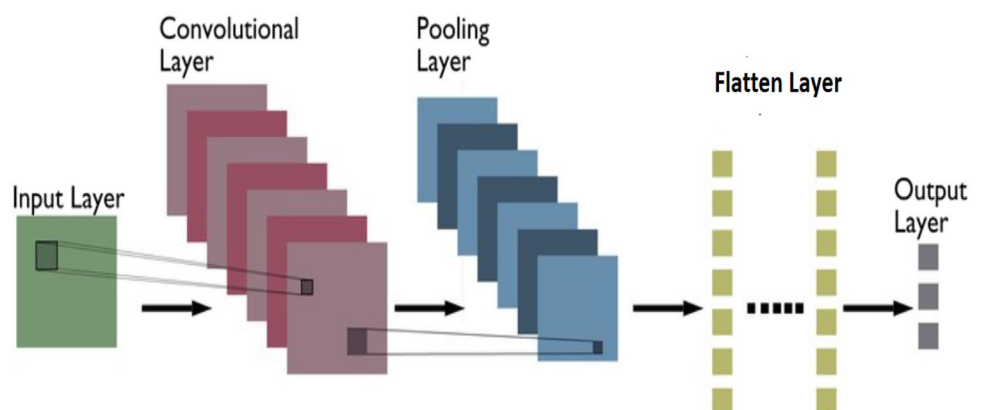
propagation) and subsequent units (backward propagation) to predict word labels. The architecture of the Bi-RNN model is exposed in Figure 2.

## 3.3 Support vector machine

The support vector machine (SVM), introduced by (Cortes and Vapnik 1995), is a popular supervised machine learning model applied in the classification and regression tasks. Recently, its application has been extended to the field of ABSA (Vanaja and Belwal 2018; Al-Smadi et al. 2019) for many reasons. First, the SVM model efficiently handles high-dimensional data where each feature represents a specific aspect or attribute, as it is the case in the ABSA task. In fact, it can analyse such data effectively by mapping it to a higher-dimensional space using the kernel trick. This ability allows it to capture complex relationships between features and provide better aspect-level sentiment analysis accuracy (Vanaja and Belwal 2018). Second, by finding an optimal hyper-plane with a maximum margin, this model may also reduce considerably the risk of overfitting and enhance the performance of the ABSA methods.

In this study, the SVM model is opted thanks to its substantial effectiveness in the aspect-based sentiment classification tasks. Previous researches (Alharbi 2021; Zhao et al. 2023) demonstrated that the classification results obtained by combining deep learning models with the traditional machine learning model (SVM) are better than those provided by using each model separately.

Therefore, in the following section, we attempt to enhance the ABSA task by combining the above-mentioned models (CNN, Bi-RNN and SVM) into a new architecture called CBRS.

## 4 Proposed method

In this study, a machine learning-based method is introduced and used for the ABSA task by applying the following steps. First, the dataset is pre-treated utilizing a set of pre-processing techniques including removing stop words, stemming, treatment of emoticons, etc. Then, a new architecture named **CBRS** is designed, combining the strength of **C**NN, **B**i-**R**NN and **S**VM models. Figure 3 illustrates the workflow of the suggested machine learning-based method.

### 4.1 Pre-processing of the dataset

Pre-processing is a fundamental step in any NLP task (Angiani et al. 2016; Palomino and Aider 2022) applied to clean and prepare the dataset for the classification task. It allows improving data quality, addressing inconsistencies, reducing noise and transforming data into a suitable format for the training phase. In this step, the dataset was pre-processed by removing stop words, stemming, treatment emoticons, etc.

- **Removing stop words:** The main objective of this step is to remove insignificant words, like conjunctions and pronouns, from the dataset used in the present work. To accomplish this step, we utilized the NLTK library containing a predefined list of French stop words. To adapt the list to the ABSA task, certain words ("**est**/*is*", etc.) that give hints about the relations between aspects and opinion words, were deleted.
- **Words stemming:** Stemming aims at converting inflected words into their base or stem (Jivani 2011). In this step, both prefixes and suffixes were removed from words. As a result, words having similar meaning were treated as the same word. In this process, words, which share a common stem and convey a similar semantic representation, were grouped.
- **Treatment of emoticons:** Emoticons are non-verbal expressions that customers often use in their comments

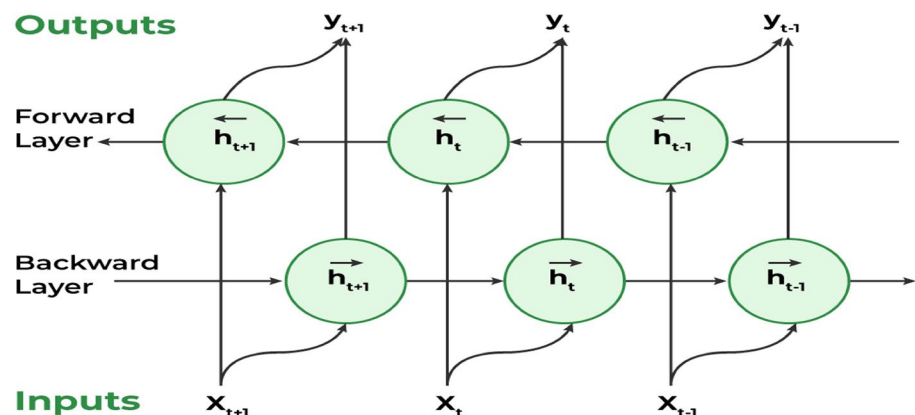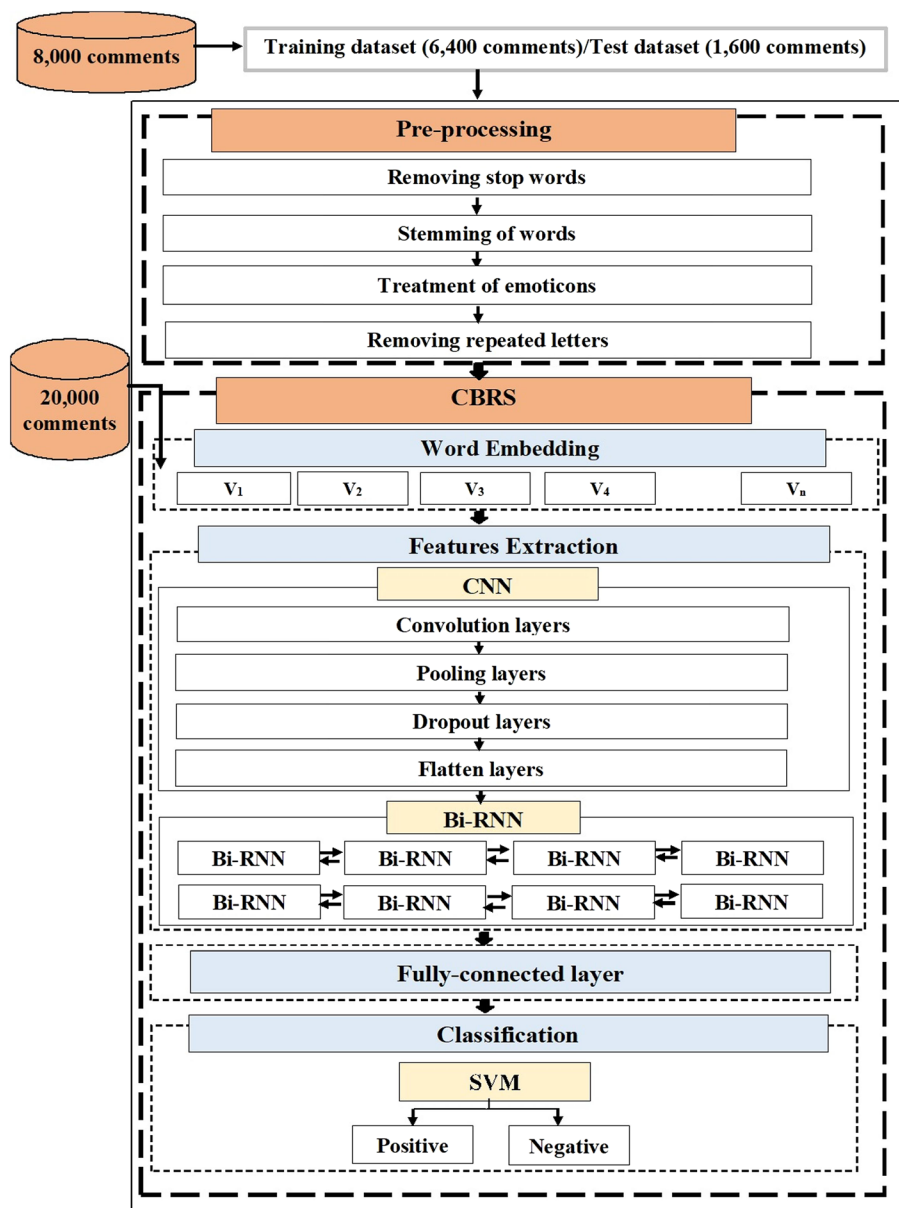**Fig. 2** The architecture of the Bi-RNN model (Deligiannidis et al. 2021)

**Fig. 3** workflow of the developed machine learning-based method



(Hammi et al. 2022). They serve as valuable indicators of sentiment polarity, reflecting the opinions and attitudes of the users towards various aspects of the product or service. In the conducted experiments, emoticons were converted into expressive tags that effectively represent their associated sentiments. From this perspective, seven types of emoticons were distinguished: happy (emot_heureux), laugh (emot_rire), love (emot_amour), sad (emot_triste), wink (emot_CO) and cry (emot_pleure).

- **Removing repeated letters:** In social networks, customers usually repeat the same letter consecutively within a word to emphasize their emotions. This repetition is often used to convey intense sentiments such as excitement, happiness or sadness. As far as the French word "**parfaitttttt**" (which means *"perfectttttt"* in English)

is concerned, it was simplified, for instance, by removing repeated letters and returning the word to its original form, "**parfait**" (*"perfect"*).

## 4.2 CBRS: an architecture for ABSA

This section depicts the overall steps applied to build the CBRS (CNN-Bi-RNN-SVM) architecture for ABSA. This architecture comprises four key components: i) word embedding to construct representations that capture the semantic meaning of words; ii) a convolutional neural network (CNN) to extract local features, a bidirectional recurrent neural network (Bi-RNN) to capture contextual information; iii) and a support vector machine (SVM) to classify the sentiments expressed towards aspects. Figure. 4 provides a detailed
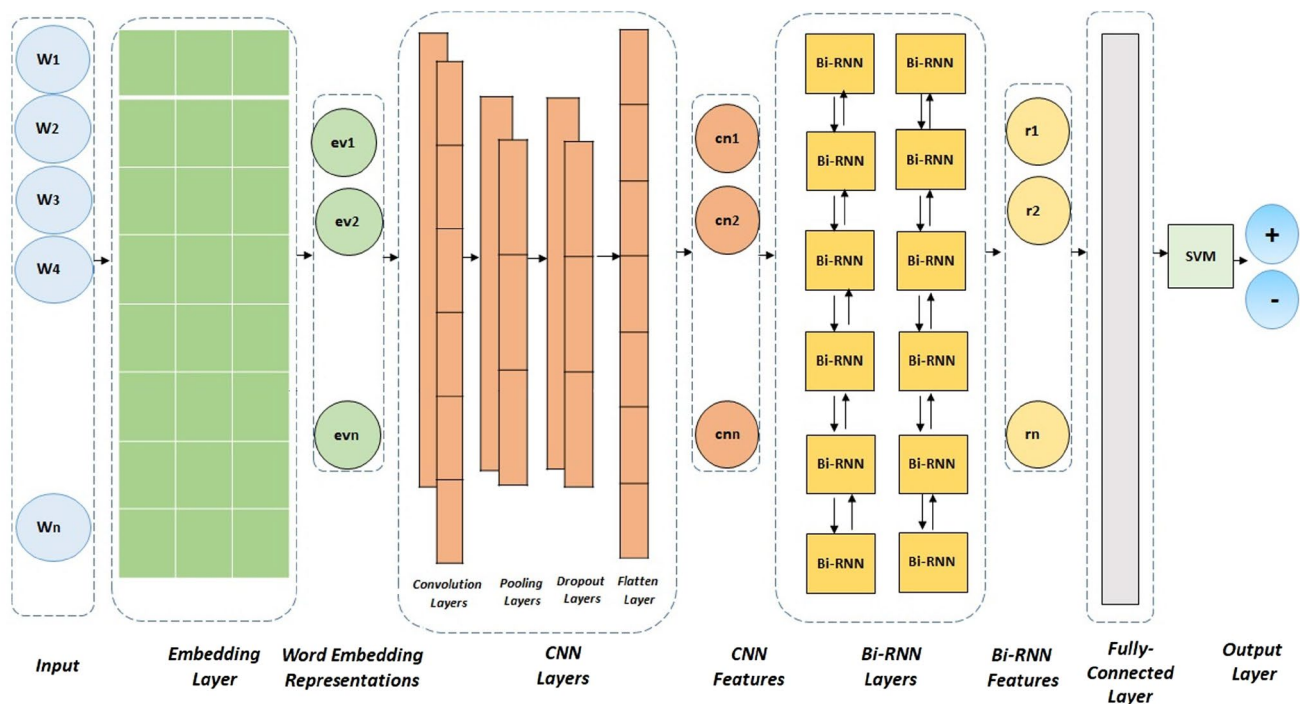
**Fig. 4** CBRS architecture

representation of the CBRS architecture. To implement the latter, vector representations, which capture the semantic meaning of words within a comment, were first constructed using the embedding layer. Subsequently, the obtained vectors were integrated into the CNN layer to generate new word vectors considering local features and patterns within the text. These word vectors were then incorporated into the Bi-RNN layer. By taking advantage from its recurrent structure to capture the contextual information and the dependencies between words, this model produced new vectors that encapsulated the nuanced relationships within the comment. Finally, the SVM model treated the vectors generated by the Bi-RNN model and classified each aspect term into positive or negative.

### 4.2.1 Embedding layer

The embedding layer is a key component in any neural network architecture (Liu et al. 2022). It is generally used to transform discrete input, such as words or categorical variables, into continuous vector representations. To perform this task, several techniques, such as term frequency–inverse document frequency and one-hot encoding, were developed. Nonetheless, the primary shortcoming of these techniques resides in their incapability to extract semantic meanings which is an important step in aspect-level sentiment classification tasks. Thus, we applied the word embedding technique as an alternative. Indeed, this technique effectively

captures the relationships among words that frequently occur in similar contexts by representing them with comparable vectors of words. To construct word embedding vectors, numerous models, such as FastText and ELMO, were designed in the literature. In the conducted experiments, the widely adopted Word2Vec model was utilized to learn knowledge about word contexts. This model was trained on 20,000 French smartphone comments (described in Section 5) and 10,000 comments related to the restaurant domain and extracted from the Yelp website. The equation employed by the Word2Vec model to convert each word into a V-dimensional word vector is determined as follows.

$$ev = [ev_1, ev_2, \ldots, ev_n] \in \mathbb{R}^{n \times V}, \tag{1}$$

where

- $ev$ represents the word embedding matrix.
- $ev_1, ev_2, ..., ev_n$: denote the word embeddings of n different words in the vocabulary.

Each word in the dataset was converted into a vector representation with a dimensionality of 250.

### 4.2.2 Convolutional, pooling, dropout and flatten layers

At time step $t$, the convolutional layer takes the word vector embeddings produced by the embedding layer and

applies several filters to generate feature maps. These feature maps capture local information. Given a sentence $Sent_1$ composed of words $Wrd_1, Wrd_2,..., Wrd_n$ transformed into vectors of words embedding $ev_1, ev_2,...,ev_n$, $f$ filters is utilized to produce feature maps, which can be formulated as follows:

$$Wrd = [Wrd_1, Wrd_2, \ldots, Wrd_n - s_i + 1] \tag{2}$$

$$\mathrm{Wrd}_{i:j} = \mathrm{ReLU}\left( \sum_{f=1}^{f} \left( F_{ev}^{(f)} * \mathrm{W}_{i:i+s-1}^{(f)} \right) + b_{conv} \right), \tag{3}$$

where

- $\mathrm{Wrd}_{i:j}$ represents the feature maps produced by the convolutional layer for the word vector embeddings at position $i$ to $j$.
- $F_{ev}^{(f)}$ refers to the $f$th filter in the network.
- $W_{i:i+s-1}^{(f)}$ denotes the weights associated with the $f$th filter applied to the input sequence from $i$ to $i + s - 1$.
- $b_{conv}$: stands for the bias term used by the convolutional layer.

Despite the CNN model's outstanding performance in selecting relevant features, its ability to capture local information often necessitates the utilization of a substantial number of stacked convolutional layers (Ellouze and Hadrich 2022). Thus, a pooling layer was utilized, in the performed experiments, to decrease the number of parameters and generate vectors of words containing pertinent information about the words.

$$Wrd^\wedge = \max\{Wrd\} \tag{4}$$

The word vectors generated from the pooling layer, denoted as $Wrd^\wedge 1, Wrd^\wedge 2, Wrd^\wedge 3, ..., Wrd^\wedge n$, were, subsequently, fed into a dropout layer. The latter was used to increase the efficiency of the model in classifying sentiments and reducing overfitting by randomly dropping out some of the neurons using a dropout rate.

$$Wrd_{\mathrm{dropout}} = Dropout(Wrd^\wedge, rate = \mathrm{dropout\_rate}) \tag{5}$$

Afterwards, a flatten layer was used to transform the multi-dimensional vectors into unidimensional vectors.

$$Wrd_{\mathrm{flatten}} = \mathrm{Flatten}(Wrd_{\mathrm{dropout}}) \tag{6}$$

In this process, two convolution layers were used; each of which comprised feature maps equal to 200 with a kernel size equal to 3. Additionally, two pooling layers, followed by two dropout layers and a flatten layer, were deployed.

### 4.2.3 Bidirectional recurrent neural network layers

The Bi-RNN layer takes the one-dimensional vectors from the flatten layer and produces new words' vectors, considering both local and contextual information. The architecture of the Bi-RNN model enables it to capture contextual information from not only previous hidden nodes, but also from subsequent hidden nodes. This model comprises two essential components: the forward RNN and the backward RNN. The former processes the input sequence, capturing information generated by the previous nodes, from left to right. On the other hand, the latter functions as a forward RNN, but in the opposite direction (i.e. from right to left). It makes predictions about the sentiments expressed towards aspects by leveraging information coming from subsequent hidden nodes (at time step $t + 1$). At each time step $t$, the Bi-RNN model receives the input vector $Wrd_{\mathrm{flatten}}$, the previous hidden state $h_{\overrightarrow{t-1}}$, and the subsequent hidden state $h_{\overleftarrow{t+1}}$ and applies a nonlinear transformation. This transformational process is used to calculate the outputs of the current hidden node at time step $t$ by capturing intricate temporal relationships and contextual information extracted from both previous and subsequent nodes. The equations employed by the Bi-RNN model to process the input data and update the hidden state at each time step are determined as follows:

$$\overrightarrow{h}_t = \tanh\left( W_{\overrightarrow{ih}} \cdot Wrd_{\mathrm{flatten}} + W_{\overrightarrow{hh}} \cdot h_{\overrightarrow{t-1}} + b_{\overrightarrow{h}} \right) \tag{7}$$

$$\overleftarrow{h}_t = \tanh\left( W_{\overleftarrow{ih}} \cdot Wrd_{\mathrm{flatten}} + W_{\overleftarrow{hh}} \cdot h_{\overleftarrow{t+1}} + b_{\overleftarrow{h}} \right) \tag{8}$$

$$h_t = [\overrightarrow{h}_t, \overleftarrow{h}_t], \tag{9}$$

where

- $h_t$, $\overrightarrow{h}_t$, and $\overleftarrow{h}_t$ represent the current hidden state, the previous hidden state, and the subsequent hidden state at time step $t$, respectively.
- $W_{\overrightarrow{ih}}$ denotes the weight matrix connecting the input to the previous hidden state.
- $W_{\overleftarrow{ih}}$ expresses the weight matrix connecting the input to the subsequent hidden state.
- $Wrd_{\mathrm{flatten}}$ corresponds to the flattened output extracted from the previous layer at time step $t$.
- $W_{\overrightarrow{hh}}$ stands for the weight matrix connecting the previous hidden state to the current hidden state.
- $W_{\overleftarrow{hh}}$ designates the weight matrix connecting the subsequent hidden state to the current hidden state.
- $h_{\overrightarrow{t-1}}$ and $h_{\overleftarrow{t+1}}$ refer to the previous hidden state and the subsequent hidden state at the time step $t - 1$ and $t + 1$, respectively.

- $b_{\overrightarrow{h}}$ and $b_{\overleftarrow{h}}$ express the bias term of the hidden state at the time step $t-1$ and $t+1$, respectively.

The vectors generated at this step were, then, passed, through a fully connected layer before being integrated into the SVM. Two Bi-RNN layers, each of which contained 250 neurons, were employed.

### 4.2.4 Fully connected layer

The fully connected layer is generally adopted to create a high-level representation of words that can be useful for the ABSA task. Its output is obtained as follows:

$$f_c = Relu(h_t.W_{fc} + b_{fc}), \tag{10}$$

where

- $f_c$ denotes the output of the fully connected layer.
- $W_{fc}$ is the weight matrix of the fully connected layer.
- $b_{fc}$ corresponds to the bias vector of the fully connected layer.

In the performed experiments, a single fully connected layer, comprising 128 neurons, and the Rectified Linear Unit (ReLU) activation function was applied.

### 4.2.5 Output layer

The SVM model predicts the sentiment label assigned to each aspect by considering the features generated by the fully connected layer. It is generally used to identify the optimal hyper-plane that can effectively separate data points belonging to different classes while maximizing the distance boundary between negative and positive classes (Rizwan et al. 2021). The SVM model relies on a set of linear functions that effectively separate classes, A and B, and can be represented as follows:

$$output = sgn(f(f_c, W_{svm}, b_{svm})) \tag{11}$$

$$f(f_c, W_{svm}, b_{svm}) = \langle W_{svm}, f_c \rangle + b_{svm} \tag{12}$$

where

- $Ws_{svm}$ is the weight vector of the SVM model.
- $b_{svm}$ denotes the bias term of the SVM model.

The $sgn()$ function represents the signum function applied to determine the predicted class label of a given input. On the other hand, the function $f(f_c, W_{svm}, b_{svm})$ is used to calculate the value of the decision function, while $sgn(f(f_c, W_{svm}, b_{svm}))$ assigns a class label to aspects based on the sign of the decision function value. If the decision function is positive, the predicted class belongs to the first label. However, if it is negative, the predicted class belongs to the second label.

## 5 Dataset collection and annotation

In the conducted study, 28,000 smartphone comments were collected from the Amazon website containing a big number of customer's reviews, ensuring a sizable and diverse dataset. This dataset, including comments belonging to more than seven prominent smartphone brands, was gathered from May 2022 to December 2022. Table 2 presents an overview of the distribution of comments made about each brand.

Given that machine learning-based methods require a substantial amount of annotated data (Sarker 2021; Taye 2023), we requested the assistance of three experts from our laboratory to perform the annotation process. During this step, the annotators labelled only 8000 comments out of 28,000 comments, while the remaining comments (20,000) were used to train the Word2Vec model (Section 4).

To annotate the dataset, the annotation guideline was first created. It includes definitions of key concepts (aspects, classes of sentiments), instructions and a set of illustrative examples related to the ABSA task. Afterwards, the dataset was segmented into four files. Each of them, comprising 2000 comments, was annotated by two annotators. In the cases of disagreement between both

**Table 2** The number of comments in the dataset according to the brand

| Brands | Number of comments | Positive aspects | Negative aspects | Neutral aspects |
|---|---|---|---|---|
| Samsung | 6856 | 3964 | 3898 | 267 |
| iPhone | 6000 | 3041 | 3405 | 128 |
| Nokia | 3564 | 1856 | 1351 | 161 |
| Huawei | 2100 | 1034 | 998 | 120 |
| Oppo | 1506 | 987 | 921 | 99 |
| Infinix | 1307 | 889 | 845 | 51 |
| Xiaomi | 1167 | 851 | 823 | 74 |
| Other brands | 5500 | 2789 | 2386 | 102 |

annotators, the first author would try to find a consensus or a resolution and make the ultimate decision. The annotators labelled each aspect in the dataset with one of three sentiment labels. They used the "POS" and the "NEG" labels to refer to aspects expressing positive sentiments and negative sentiments, respectively. On the other hand, the "NEU" label was utilized to show aspects where the customer expressed neither positive nor negative sentiment. Finally, a set of 31,040 aspects, including 15,411 positive aspects; 14,627 negative aspects and 1,002 neutral aspects, was obtained. It is worth to note that, in the performed study, the neutral aspects were not considered because of their small number.

To annotate the dataset, the annotation guideline was first created. It includes definitions of key concepts (aspects, classes of sentiments), instructions and a set of illustrative examples related to the ABSA task. Afterwards, the dataset was segmented into four files. Each of them, comprising 2000 comments, was annotated by two annotators. In the cases of disagreement between both annotators, the first author would try to find a consensus or a resolution and make the ultimate decision. The annotators labelled each aspect in the dataset with one of three sentiment labels. They used the "POS" and the "NEG" labels to refer to aspects expressing positive sentiments and negative sentiments, respectively. On the other hand, the "NEU" label was utilized to show aspects where the customer expressed neither positive nor negative sentiment. Finally, a set of 31,040 aspects, including 15,411 positive aspects; 14,627 negative aspects and 1,002 neutral aspects, was obtained. It is worth to note that, in the performed study, the neutral aspects were not considered because of their small number.

# 6 Experiments

This section presents the hyper-parameters utilized to implement the architecture. It also describes a series of experiments conducted to assess the performance of the introduced ABSA method. The first experiment was performed to evaluate the overall performance of the proposed machine learning-based method. It demonstrated that the combination of multiple machine learning models enhances the overall performance of the ABSA task. However, in the second experiment, the influence of the number of Bi-RNN layers on the classification results was analysed. In the third experiment, the efficiency of word embedding models, namely Word-2Vec and Glove, in the ABSA task were compared. Finally, in the last experiment, the performance of the introduced machine learning-based method and that of several benchmark studies were compared.

## 6.1 Hyper-parameter settings

This section describes the hyper-parameter settings used in the applied models. The process of selecting the optimal hyper-parameters is a challenging task that varies based on the characteristics of the dataset such as its size and the structure of sentences. To this end, several trials were made to choose the hyper-parameters with which the architecture could yield the best performance results.

The introduced CBRS architecture consists of multiple layers. First, a word embedding layer is employed to generate word vectors with a dimension of 250. Then, the first convolutional layer is applied with 200 feature maps and a kernel size of 3. It uses the Rectified Linear Unit (ReLU) as the activation function. After that, the first pooling layer is utilized to reduce the number of parameters. Subsequently, the first dropout layer, with a dropout rate equal to 0.3, is employed. This layer reduces overfitting by randomly dropping out some of the neurons from the output of the pooling layer. After that, the second convolutional layer, having the same parameters as the first convolutional layer, is introduced. It is followed by the second pooling layer and the second dropout layer. Additionally, a flatten layer is applied to convert the multi-dimensional vectors obtained from the previous layers into a one-dimensional vector. Afterwards, two Bi-RNN layers (each containing 250 neurons) are used to capture contextual features. Following the Bi-RNN layers, a fully connected layer, with 128 neurons, is utilized together with the ReLU activation function. It allows integrating the learned features and preparing the data for the final classification step. Finally, a classification layer, based on SVM model, is used to classify the sentiments expressed towards aspect terms as positive or negative. In the CBRS, 50 epochs, a batch size of 64, a dropout rate of 0.3 and a learning rate of 0.0025 are considered. The optimization process of the adopted model relies on the Adam optimizer. Additionally, a loss function, chosen based on cross-entropy considerations, is employed. Table 3 represents the hyper-parameters used in the designed CBRS architecture.

Additionally, in the performed study, we evaluated the performance of the introduced architecture by comparing it with that of two deep learning models: CNN and Bi-RNN. The CNN model consists of two convolution layers; each of which comprises feature maps equal to 200 and uses a kernel size of 3 with Rectified Linear Unit (ReLU) activation function. It also includes two pooling layers, two dropout layers with a dropout rate equal to 0.3, a flatten layer, a fully connected layer comprising 128 neurons and an output layer. However, the Bi-RNN model is concerned, it comprises two Bi-RNN layers, each of which includes 250 neurons, a fully connected layer, comprising 128 neurons, and an output layer. In these two models, a batch size of 64, a

**Table 3** CBRS hyper-parameters

| Hyper-parameter | Value |
| --- | --- |
| Number of epochs | 50 |
| Batch size | 64 |
| Embedding dimension | 250 |
| Number of convolutional layers | 2 |
| Number of pooling layers | 2 |
| Kernel size | 3 |
| dropout rate | 0.3 |
| Number of Bi-RNN layers | 2 |
| Number of fully connected layers | 1 |
| Activation function type | ReLU |
| Learning rate | 0,0025 |
| Optimizer | Adam |
| Loss function | Binary cross-entropy |

**Table 4** Evaluation of the learning-based method

| Dataset | Model | Precision (%) | Recall (%) | F-measure (%) |
| --- | --- | --- | --- | --- |
| Smart-phone Dataset | CNN | 85.41 | 96.83 | 90.67 |
| | Bi-RNN | 87.39 | 84.98 | 86.17 |
| | SVM | 73.25 | 62.32 | 67.34 |
| | **CBRS** | **90.79** | **97.57** | **94.05** |
| Seme-val-2016 restaurant dataset | CNN | 77.25 | 81.15 | 79.15 |
| | Bi-RNN | 80.61 | 72.20 | 76.17 |
| | SVM | 68.51 | 56.31 | 61.80 |
| | **CBRS** | **83.41** | **88.12** | **85.70** |

number of epochs equal to 50 and a learning rate of 0.0025 are considered.

## 6.2 Evaluation of the proposed method

This section provides a comprehensive assessment of the proposed method showcases its relevance and effectiveness in accurately classifying the sentiments expressed towards aspects into positive and negative. In the experiments, two benchmark datasets, namely the smartphone dataset (containing 8000 comments) and the SemEval-2016 restaurant dataset (including 335 comments) (Pontiki et al. 2016), were used. Each dataset was divided into two subsets. The first one comprised 80% of the data and was utilized to train the CBRS architecture. On the other hand, the second subset contained the remaining 20% of the data and was employed

**Table 5** Comparison of the results provided by CBRS architecture with those obtained by other architectures

| Dataset | Model | Precision (%) | Recall (%) | F-measure (%) |
| --- | --- | --- | --- | --- |
| Smart-phone Dataset | CBRN | 84.20 | 89.51 | 86.77 |
| | CBRL | 85.27 | 90.31 | 87.71 |
| | CBRC | 87.60 | 93.04 | 90.23 |
| | **CBRS** | **90.79** | **97.57** | **94.05** |

**Table 6** Assessment of the CBRS architecture for classifying positive, negative and neutral aspects

| Model | Precision (%) | Recall (%) | F-measure (%) |
| --- | --- | --- | --- |
| CBRC | 74.28 | 79.99 | 77.02 |

to assess the performance of CBRS architecture in classifying the sentiments towards aspects.

As revealed in Table 4, the proposed architecture exhibited high performance when applied on the smartphone dataset by providing excellent precision, recall and F-measure values of 90.79%, 97.57% and 94.05%, respectively. These results highlighted the robustness, feasibility and effectiveness of the designed architecture in handling ABSA tasks. It was also obvious that, in terms of F-measure value, this architecture significantly outperformed CNN, Bi- RNN and SVM models by 4, 8 and 27%, respectively. The performance of the introduced architecture was also evaluated on the SemEval-2016 restaurant dataset. In fact, it achieved competitive precision, recall and F-measure values equal to 83.41, 88.12 and 85.70%, respectively. It also boosted the performance of the CNN, Bi-RNN and SVM models by 5, 9 and 24%, respectively.

To assess the performance of the SVM model in aspect-level sentiment classification, it was replaced with one of three commonly used machine learning models: Naive Bayes (NB), logistic regression (LR) and conditional random fields (CRF). The obtained results (Table 5) confirmed that the SVM model outperformed the three other models when applied on the smartphone dataset. Indeed, it enhanced the performance of the CBRN (CNN-Bi-RNN-NB), CBRL (CNN-Bi-RNN-LR) and CBRC (CNN-Bi-RNN-CRF) models by approximately 7, 6 and 4%, respectively.

As we stated above, the experiments were conducted on two classes of sentiment: positive and negative. The neutral class was excluded due to the limited number of neutral aspects in the dataset (Section 5). However, the performance of the CBRS architecture in detecting neutral sentiments together with positive and negative sentiments was also evaluated by performing an experiment on a small sample of the studied smartphone dataset. This sample consisted

of 2,691 comments (1,300 positive aspects; 1,300 negative aspects and 1,000 neutral aspects). The obtained results, as shown in Table 6, demonstrated the high performance of the introduced architecture in identifying the three sentiment classes (F-measure of 77.02). Nevertheless, these results remain modest due to the limited number of aspects in the examined dataset.

## 6.3 The effect of the number of Bi-RNN layers on the CBRS architecture

In this subsection, the impact of the number of Bi-RNN layers on the performance of the CBRS architecture is assessed. In this regard, the performance of three variations of the CBRS architecture, made up of one, two and three Bi-RNN layers, respectively, is compared. Each Bi-RNN layer in the three variations consists of 250 neurons. As displayed, in Table 7 the optimal classification results were obtained by using the architecture containing two layers of Bi-RNN. This variation demonstrated a significant improvement in the aspect-based sentiment classification task, compared to both the single-layer and three-layer CBRS architecture. It also enhanced the F-measure value by approximately 3% for the smartphone dataset, thanks to the balance between the complexity of the used model and its capacity to capture relevant patterns. In fact, with a single Bi-RNN layer, the model may lack the necessary capacity to adequately learn and represent intricate relationships between words within the data. As a result, the classification accuracy decreased due to the limited potential of the model. On the other hand,

the use of three Bi-RNN layers entailed a degradation in the classification results, compared to the two-layer configuration because of overfitting. In fact, the incorporation of an extra layer led to the rise of the model's complexity and the number of the used parameters, which resulted in the over-optimization of the training data. Therefore, the performance of the proposed architecture has been decreased.

It is also obvious that, when applied on the smartphone dataset, two-layer CBRS architecture struck an optimal balance. This configuration allowed the architecture to capture effectively relevant patterns and dependencies within the data. It also allows to shun the limitations of a single-layer architecture while mitigating the risk of overfitting resulting from the high number of layers.

## 6.4 Comparison of the word embeddings models: Word2Vec and Glove

This subsection evaluates the efficiency of both Word2Vec and Glove (Global Vectors for Word Representation) models in constructing word embeddings for the ABSA task. In fact, Glove corresponds to an unsupervised learning model applied to generate word embeddings. It was introduced by (Pennington et al. 2014) to learn word representations based on the co-occurrence statistics of words in a large dataset. This model constructs first a global co-occurrence matrix that represents the word co-occurrence frequencies in the corpus. Then, it factorizes this matrix to obtain word embeddings that capture semantic relationships between words. As revealed in Figure 5, the Word2Vec model showed higher

**Table 7** Assessment of the optimal number of Bi-RNN layers

| Dataset | Number of Bi-RNN layers | Precision (%) | Recall (%) | F-measure (%) |
|---|---|---|---|---|
| Smartphone dataset | 1 | 88.67 | 93.48 | 91.01 |
| | 2 | 90.79 | 97.57 | **94.05** |
| | 3 | 87.49 | 95.61 | 91.37 |

**Fig. 5** Comparison of the performance of Word2Vec and Glove word embedding models
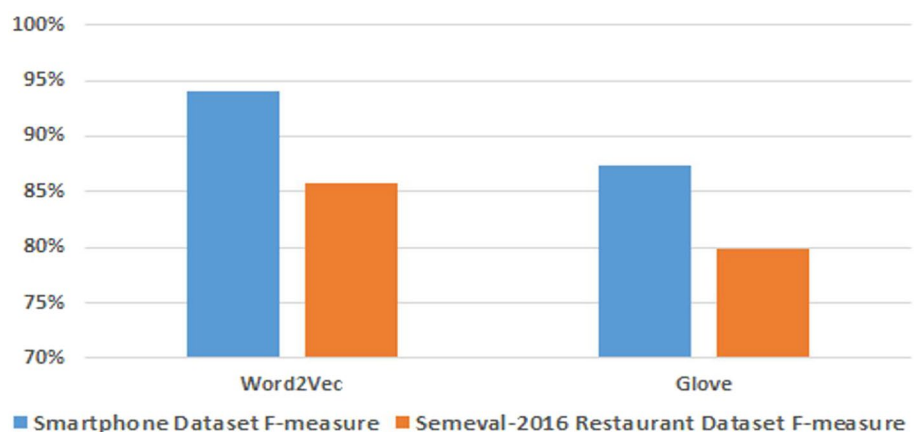
performance than the Glove model in the ABSA task. It improved the F-measure value provided by the Glove model by approximately 7% and 6% for the smartphone dataset (87,42%) and the SemEval-2016 restaurant dataset (79.85%), respectively. These outstanding results are attributed to several factors. First, the Word2Vec model focuses on capturing local context information by considering the neighbouring words, which allows extracting more fine-grained sentiment-related relations within a specific context. Moreover, as sentiment analysis requires understanding the sentiment in the local contexts, Word2Vec's ability to capture local context can be advantageous. Second, the Word2Vec can generate meaningful representations for Out-Of-Vocabulary (OOV) words. By leveraging the character n-grams, Word2Vec approximates the embeddings of OOV words by considering their similarities to known words. This is particularly useful in sentiment analysis tasks where domain-specific or rare sentiment-related words may be found.

### 6.5 Comparison of the proposed method with the state-of-the-art methods

To evaluate the effectiveness and goodness of fit of the developed machine learning-based method, a comparative analysis of the performance of the latter with that of other benchmark studies (Ruder et al. 2016; Brun et al. 2016; Akhtar et al. 2017; Pigneul and Kooli 2018; Brun 2018; García-Pablos et al. 2018; Akhtar et al. 2019; Essebbar et al. 2021; Kumar et al. 2016) was conducted. The authors, in previous studies, used the French SemEval-2016 restaurant dataset to evaluate the performance of their suggested methods. For instance, (Ruder et al. 2016) employed a CNN model to identify the sentiments expressed towards aspects. They first pre-processed the dataset by removing stop words. Subsequently, they incorporated aspect-augmented embedding vectors into their CNN model. These vectors were generated by concatenating the aspect vector with the word embedding vector of each word. (Brun et al. 2016) introduced a system combining syntactic and semantic knowledge with the machine learning algorithm SVM to identify the sentiments of customers about aspect terms. In this study, the syntactic and semantic knowledge was extracted from dependency parsers and lexicons. Then, it was incorporated into the SVM model. Similarly to (Brun et al. 2016), (Kumar et al. 2016) and (Akhtar et al. 2017) also used the SVM algorithm in the ABSA task. In their research, (Kumar et al. 2016) first pre-processed the dataset by removing stop words, punctuation, etc. Subsequently, they extracted multiple features including domain-dependency graph features and distributional thesaurus features. These features were, then, fed into the SVM model to perform the ABSA task. (Pigneul and Kooli 2018) elaborated an attention-based memory network to classify the sentiments

expressed towards aspects. This network consists of a series of blocks, each of which comprises a linear layer and an attention layer. In fact, the former applies a linear transformation to the embedding vectors of the aspects. However, the latter assigns importance to specific contextual words by considering semantic information and positional connections between words. The attention-based memory network takes, as input, the word embedding vectors of the aspects and generates sentiment polarities about aspect terms as output. In their study, (Brun 2018) proposed a new method combining semantic resources with machine learning algorithms (CRF and LR) to improve the performance of ABSA task. First, (Brun 2018) extracted lexical, syntactic and semantic information from the dataset using a robust parser and the CRF model. Afterwards, they classified the sentiment towards aspects applying the LR model. (García-Pablos et al. 2018) constructed the W2VLDA model in order to perform the ABSA based on the topic modelling model LDA. This model was combined with the Maximum Entropy (ME) classifier, seed words and word embeddings to detect the sentiments expressed towards aspects. (Akhtar et al. 2019) adopted a Bi-LSTM model to perform the ABSA task. The authors incorporated word embeddings and manually crafted features into the Bi-LSTM model to improve the accuracy of classifying sentiments. Moreover, (Essebbar et al. 2021) applied three multi-lingual pre-trained models (multilingual BERT, CamemBERT and FlauBERT) in ABSA. These models were, then, fine-tuned using the collected dataset and three methods, namely fully connected method, Sentences Pair Classification and Attention Encoder Network. Finally, they were applied to classify the sentiment associated with each aspect. As shown in Table 8, the proposed architecture outperformed the existing models when applied on the French SemEval-2016 restaurant dataset. It can be considered as an enhancement of the studies conducted by (Ruder et al. 2016), (Brun et al. 2016), (Kumar et al. 2016), (Akhtar et al. 2017), (Pigneul and Kooli 2018), (Brun 2018),

**Table 8** Comparison between the CBRS architecture and benchmarks

|  | Study | Accuracy (%) |
|---|---|---|
| SemEval-2016 restaurant Dataset | (Ruder et al. 2016) | 73.16 |
|  | (Brun et al. 2016) | 78.82 |
|  | (Kumar et al. 2016) | 72.20 |
|  | (Akhtar et al. 2017) | 70.91 |
|  | (Pigneul and Kooli 2018) | 74.23 |
|  | (Brun 2018) | 78.83 |
|  | (García-Pablos et al. 2018) | 62.80 |
|  | (Akhtar et al. 2019) | 75.30 |
|  | (Essebbar et al. 2021) | 84.06 |
|  | CBRS architecture | 85.70 |

**Table 9** Comparison between the CBRS architecture and English benchmarks

| Study | SemEval-2016 English Restaurant dataset | SemEval-2016 French Restaurant dataset |
|---|---|---|
| (Ramaswamy and Chinnappan 2022) | 84.81% | – |
| (Ayetiran 2022) | 74.44% | – |
| CBRS architecture | – | 85.70% |

(García-Pablos et al. 2018), (Akhtar et al. 2019) and (Essebbar et al. 2021) by approximately 12, 6, 13, 15, 11, 7, 23, 10 and 2% respectively.

Additionally, given the significant advancements of the ABSA in the English language, the results provided by the method suggested in the present study were compared to those obtained by similar methodologies applied on datasets written in the same language Among these studies, we cite the work of (Ramaswamy and Chinnappan 2022) who proposed an enhanced LSTM-CNN architecture for aspect-level sentiment classification. To implement this architecture, the knowledge database RecogNet was first used to extract semantic and sentiment information. This information was, then, integrated into the LSTM model to capture long-term dependencies between words. After that, the CNN model with an attention mechanism was employed to learn local features. These features were, subsequently, incorporated into the output layer to classify the sentiments towards aspects. The performance of the proposed architecture was evaluated on the SemEval-2014 English restaurant dataset (Pontiki et al. 2016) for the ABSA task and achieved the state-of-the-art results. (Ayetiran 2022) suggested a novel attention-based CNN-Bi-LSTM model to classify the sentiments. The model initially employed a CNN to extract local features from the sentence. Subsequently, Bi-LSTM was applied to capture the long-term dependencies between words. Finally, an attention mechanism was utilized to detect the most important words in sentiment classification.

As shown in Table 9, the architecture developed in the present work achieved results that are comparable to those provided in the state-of-the-art studies.

# 7 Conclusion

In this research paper, a machine learning-based method applied in ABSA was set forward. The smartphone dataset, composed of 8000 Amazon comments, was first annotated. It was pre-processed by removing stop words, stemming, treating emoticons, etc. Subsequently, a new architecture, called CBRS (CNN-Bi-RNN-SVM) and combining the strengths of the deep learning models (CNN and Bi-RNN)

with the classical machine learning model (SVM) was introduced. The CNN model was utilized to capture local features in the textual data. It employed convolutional layers to extract relevant features by convolving filters over the input data. Then, the Bi-RNN model was used to learn contextual information from the input text. It consisted of two recurrent layers, each of which processed the input text in forward direction and backward direction. On the other hand, the SVM model was used to classify the sentiments towards aspects and determine whether an aspect term expresses a positive or negative sentiment.

This introduced method provided interesting F-measure values equal to 94.04 and to 85.70% for the smartphone dataset and the Semeval-2016 restaurant dataset, respectively. In fact, it outperformed the deep learning models (Bi-RNN as well as CNN) and the classical machine learning model (SVM) with 8%, 4% and 27%, respectively. It was, therefore, obvious that the results given by combining the machine learning models were better than those provided by using each model separately. In addition, the developed method was compared to a set of benchmark methods. The obtained findings corroborated the effectiveness of the proposed architecture that outperformed the other benchmarks with impressive values of F-measure and accuracy.

In conclusion, the suggested machine learning-based method promoted the ABSA task for the French language. It exhibited high ability in capturing contextual and local features as well as remarkable robust classification capabilities, which makes it efficiently applied in the field of sentiment analysis. This method did not only achieve high accuracy in the ABSA task, but it also displayed an excellent performance in handling various datasets and domains.

Therefore, we can assert that the current research work may be extended and improved, in the future, by increasing the number of the annotated comments. In addition, we intend to promote the machine learning-based method through incorporating new features including POS, syntactic and semantic features. We will also integrate attention mechanism within the CBRS architecture to make it focus on relevant parts of the input data. This mechanism allows assigning more weight to important word and significantly improving the architecture's ability to capture fine-grained sentiment information. Furthermore, we will enhance the introduced architecture using transfer learning techniques such as leveraging pre-trained models on large-scale sentiment-related tasks.

## Declarations

**Conflict of interest** The authors declare that they have no competing interests.

**Ethical approval** The current manuscript complies with the Ethical Rules applicable to this journal. This project is part of a thesis. There is no conflict between the authors and everyone is involved in this project according to their roles fixed from the start.

## References

Akhtar MS, Kohail S, Kumar A, Ekbal A, Biemann C (2017) Feature selection using multi-objective optimization for aspect based sentiment analysis. In: Natural language processing and information systems: 22nd international conference on applications of natural language to information systems, NLDB 2017, Liège, Belgium, June 21--23, 2017, Proceedings 22, pp. 15–27. Springer

Akhtar MS, Kumar A, Ekbal A, Biemann C, Bhattacharyya P (2019) Language-agnostic model for aspect-based sentiment analysis. In: Proceedings of the 13th international conference on computational semantics-long papers, pp 154–164

Al-Smadi M, Al-Ayyoub M, Jararweh Y, Qawasmeh O (2019) Enhancing aspect-based sentiment analysis of arabic hotels' reviews using morphological, syntactic and semantic features. Inf Proces Manag 56(2):308–319

AlAjlan SA, Saudagar AKJ (2021) Machine learning approach for threat detection on social media posts containing arabic text. Evol Intell 14(2):811–822

Alharbi O (2021) A deep learning approach combining cnn and bi-lstm with svm classifier for arabic sentiment analysis. Int J Adv Comput Sci Appl 12(6)

Alqaryouti O, Siyam N, Abdel Monem A, Shaalan K (2020) Aspect-based sentiment analysis using smart government review data. Appl Comput Inf

Althobaiti M, Kruschwitz U, Poesio M Aranlp (2014) A java-based library for the processing of arabic text. In: Proceedings of the 9th international conference on language resources and evaluation, LREC 2014, pp 4134–4138. European Language Resources Association (ELRA)

Angiani G, Ferrari L, Fontanini T, Fornacciari P, Iotti E, Magliani F, Manicardi S (2016) A comparison between preprocessing techniques for sentiment analysis in twitter. In: KDWeb

Aubaid AM, Mishra A (2020) A rule-based approach to embedding techniques for text document classification. Appl Sci 10(11):4009

Ayetiran EF (2022) Attention-based aspect sentiment classification using enhanced learning through cnn-bilstm networks. Knowl Based Syst 252:109409

Banjar A, Ahmed Z, Daud A, Abbasi RA, Dawood H (2021) Aspect-based sentiment analysis for polarity estimation of customer reviews on twitter. Comput Mater Continua 67(2):2203–2225

Brauwers G, Frasincar F (2022) A survey on aspect-based sentiment classification. ACM Comput Surv 55(4):1–37

Brun C, Perez J, Roux C (2016) Xrce at semeval-2016 task 5: Feed-backed ensemble modeling on syntactico-semantic knowledge for aspect based sentiment analysis. In: Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016), pp 277–281

Brun C (2018) Transfert de ressources sémantiques pour l'analyse de sentiments au niveau des aspects. In: Actes de la Conférence Traitement Automatique de la Langue Naturelle, TALN 2018, p 547

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297

Deligiannidis S, Mesaritakis C, Bogris A (2021) Performance and complexity analysis of bi-directional recurrent neural network models versus volterra nonlinear equalizers in digital coherent systems. J Lightwave Technol 39(18):5791–5798

Ellouze M, Hadrich L (2022) A hybrid approach for the detection and monitoring of people having personality disorders on social networks. Social Netw Anal Min 12(1):67

Essebbar A, Kane B, Guinaudeau O, Chiesa V, Quénel I, Chau S (2021) Aspect based sentiment analysis using french pre-trained models. In: ICAART (1), pp 519–525

Fukushima K (1980) Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 36(4):193–202

García-Pablos A, Cuadros M, Rigau G (2018) W2vlda: almost unsupervised system for aspect based sentiment analysis. Exp Syst Appl 91:127–137

Graves A (2013) Generating sequences with recurrent neural networks. arXiv preprint arXiv:1308.0850

Grewal D, Roggeveen A, Sisodia R, Nordfält J (2017) Enhancing customer engagement through consciousness. J Retail 93(1):55–64

Hamdan H, Bellot P, Bechet F Lsislif (2015) Crf and logistic regression for opinion target extraction and sentiment polarity analysis. In: Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015), pp 753–758

Hammi S, Hammami SM, Belguith LH (2022) Aspect term extraction improvement based on a hybrid method. In: International symposium on methodologies for intelligent systems, pp 85–94 Springer

Hammi S, Hammami SM, Belguith LH (2022) An improved hybrid method for sentiment analysis. In: 2022 International conference on innovations in intelligent systems and applications (INISTA), pp 1–6. IEEE

Hidaka A, Kurita T (2017) Consecutive dimensionality reduction by canonical correlation analysis for visualization of convolutional neural networks. In: Proceedings of the ISCIE international symposium on stochastic systems theory and its applications, vol 2017, pp 160–167

Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining, pp 168–177

Jagtap V, Pawar K (2013) Analysis of different approaches to sentence-level sentiment classification. Int J Sci Eng Technol 2(3):164–170

Jivani AG et al (2011) A comparative study of stemming algorithms. Int J Comp Tech Appl 2(6):1930–1938

Kumar V, Sundaram S (2022) Offline text-independent writer identification based on word level data. arXiv preprint arXiv:2202.10207

Kumar A, Kohail S, Kumar A, Ekbal A, Biemann C (2016) Iit-tuda at semeval-2016 task 5: Beyond sentiment lexicon: combining domain dependency and distributional semantics features for aspect based sentiment analysis. In: Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016), pp 1129–1135

Liang B, Su H, Gui L, Cambria E, Xu R (2022) Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. Knowl Based Syst 235:107643

Liu C, Ta V, Le N, Tadesse DA, Shi C (2022) Deep neural network framework based on word embedding for protein glutarylation sites prediction. Life 12(8):1213

Medhat W, Hassan A, Korashy H (2014) Sentiment analysis algorithms and applications: a survey. Ain Shams Eng J 5(4):1093–1113

Moghaddam, S., Ester, M.: Opinion digger: an unsupervised opinion miner from unstructured product reviews. In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management, pp. 1825–1828 (2010)

Mubarok MS, Adiwijaya A, Aldhi MD (2017) Aspect-based sentiment analysis to review products using naïve bayes. In: AIP conference proceedings, vol 1867

Nasr L, Burton J, Gruber T, Kitshoff J (2014) Exploring the impact of customer feedback on the well-being of service entities: a tsr perspective. J Serv Manage 25(4):531–555

Palomino MA, Aider F (2022) Evaluating the effectiveness of text pre-processing in sentiment analysis. Appl Sci 12(17):8765

Pang B, Lee L, Vaithyanathan S (2002) Thumbs up? sentiment classification using machine learning techniques. arXiv preprint cs/0205070

Pennington J, Socher R, Manning CD (2014) Glove: global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP), pp 1532–1543

Pigneul N, Kooli E (2018) Analyse de sentiments à base d'aspects par combinaison de réseaux profonds: application à des avis en français

Piryani R, Gupta V, Singh VK, Ghose U (2017) A linguistic rule-based approach for aspect-level sentiment analysis of movie reviews. In: Advances in computer and computational sciences: proceedings of ICCCCS 2016, vol 1, pp 201–209

PontikiM, GalanisD, Papageorgiou H, Androutsopoulos I, Manandhar S, AL-Smadi M, Al-Ayyoub M, Zhao Y, Qin B, Clercq OD et al. (2016) Semeval-2016 task 5: Aspect based sentiment analysis. In: Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016), pp 19–30. Association for Computational Linguistics

Potisuk S (2010) Typed dependency relations for syntactic analysis of thai sentences. In: Proceedings of the 24th Pacific Asia conference on language, information and computation, pp 511–518

Ramaswamy SL, Chinnappan J (2022) Recognet-lstm+ cnn: a hybrid network with attention mechanism for aspect categorization and sentiment classification. J Intell Inf Syst 58(2):379–404

Ray P, Chakrabarti A (2022) A mixed approach of deep learning method and rule-based method to improve aspect level sentiment analysis. Appl Comput Inf 18(1/2):163–178

Rizwan A, Iqbal N, Ahmad R, Kim D (2021) Wr-svm model based on the margin radius approach for solving the minimum enclosing ball problem in support vector machine classification. Appl Sci 11(10):4657

Ruder S, Ghaffari P, Breslin JG (2016) Insight-1 at semeval-2016 task 5: deep learning for multilingual aspect-based sentiment analysis. arXiv preprint arXiv:1609.02748

Ruder S, Ghaffari P, Breslin JG (2016) A hierarchical model of reviews for aspect-based sentiment analysis. arXiv preprint arXiv:1609.02745

Sarker IH (2021) Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. SN Comput Sci 2(6):420

Straka M, Straková J (2017) Tokenizing, pos tagging, lemmatizing and parsing ud 2.0 with udpipe. In: Proceedings of the CoNLL 2017 shared task: multilingual parsing from raw text to universal dependencies, pp 88–99

Taye MM (2023) Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. Computers 12(5):91

Vanaja S, Belwal M (2018) Aspect-level sentiment analysis on e-commerce data. In: 2018 International conference on inventive research in computing applications (ICIRCA), pp 1275–1279

Villena RPCSL, Cristóbal J (2011) Hybrid approach combining machine learning and a rule-based expert system for text categorization. AAAI

Wang J, Li J, Li S, Kang Y, Zhang M, SiL, Zhou G (2018) Aspect sentiment classification with both word-level and clause-level attention networks. In: IJCAI, vol 2018, pp. 4439–4445

Yan Z, Xing M, zhang D (2015) Exprs: An extended pagerank method for product feature extraction from online consumer reviews. Inf Manage 52(7)

Yao L, Mao C, Luo Y (2019) Clinical text classification with rule-based features and knowledge-guided convolutional neural networks. BMC Med Inf Decis Making 19(3):31–39

Zhao Y, Mamat M, Aysa A, Ubul K (2023) Multimodal sentiment system and method based on crnn-svm. Neural Comput App, pp 1–13

Zhao Y, Mamat M, Aysa A, Ubul K (2023) Multimodal sentiment system and method based on crnn-svm. Neural Comput Appl, pp 1–13