Electronic Research Archive

ERA, 31(6): 3495–3509.
DOI: 10.3934/era.2023177
Received: 08 January 2023
Revised: 27 March 2023
Accepted: 06 April 2023
Published: 20 April 2023

http://www.aimspress.com/journal/ERA

Research article

Sentence opinion mining model for fusing target entities in official government documents

Xiao Ma[1], Teng Yang[2], Feng Bai[2] and Yunmei Shi[2,*]

[1] Department of Big Data Development, The State Information Center, Beijing 100045, China
[2] School of Computer Science, Beijing Information Science and Technology University, Beijing 100101, China

* **Correspondence:** Email: sym@bistu.edu.cn; Tel: +8613261853839.

Abstract: When drafting official government documents, it is necessary to firmly grasp the main idea and ensure that any positions stated within the text are consistent with those in previous documents. In combination with the field's demands, By taking advantage of suitable text-mining techniques to harvest opinions from sentences in official government documents, the efficiency of official government document writers can be significantly increased. Most existing opinion mining approaches employ text classification methods to directly mine the sentential text of official government documents while disregarding the influence of the objects described within the documents (i.e., the target entities) on the sentence opinion categories. To address these issues, this study proposes a sentence opinion mining model that fuses the target entities within documents. Based on the Bi-directional long short-term (BiLSTM) and attention mechanisms, the model fully considers the attention given by a official government document's target entity to different words within the corresponding sentence text, as well as the dependency between words of the sentence. The model subsequently fuses two by using feature vector fusion to obtain the final semantic representation of the text, which is then classified using a fully connected network and softmax function. Experimental results based on a dataset of official government documents show that the model significantly outperforms baseline models such as Text-convolutional neural network (TextCNN), recurrent neural network (RNN), and BiLSTM.

Keywords: sentence opinion mining; target entity; BiLSTM; attention mechanism

1. Introduction

Official government documents are written materials that follow a specific style and undergo certain processing procedures to be formulated and used by statutory bodies and organizations in their official activities. Government functionaries engaging in professional work or administrative services must learn to convey governmental policies and handle official matters through official government documents, thereby ensuring accuracy and efficiency in coordinating various relationships and decision-making tasks [1]. With the emergence of Internet-based technologies, paperless office work has become a mainstream practice in the modern workplace. The increasing prevalence of electronic documents reduces the storage problems associated with official government documents, further improving the speed at which government agencies handle official duties [2].

As tools for exchanging information between various governmental agencies, official government documents reflect the authority exercised by these agencies. When drafting official government documents, government functionaries must understand the main idea of each document to ensure that it does not conflict with documents already issued by higher authorities or departments. Therefore, to maintain consistency, it is crucial to refer to related documents previously issued by higher authorities, issuing departments, and other related departments at the same level [3,4]. However, it is time-consuming and laborious to perform a manual comparative search. Accordingly, this study aimed to address the above core problem by constructing a sentence opinion-mining model to help official document writers to accurately and quickly judge the consistency of sentence opinion and maintain the continuity and consistency of policies.

Opinion mining, also known as tendency or sentiment analysis [5], is essentially a text classification task. Therefore, it can be conducted via traditional machine learning or deep learning.

Traditional machine learning-based text classification algorithms, such as K-nearest neighbor [6] support vector machine [7], and naive Bayes [8], require manual labeling of textual features and are prone to issues such as excessively high dimensionality and local optimality. Furthermore, parametric variability significantly impacts classification results, the robustness of the algorithms is poor, and the effectiveness of text classification is suboptimal.

With the development and application of deep learning technology in the fields of image processing and speech recognition, various text classification algorithms based on deep learning have recently emerged. Deep learning-based neural network models include the convolutional neural network (CNN), recurrent neural network (RNN), and many others. In the task of text classification, deep learning models can not only automatically learn and extract the key features of text, but also involve relatively few parameters in the training process of multilevel network, which can avoid the phenomenon of overfitting.Bi-directional long short-term memory (BiLSTM) [9] is a neural network model relatively widely used in the field of text classification. However, BiLSTM is not sufficiently competent in text feature extraction and cannot account for the local dependencies among textual data.

The attention mechanism is a mechanism to improve the observation accuracy of a specific region. It can selectively focus on certain parts of the observation area, and can quickly extract the key features of sparse data.The self-attention mechanism, a special variant of the attention mechanism, is more adept at obtaining correlations within the data and performs well in extracting local key information from text. Therefore, by combining self-attention and BiLSTM [10,11], the disadvantages of BiLSTM in feature extraction can be supplemented appropriately.

In the case of opinion mining in official government documents, the opinion category of a sentence is often closely related to the target entity discussed in the article. Consider the following sentence: "Crackdown on *domestic violence* and provide assistance to *victims of domestic violence*". If the target entity is "domestic violence", the corresponding opinion category should be regarded as negative, indicating that the government is opposed to domestic violence. However, if the target entity is "victims of domestic violence", the opinion category is instead regarded as positive, indicating that the government will help the victims of domestic violence and punish the perpetrators accordingly. Therefore, it is important to consider the semantics of the sentence text as well as the influence exerted by the target entities.

The purpose of this paper is to help document writers to accurately and quickly judge the opinions of different materials. Owing to the wide range of fields, complex structure, and relatively concise language of official government documents, existing article writing tools do not have a sufficient sentence semantic analysis function. This is not conducive to improving the efficiency of official government document writing and also brings challenges to the maintenance of policy continuity and consistency. Therefore, based on BiLSTM and the Attention mechanism, this study combined the description object (i.e., target entities) of the official government document with the text semantics of the document, considering the correlation within the text and correlation between the target entity and different words in the text, and then used feature fusion technology to obtain the final semantic representation of the text. This paper proposes an opinion mining model based on specific target entity. The experimental results show that the model achieved good performance in the task of sentence opinion mining in official government documents. The main contributions of this paper are as follows:

- This paper proposes a sentence opinion mining model fused with target entity for official government documents. The model combines Self-Attention and BiLSTM to overcome the shortcomings of each. As such, the model can not only consider the dependence between words in the text but also analyze the attention of the Target entity to different words in the text.
- Feature fusion technology is applied to the field of official government document analysis, so that the semantic information contained in document sentences can be more comprehensively expressed, and the results of sentence opinion mining are more accurate.

Within the paper, relevant prior studies concerning text classification are summarized in Section 2; the structure and principles of the opinion mining model are introduced in Section 3; the dataset and experimental setup are presented in Section 4; experimental results are analyzed in Section 5; the final section concludes the paper and provides relevant future prospects.

2. Literature review

Opinion mining, which essentially entails the classification of sentence text, may be conducted via machine learning or deep learning. Traditional machine learning models have relatively simple structures and rely primarily on text features obtained manually. Although such models require relatively few parameters, they often produce satisfactory results in complex tasks and are highly adaptable across multiple fields. Some of the most used algorithms are the decision tree (DT), support vector machine (SVM), K-nearest neighbor (KNN), and naive Bayes (NB).

As feature engineering is a crucial component of machine learning, selecting appropriate features or adding relevant external features are common methods to enhance machine learning models. For example, Li et al. [12] proposed a sentiment analysis method based on SVM and conditional random field (CRF) by generating different combinations of text features and selecting the optimal combination from experimental results. Goel et al. [13] further improved the accuracy of the conventional Bayesian classification model by optimizing the naive Bayesian algorithm using the WordNet sentiment lexicon. In addition, many researchers have investigated the strengths and weaknesses of different machine learning algorithms. In 2019, Ababneh [14] examined the performance of three KNN, DT, and NB on the Saudi Press Agency dataset and discovered that the NB algorithm outperformed the other algorithms in terms of accuracy, recall rate, and F1 score. In 2022, Ahuja et al. [15] compared the performance of six different machine learning algorithms (SMV, logistic regression, KNN, random forest, Bayesian, and DT) on datasets from three different fields, concluding that the best-performing models were logistic regression and the naive Bayesian algorithm. Notably, the strengths and weaknesses of machine learning algorithms often depend on selected features and the specific field of study. Therefore, although machine learning methods are highly generalizable, they are often demanding in terms of labor and are not applicable when a large corpus is involved.

Deep-learning-based neural network models have recently become popular methods for text classification. These methods can be classified as CNN-based, RNN-based, or attention-based.

CNNs, first employed for image classification, contain convolutional filters that extract local data features. These networks can simultaneously perform convolutional operations on text sequences using different convolutional kernels to obtain corresponding n-gram features. By proposing the TextCNN model in 2014, Kim [16] instantly established the presence of CNNs in the field of natural language processing. Consequently, Johnson et al. [17] proposed a low-complexity word-level deep CNN architecture for text classification called deep pyramidal CNN (DPCNN) in 2017. DPCNN can effectively represent remote associations in text. In 2021, Yu et al. [18] proposed a novel deep CNN for short text classification, known as the deep pyramidal temporal convolutional network (DPTCN) as it was inspired by the temporal convolutional network and DPCNN.

Although CNNs can obtain the local features of text sequences through convolution and pooling operations, they cannot obtain the long-distance dependencies between words in longer sentences. However, RNNs avoid this problem. RNNs, widely employed in the field of natural language processing, are characterized by a unique time sequence processing approach consistent with human text reading, giving them the ability to learn historical and local information of input sequences effectively. Wang [19] proposed an RNN-based capsule model called RNN-Capsule for sentiment analysis in 2018 and used it to achieve optimal performance on several datasets. In 2019, Wang [20] proposed a novel method that combines the strengths of RNNs and CNNs by first obtaining the word representation of text through a bi-directional RNN and then learning the importance of each word with respect to text classification using a CNN. In general, the initial and control parameter settings of RNNs, which are selected based on trial and error, can significantly affect model performance. Accordingly, Singh et al. [21] proposed an evolutionary LSTM network (ELSTM) and optimized its architecture and weights using a multi-objective genetic algorithm, thereby addressing the parameter tuning problem of LSTM and enhancing model performance.

As the presence of hidden states in RNNs and CNNs incurs a certain degree of uninterpretability in the models, the attention mechanism has emerged as an alternative solution. In 2015, Bahdanau et al. [22] from Jacobs University in Bremen, Germany, first proposed an attention mechanism for a machine translation task, achieving good results. Subsequently, many researchers began applying the attention mechanism to text classification and sentiment analysis tasks. In 2016, Wang et al. [23] fused the LSTM and attention mechanisms to propose the ATAE-LSTM model structure for aspect-level fine-grained sentiment classification tasks. This approach, which incorporates aspect words into the model's training process, has been highly inspirational for the work presented in this paper. In 2022, Liu et al. [24] used Probase as an external knowledge source to enrich semantic representation and combined contextually relevant features with a multi-stage attention model based on a temporal convolutional network (TCN) and CNN to develop a novel short text classification method known as CRFA, which effectively addresses the issues of data sparsity and ambiguity of short texts. In 2019, Hu et al. [25] used BiLSTM to train the topic and text word vectors and performed feature fusion on the obtained topic and text features, which were then processed with the deep attention mechanism to enhance the model's focus on topic and text information.

Presently, most solutions for opinion mining tasks are based on deep learning. In contrast to traditional machine learning methods, deep learning models have relatively complex structures, do not rely on manually obtained text features, and can learn and model text content directly. However, these models are highly dependent on data and exhibit poor adaptability across multiple fields. As the sentence opinion mining task for official government documents depends on specialized data within the corresponding field, it is necessary to adopt deep learning technologies and design a dedicated opinion mining model targeting the characteristics of the official government document corpus. Given these characteristics, the method proposed in this paper adopts a feature fusion approach to fuse target entities into a semantic representation of sentence texts, strengthening the connection between sentence contents and target entities and yielding a significant improvement in the opinion mining model's performance.

3. Sentence opinion mining model for fusing target entities

The model proposed in this paper comprises several steps during execution. First, the word2vec word embedding technology is used to model the *Target* entity and sentence *Text*. The advantage of word2vec is that it can learn the relationship between words, such as their semantic and grammatical relationship, and has an advantage in training speed compared with Glove and other similar methods [26]. Thus, it is suitable for application in the proposed model, with word embedding used as input to obtain the hidden states of *Target* and *Text* through BiLSTM. Then, the attention mechanism collects important information from *Text* and *Target* to obtain the self-attention representation (self-representation) of *Text* and attention representation (target representation) between *Target* and *Text*. Next, the two vectors are fused using feature fusion to obtain the final representation. The final representation accounts for the correlation within *Text* and fuses the attention levels given to different words in *Text* by *Target*, thereby producing a more accurate representation of semantic information in official government document sentences, make the result of sentences opinion mining more accurate. The final representation is then fed into the fully connected layers and softmax function to obtain the final sentence opinion mining results. The overall structure of the model is illustrated in Figure 1.

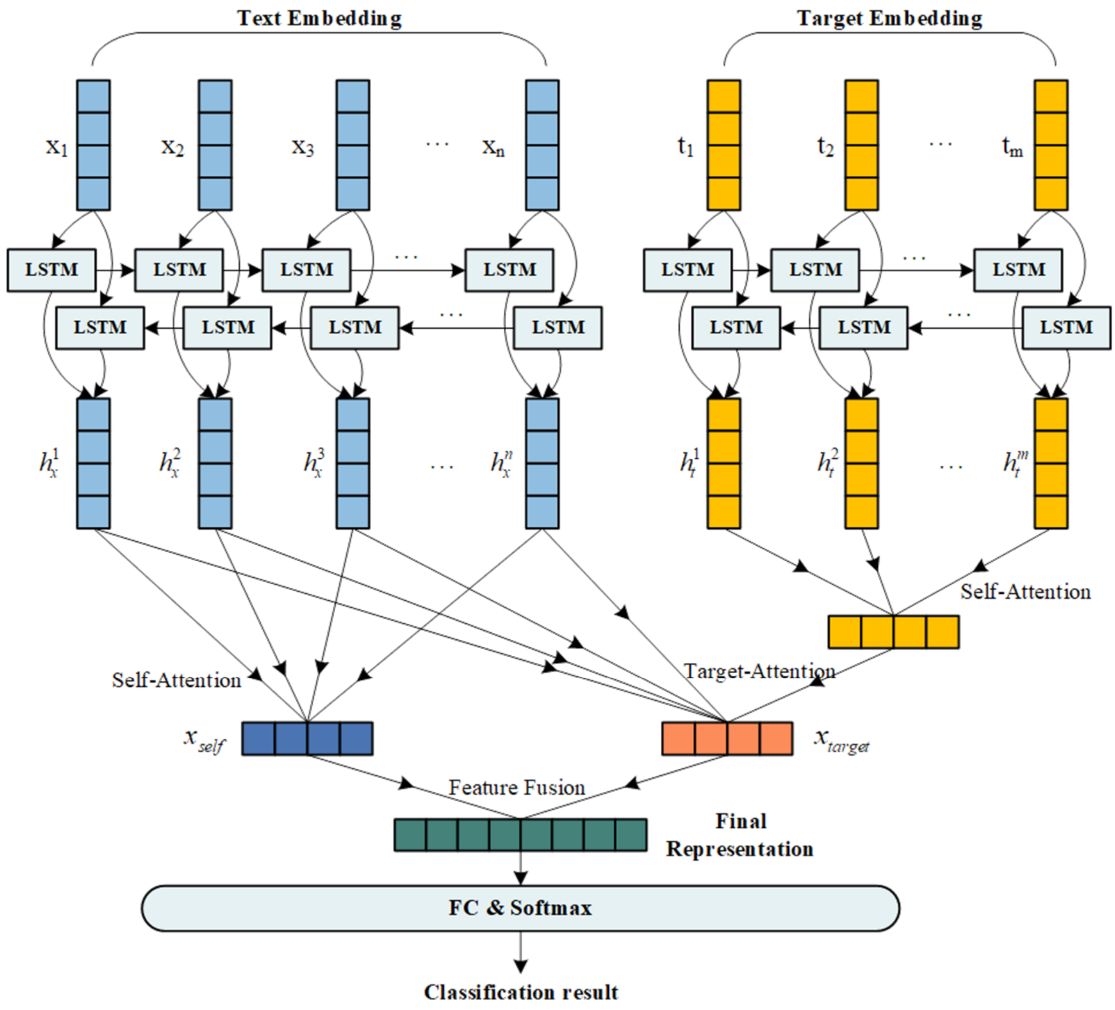

Figure 1. Structure of opinion mining model.

First, assume that the sentence text "*Text*" consists of n words $[\mathrm{w}_x^1, w_x^2, \cdots, \mathrm{w}_x^n]$, and the target entity "*Target*" consists of m words $[w_t^1, w_t^2, \cdots, w_t^m]$. Then, each word is mapped into a k-dimensional real space using the word embedding technology to obtain the vector representations of *Text* and *Target*, respectively: $X = [x_1, x_2, \cdots, x_n]$ and $T = [t_1, t_2, \cdots, t_m]$, where $x_i, t_i \in R^k$.

Then, the word vector is fed into BiLSTM to obtain the hidden state representations of *Text* and *Target*: $H_x = [\mathrm{h}_x^1, h_x^2, \cdots, \mathrm{h}_x^n] \in R^{n\times 2*h}, H_t = [h_t^1, h_t^2, \cdots, h_t^m] \in R^{m\times 2*h}$, where $h_x^i, h_t^i \in R^h$, h is the dimension of the hidden state.

$$H_x = BiLSTM(X) \tag{1}$$

$$H_t = BiLSTM(T) \tag{2}$$

Next, obtain the dependencies between the words of *Target* and *Text* using the self-attention mechanism to derive the self-attentive representations t_{self} and x_{self}, respectively. t_{self} represents the semantic representation of *Target* involved in the next stage of the target representation calculation, whereas x_{self} is involved in the final feature fusion as the semantic representation of sentence text. The corresponding computational formula is shown in Eqs (3) and (4).

$$\mathrm{T_{self}} = \mathrm{avg}\big(\mathrm{selfAttention}(\mathrm{H_t})\big) \tag{3}$$

$$x_{self} = avg\big(selfAttention(H_x)\big) \tag{4}$$

where $t_{self}, \mathrm{x_{self}} \in R^{2*h}$, $avg(*)$ is the averaging function, and $selfAttention(*)$ is the self-attention layer, whose output is a matrix with the same dimension as the input. This matrix is calculated as

$$\mathrm{Q} = \mathrm{f}\big(\mathrm{W_Q I} + \mathrm{b_Q}\big) \tag{5}$$

$$K = \mathrm{f}(\mathrm{W_K I} + \mathrm{b_K}) \tag{6}$$

$$V = \mathrm{f}(\mathrm{W_V I} + \mathrm{b_V}) \tag{7}$$

$$\mathrm{selfAttention\ (I)} = \mathrm{softmax}\left(\frac{\mathrm{QK^T}}{\sqrt{\mathrm{d_k}}}\right)\mathrm{V} \tag{8}$$

where I is the input matrix; W_Q, W_K, and W_V are the weight matrices; b_Q, b_K and b_V are the bias values; $f(\cdot)$ is the nonlinear activation function; d_K is the dimensionality of K.

This step is followed by calculating the Target representation $\mathrm{x_{target}}$ of *Text*, capturing the attention level of *Target* to different words in *Text*, i.e., the attention level t_{self} to different hidden states in H_x. The attention score based on the attention level is then calculated, with a vector representation $x_{target} \in R^{2*h}$ obtained by weighted summation, where x_{target} contains the part of semantic information most relevant to *Target*, as in Eq (9).

$$x_{target} = softmax\big(t_{self} \cdot H_x^T\big) \cdot H_x \tag{9}$$

where $H_x^T \in R^{h\times n}$ is the transpose matrix of H_x.

Lastly, x_{self} and x_{target} are fused into a vector using feature fusion technology, i.e., the semantic information that *Text* itself is more attentive to and the semantic information that *Target* is more attentive to fused toobtain the final representation "$final$". This representation is fed into the fully connected and $softmax$ layers to obtain the final sentence opinion classification results, as shown in Eqs (10) and (11).

$$final = \Phi\big(\mathrm{x_{self}}, \mathrm{x_{target}}\big) \tag{10}$$

$$result = softmax\big(fc(final)\big) \tag{11}$$

where $\Phi(*)$ is the feature fusion function and $fc(*)$ is the fully connected layer.

4. Experimental data and evaluation criteria

4.1. Official government documents dataset

The objective of this experiment was to perform opinion mining on sentences in official government documents and determine the types of opinions held within the sentences for specific target entities. The corpus used in the experiment comes from official government documents obtained from nationwide government websites, with a total of 1995 experimental data remaining after sentence segmentation. The training, validation, and test sets were divided according to a 3:1:1 ratio, with the numbers of positive and negative samples in each dataset equalized. Dataset composition listed in Table 1.

Table 1. Dataset composition.

	Positive	Negative	Total
Training set	651	546	1197
Validation set	217	182	399
Test set	217	182	399
Total	1085	910	1995

Each input data consists of three components, as listed in Table 2: sentence, target entity, and opinion category: the sentence is the text from which opinions are to be mined; the target entity is the object of description in the official government document, which may be a word or phrase, obtained from the document's title via named entity identification; the opinion category indicates the opinion and attitude expressed by the sentence towards the target entity, with the positive category corresponding supportive or helpful attitudes toward the target entity and the negative attitude associated with opposing or forbidding attitudes toward the target entity.

These opinion categories require manual labeling. The sentences belonging to the same official government document should have the same opinion about the target entity because they have the same target entity. Therefore, the same label is used for all corresponding sentences within a document.

Table 2. Examples of training data structure.

Sentence	Target entity	Opinion category
People's governments at or above the county level should take measures to provide temporary assistance to people wandering on the street and beggars, such as providing food and lodging, emergency medical treatment, and assistance in returning to their homes.	People wandering on the street and beggars	Positive
The various public service facilities and places shall not accommodate any illegal societal organizations.	Illegal societal organizations	Negative

4.2. Public dataset

To verify the generalization ability of the proposed model on other domain datasets, we applied it to the Aspect Category Sentiment Analysis (ACSA) task, which is similar to the task considered in this study, and conducted experiments on the public ASAP dataset [27]. The dataset covers 46,730 starred restaurant user reviews from Chinese e-commerce platforms, each manually annotated into sentiment categories across 18 fine-grained aspect categories. To meet the data input requirements of this model, we reorganized the data and generated 270,834 experimental data.

4.3. Evaluation criteria

An experiment was conducted to evaluate the opinion mining model's performance using the accuracy (Acc) and F1 score metrics commonly used for classification tasks.

Acc [28]: Measures the accuracy of the predicted result or the number of correctly predicted samples divided by the total number of samples. The calculation formula is expressed in Eq (12), where $y_i^{predict}$ and y_i^{true} denote the predicted and actual values, respectively, n is the total number of samples, and $I(\cdot)$ is the indicator function.

$$Acc = \frac{\sum_{i=1}^{n} I\left(y_i^{predict} = y_i^{true}\right)}{n} \tag{12}$$

F1 value [28]: Supposing 10 samples to be tested (with 9 positive and 1 negative), the model can achieve 90% accuracy by simply predicting all cases as positive. Therefore, using Acc as the sole performance metric is unreasonable. In contrast, the F1 score accounts for both the precision rate P and recall rate R of the classification model, making it an effective evaluation metric. The precision rate denotes the probability of correctly predicting a positive sample among all samples predicted to be positive, whereas the recall rate indicates the probability of being correctly predicted as a positive sample among all positive samples of the original sample. The F1 score, regarded as a kind of weighted average between model accuracy rate and recall rate, is calculated according to Eqs (13)–(15):

$$P = \frac{TP}{TP+FP} \tag{13}$$

$$R = \frac{TP}{TP+FN} \tag{14}$$

$$F_1 = \frac{2*P*R}{P+R} \tag{15}$$

where TP, FP, and FN denote the number of true positives, false positives, and false negatives, respectively.

5. Experimental evaluation

5.1. Hyperparameter settings

The model's hyperparameters exert a certain influence over the effectiveness of opinion mining. Three parameters—embedding_size, hidden_num, and dropout_rate—are particularly important and were optimized through numerous experiments. These parameters determine the input and output dimensions of the BiLSTM layers and the probability of dropping each neuron in the last fully connected layer, respectively. The experimental results listed in Table 3 indicate that the model achieves optimal performance when the three hyperparameters are set to 512, 256, and 0.4, with Acc and F1 scores reaching 0.9447 and 0.9484, respectively.

Table 3. Hyperparameter settings.

Embedding_size	Hidden_num	Dropout_rate	Acc	F1
256	256	0.4	0.9397	0.9450
256	256	0.5	0.9095	0.9193
256	512	0.4	0.9347	0.9390
256	512	0.5	0.9171	0.9177
512	256	0.4	**0.9447**	**0.9484**
512	256	0.5	0.8995	0.9061
512	512	0.4	0.9246	0.9309
512	512	0.5	0.8693	0.8785

5.2. Feature fusion methods

Two different feature fusion methods were used with the proposed opinion mining model developed in this study: addition of vectors (Add) and concatenation of vectors (Concatenate). A comparison was performed experimentally to determine the fusion method for self-representation x_{self} and target representation x_{target}, with results listed in Table 4.

Table 4. Feature fusion methods.

Method	Acc	Macro-F1
Add	0.9095	0.9192
Concatenate	**0.9447**	**0.9484**

Clearly, the Concatenate method exhibits a significant advantage over the Add method, primarily because the latter confuses two different feature representations, making it difficult for the model to obtain useful information. Conversely, the Concatenate method allows the model to obtain the required information by changing its parameters. Therefore, Concatenate was adopted as the feature fusion method in the proposed opinion mining model.

5.3. Comparative experiment

5.3.1. Main baseline models

To verify the proposed model's effectiveness, comparison experiments were conducted with the following five traditional baseline sentiment classification models:

Majority [29]: A basic method that assigns the maximum sentiment polarity in the training set to each sample in the test set.

TextCNN [16]: The word embeddings are convolved separately using four convolutional kernels of different sizes. The feature maps obtained are then concatenated after maximum pooling and fed into the fully connected layer and the softmax function to obtain the sentiment classification results.

BiLSTM [9]: The hidden states of each word are obtained by modeling the context using a bi-directional LSTM network. Subsequently, the average of all hidden states is considered the final representation and is fed into the fully connected layer and softmax function to obtain the sentiment classification results.

Self-Attention [30]: The self-attention mechanism directly obtains the context vectors of word embeddings, which are averaged and fed into the fully connected layer and softmax function as the final representation to obtain the sentiment classification results.

BiLSTM+SelfAttention [11]: The hidden states are first obtained using the BiLSTM network and subsequently fed into the self-attention mechanism.

5.3.2. Analysis of experimental results based on official government documents dataset

Table 5 presents a comparison of the experimental results between the model proposed in this paper and the main baseline models.

Table 5. Experimental results.

Model	Acc	F1 value
Majority	0.5439	0.3523
TextCNN	0.8543	0.8711
Self-Attention	0.8869	0.8863
BiLSTM	0.8819	0.8920
BiLSTM+SelfAttention	0.8995	0.9083
Proposed Model	**0.9447**	**0.9484**

Notably, the worst results were observed for Majority, with all other baseline models exhibiting similar results, possibly explained by the fact that the Majority classification method only considers the proportion of different data categories in the training set, without considering any semantic information.

TextCNN, which uses four different sizes of convolutional kernels to convolve word embeddings, extracts the n-gram information of textual data (where n denotes the size of a convolutional kernel). However, it ignores the location information of text sequences and is relatively weak in capturing sequence information and long-range dependencies, thus achieving results of only 0.8543 and 0.8711.

The BiLSTM method differs from TextCNN in that it fully considers the location information of the input sequence, feeds each word into the model sequentially, and finally yields an output that considers a combination of past and future semantic information. Therefore, the Acc and F1 values obtained are superior to those of TextCNN because, in those cases, the cellular state is used to store long-distance information and the gating structure is used to control the information tradeoff, somewhat alleviating the problem of long-distance dependency.

The self-attention method can obtain the dependency between the current word and other position words through the attention mechanism, addressing the problem of long-distance dependency. However, similar to TextCNN, it ignores position information of the text sequence, performing better than TextCNN and slightly worse than BiLSTM on the dataset.

The BiLSTM+Self-Attention is a combination of the two aforementioned methods addressing the long-distance dependency problem while considering the position relationships between words, thereby yielding improved Acc and F1 results compared to those attained by its individual constituent models.

It is apparent from the experimental results that the proposed opinion mining model achieves the best results compared to all traditional models, mainly because none of the above baselines account for the target entity's influence on model performance. The model proposed in this paper obtains the final representation by embedding the target entity into itself, extracting the initial representation using BiLSTM, calculating the self-attention representation of the text and the target-directed attention representation separately, and fusing the two to obtain the final representation. Thus, the model accounts for dependencies between words and the target entity's attention to those words.

In particular, we conducted a set of ablation experiments to verify the effectiveness of the fused target entity. Comparing our model with the BiLSTM+SelfAttention model that removes the target entity, the results show that the fused target entity model improves the ACC and F1 score by 0.0452 and 0.0401, respectively. This indicates that, by fusing the semantic information of the target entity, the performance of the model is significantly improved and opinion information in sentences can be more accurately mined.

5.3.3. Analysis of experimental results based on public dataset

Table 6 lists the experimental results of the ACSA task performed on the ASAP dataset by the proposed model proposed. The baseline models used for comparison were TextCNN [16], BiLSTM+Attention [31], ATAE-LSTM [23], and CapsNet [32]; the experimental results of these models were all taken from the literature [27].

Table 6. Experimental results for the ACSA task.

Model	Acc	F1 value
TextCNN	0.7110	0.6041
BiLSTM+Attention	0.7778	0.7053
ATAE-LSTM	0.8194	0.7660
CapsNet	0.8166	0.7554
Proposed Model	**0.8152**	**0.7513**

The proposed model performed well on the ACSA task. Compared with TextCNN and BiLSTM+Attention, our model improved ACC by 0.1042 and 0.0374, respectively. Although our model is slightly deficient in ACC compared with ATAE-LSTM and CapsNet, the differences were only 0.0042 and 0.0014, which is sufficient to prove that our model also performs well on other domain datasets, further verifying the effectiveness and generalization performance of the model.

6. Conclusions and future prospects

6.1. Conclusions

This study designed an opinion mining model to mine the opinion categories of sentences for specific target entities in official government documents. The underlying concept of the model is to fuse the self-attention representation of the text with the attention representation directed by the target entity to obtain the final representation. The model is thereby able to pay close attention to important elements of the text itself while also considering the attention level of the target entity towards different words, enabling accurate extraction of semantic information embedded in the text and target entity. Experimental results on a dataset of official government documents indicate that the proposed model can effectively learn features and provide sufficient information for opinion classification. These results also show that the proposed model can reasonably focus on words more important in determining opinion categories. Thus, the performance of the model is further improved. The program code of the model can be found at: https://github.com/Yangt524/OpinionMining.

Although traditional machine learning methods have some drawbacks, they are still highly generalizable and interpretable. In future research, there are plans to incorporate traditional machine learning methods into the model to investigate whether these qualities can be further improved.

6.2. Future prospects

The proposed opinion mining model provides a new solution for opinion classification in official government documents. In the future, as the amount of text data in official government documents and other fields continues to increase, its application prospects will become increasingly broad. In practical applications, the model can be used in the fields of opinion consistency detection, assistant writing, and policy analysis of government documents, which greatly improves the work efficiency of public officials. In addition, research based on this model can also be extended to other fields, such as product reviews and social media text, to better meet people's needs for text data mining. Future research can also focus on optimizing the efficiency of the model and enhancing the robustness of the model to improve its performance in practical applications. Therefore, the proposed opinion mining model has broad application prospects and research value.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. W. Zhang, F. Wang, H. Zhao. J. Zhang, Topic extraction and analysis of S&T policies based on structural decomposition of government documents in Chinese, *Stud. Sci. Sci.*, **38** (2020), 1185–1196. https://doi.org/10.16192/j.cnki.1003-2053.2020.07.005
2. Y. Zhai, The theoretical evolution and practical exploration of China's e-government development in the past 40 years of reform and opening up: from business online to service online in Chinese, *E-Gov.*, **12** (2018), 80–89. https://doi.org/10.16582/j.cnki.dzzw.2018.12.008
3. Q. Qin, X. Zhang, On the institutional attribute of documents jointly issued by the party and government in Chinese, *J. Party School Cent. Comm. CPC. (Chin. Acad. Governance)*, **25** (2021), 120–128. https://doi.org/10.14119/j.cnki.zgxb.2021.04.014
4. Y. Yang, F. Wang, Whether deleveraging task conflicts with industrial policy: an empirical study based on enterprises' debt structure in Chinese, *J. Zhongnan Univ. Econ. Law*, **2** (2020), 3–13. https://doi.org/10.19639/j.cnki.issn1003-5230.20200427.003
5. S. Sun, C. Luo, J. Chen, A review of natural language processing techniques for opinion mining systems, *Inf. Fusion*, **36** (2017), 10–25. https://doi.org/10.1016/j.inffus.2016.10.004

6. J. Gong, X. Huang, Text categorization framework based on improved TF-IDF and k-nearest neighbor in Chinese, *Comput. Eng. Des.*, **39** (2018), 1340–1344+1349. https://doi.org/10.16208/j.issn1000-7024.2018.05.024
7. M. T. Zulfikar, Suharjito, Detection traffic congestion based on Twitter data using machine learning, *Procedia Comput. Sci.*, **2019** (2019), 118–124. https://doi.org/10.1016/j.procs.2019.08.148
8. B. Zhao, L. Wang, H. Guo, Weighted naive bayes text classification algorithm based on poisson distribution in Chinese, *Comput. Eng.*, **46** (2020), 91–96. https://doi.org/10.19678/j.issn.1000-3428.0054056
9. X. Huang, G. Liu, X. Liu, A. Yang, Sentiment classification depth model based on word2vec and bi-directional LSTM in Chinese, *Appl. Res. Comput.*, **36** (2019), 3583–3587+3596. https://doi.org/10.19734/j.issn.1001-3695.2018.08.0599
10. K. Cheng, Y. Yue, Z. Song, Sentiment classification based on part-of-speech and self-attention mechanism, *IEEE Access*, **2020** (2020), 16387–16396. https://doi.org/10.1109/ACCESS.2020.2967103
11. X. Wu, L. Chen, T. Wei, T. T. Fan, Sentiment analysis of chinese short text based on self-attention and Bi-LSTM in Chinese, *J. Chin. Inf. Process.*, **33** (2019), 100–107. https://doi.org/10.3969/j.issn.1003-0077.2019.06.015
12. T. Li, D. Ji, Sentiment analysis of micro-blog based on SVM and CRF using various combinations of features in Chinese, *Appl. Res. Comput.*, **32** (2015), 978–981. https://doi.org/10.3969/j.issn.1001-3695.2015.04.004
13. A. Goel, J. Gautam, S. Kumar, Real time sentiment analysis of tweets using naive bayes, in *2016 2nd International Conference on Next Generation Computing Technologies (NGCT)*, (2016), 257–261. https://doi.org/10.1109/NGCT.2016.7877424
14. J. Ababneh, Application of naive bayes, decision tree, and k-nearest neighbors for automated text classification, *Mod. Appl. Sci.*, **13** (2019), 31–36. https://doi.org/10.5539/mas.v13n11p31
15. R. Ahuja, S. C. Sharma, Sentiment analysis on different domains using machine learning algorithms, *Adv. Data Inf. Sci.*, **2022** (2022), 143–153. https://doi.org/10.1007/978-981-16-5689-7_13
16. Y. Kim, Convolutional neural networks for sentence classification, *arXiv preprint*, (2014), Arxiv:1408.5882. https://doi.org/10.48550/arXiv.1408.5882
17. R. Johnson, Z. Tong, Deep pyramid convolutional neural networks for text categorization, in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, (2017), 562–570. https://doi.org/10.18653/v1/p17-1052
18. S. Yu, D. Liu, Y. Zhang, S. Zhao, W. Wang, DPTCN: A novel deep CNN model for short text classification, *J. Intell. Fuzzy Syst.*, **6** (2021), 7093–7100. https://doi.org/10.3233/JIFS-210970
19. Y. Wang, A. Sun, J. Han, Y. Liu, X. Zhu, Sentiment analysis by capsules, in *the Web Conference*, (2018), 1165–1174. https://doi.org/10.1145/3178876.3186015
20. R. Wang, Z. Li, J. Cao, C. Tong, W. Lei, Convolutional recurrent neural networks for text classification, in *2019 International Joint Conference on Neural Networks*, (2019), 1–6. https://doi.org/10.1109/IJCNN.2019.8852406
21. A. Singh, S. K. Dargar, A. Gupta, A. Kumar, A. K. Srivastava, M. Srivastava, et al., Evolving long short-term memory network-based text classification, *Comput. Intell. Neurosci.*, **2022** (2022), 1687–5265. https://doi.org/10.1155/2022/4725639

22. D. Bahdanau, K. Cho, Y. Bengio, Neural machine translation by jointly learning to align and translate, *arXiv preprint*, (2014), arXiv:1409.0473. https://doi.org/10.48550/arXiv.1409.0473
23. Y. Wang, M. Huang, X. Zhu, L. Zhao, Attention-based LSTM for aspect-level sentiment classification, in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, (2016), 606–615. https://doi.org/0.18653/v1/D16-1058
24. Y. Liu, P. Li, X. Hu, Combining context-relevant features with multi-stage attention network for short text classification, *Comput. Speech Lang.*, **71** (2021), 1–14. https://doi.org/10.1016/j.csl.2021.101268
25. C. Hu, N. Liang, Deeper attention-based LSTM for aspect sentiment analysis in Chinese, *Appl. Res. Comput.*, **36** (2019),1075–1079, https://doi.org/10.19734/j.issn.1001-3695.2017.11.0736
26. T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, *arXiv preprint*, (2013), arXiv:1301.3781. https://doi.org/10.48550/arXiv.1301.3781
27. J. Bu, L. Ren, S. Zheng, Y. Yang, J. Wang, F. Zhang, et al., ASAP: A chinese review dataset towards aspect category sentiment analysis and rating prediction, *arXiv preprint*, (2021), arXiv:2103.06605. https://doi.org/10.48550/arXiv.2103.06605
28. F. Song, L. Gao, Performance evaluation metric for text classifiers in Chinese, *Comput. Eng*., **30** (2004), 107–127. https://doi.org/10.3969/j.issn.1000-3428.2004.13.044
29. D. Ma, S. Li, X. Zhang, H. Wang, Interactive attention networks for aspect-level sentiment classification, in *Twenty-Sixth International Joint Conference on Artificial Intelligence*, (2017), 1–7. https://doi.org/10.24963/ijcai.2017/568
30. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. Gomez, et al., Attention is all you need, *arXiv preprint*, (2017), arXiv:1706.03762. https://doi.org/10.48550/arXiv.1706.03762
31. P. Zhou, W. Shi, J. Tian, Z. Qi, B. Li, H. Hao, et al., Attention-based bidirectional long short-term memory networks for relation classification, in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, (2016), 207–212. https://doi.org/10.18653/v1/P16-2034
32. S. Sabour, N. Frosst, G. E. Hinton, Dynamic routing between capsules, in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, (2017), 3859–3869.