ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20221018031450
Preliminary communication

Logistics Service Quality Sentiment Analysis with Deeper Attention LSTM Model with Aspect Embedding

Wenjing XUAN, Min DENG*

Abstract: To understand the audience's subjective perception of quality of service (QoS), it is important to analyze the data acquired from the logistics service logs and online evaluation system reasonably and effectively. Based on the analysis, rational improvement measures and decision suggestions can be developed to enhance the QoS. However, modern logistics service departments often face various business needs and service objects at the same time. If the evaluation subjects and their relationships are unclear in the service evaluation data, the sentiment analysis result of the text is a coarse-grained evaluation of the service as a whole. The lack of fine-grained pertinent evaluation results will hinder the improvement of specific management measures. To solve the problem, this paper designs an attention-based long short-term memory network (AT-LSTM) to divide the service reviews into ten topic relations, and then builds a deeper attention LSTM with aspect embedding (AE-DATT-LSTM). The weight-sharing bidirectional LSTM (BiLSTM) trains the topic word vectors and the text word vectors, and fuses the resulting topic features and text features. After the processing of the deep attention mechanism, the sentiment class of each evaluation topic is obtained by the classifier. Finally, several experiments were carried out on different public datasets. The results show that our approach surpasses the previous attention-based sentiment analysis models in accuracy and stability of service quality sentiment analysis. The introduction of topic features and deep attention mechanism is of great significance for the QoS-based sentiment classification b, and provides a feasible method for other fields like public opinion analysis, question answering system, and text reasoning.

Keywords: deep attention; deep learning; logistics management; long short-term memory (LSTM); service quality sentiment analysis; semantic relation

1 INTRODUCTION

With the development of the Internet industry, the accumulation of e-commerce reviews, and the promotion of social platforms, a large amount of short text data has been accumulated. Better decisions and promotions are possible, if the opinions and emotions can be extracted from these data. Text sentiment analysis through natural language processing (NLP) has become a hot topic at home and abroad [1-3]. Thanks to the in-depth development of informatization, network-based service provision can cut down the logistics management cost, and provide a convenient platform for the service objects to express their personal opinions [4]. The emotional tendencies embed in the personal opinions help to understand the problems arising in the logistics services, and to reasonably evaluate the services accepted by users. Different users tend to hold different views and emotional tendencies towards the process and results of the same service. Understanding this phenomenon provides strong support for formulating reasonable management measures to improve the quality of service (QoS) and user experience. However, modern logistics service departments often face various business needs and service objects at the same time.

In addition to face-to-face communication with customers, comments left by customers are important raw data for analyzing service quality. However, due to the large differences in presentation and the large amount of review data, it is extremely expensive to analyze articles by manual evaluation. However, using NLP analysis faces a series of challenges. When the service evaluation data lacks clear evaluation subjects and relationships, the results of the emotional analysis of the evaluation text are coarse-grained evaluations of the whole service, but lack precise analysis of fine granularity and directivity. This kind of analysis results can only understand the "good" and "bad" from the macro level, but cannot find specific problems in the service, which is unfavourable to the improvement of specific management measures. Therefore, we need to establish a more granular comment analysis method to find problems.

Sentiment analysis, also known as opinion mining, is a fundamental task of NLP and computational linguistics. In recent years, it has received a growing attention from the industry and academia. The key to sentiment analysis is to understand the textual information generated by users in social media and product reviews, and to mine sentiments from the information [5]. Recently, deep learning-based methods have performed excellently in many NLP tasks, such as machine translation, semantic recognition, question answering, and text summarization. Meanwhile, deep learning models improve the classification accuracy in sentiment analysis, without the feature engineering of manual labelling. Socher et al. [6] applied the recurrent neural network (RNN) to the construction of emotion tree, and enhanced the detection accuracy of emotions. Tang et al. [7] developed a text-level RNN, which is superior to the standard RNN in sentiment classification. Tai et al. [8] improved the standard long short-term memory (LSTM) into the tree-LSTM model. The tree-like LSTM topology of the model performed well in sentiment classification.

The attention mechanism comes from the fact that the human brain pays more attention to the key parts of things. By calculating the probability distribution of attention, the key input is highlighted to optimize the traditional model. Attention mechanism was proposed by Lin et al. [9], who introduced the attention mechanism into the RNN to elevate the accuracy of image classification. Since then, the attention mechanism has achieved good results in various aspects of NLP, such as machine translation [10] and text summarization [11]. Yang et al. [12] proposed the hierarchical attention network for text classification, which achieved a better classification effect than the previous models. Pavlopoulos et al. [13] developed a deep attention mechanism, and applied it to review user comments. The deep attention mechanism achieves better results than RNN. In NLP, the attention mechanism guides models to focus on the important parts of sentences. In sentiment analysis tasks, many models have taken advantage of the fact that attention can capture important information in sentences, and thus achieve ideal results.

2 RELATED WORKS

2.1 Topic Sentiment Analysis

Topic sentiment classification, a fine-grained task of sentiment classification, aims to infer the sentiment polarity corresponding to a given sentence and the topics of that sentence [14, 15], such as positive, negative and neutral. Taking "The food is great and they have a good selection of wines at reasonable prices" for example, the sentiment polarity of the topic "food" is positive, while the sentiment polarity of the topic "price" is negative. In this paper, the topic information comes from the extracted training set and test set. In practical applications, the topic information can be extracted automatically from keywords through term frequency-inverse document frequency (TF-IDF) or latent Dirichlet allocation (LDA). Most of the existing methods attempt to mine the sentiment polarity of whole sentences, without considering the situation of a given topic. Traditionally, this problem is solved with a series of manually labelled features. With the continuous expansion of sentiment dictionaries, dictionary-based features are increasingly adopted in sentiment analysis tasks. Most scholars focus on building sentiment classifiers using features, including bag-of-words (BOW) model and sentiment dictionary. The representative methods are grounded on the support vector machine (SVM) [16] or the neural network (NN) [17]. Without paying attention to feature engineering, NN-based methods can learn the feature representation in the data excellently, and obtain the exact semantic relationship and sentiment information between the context and topic information.

The NN-based methods can effectively learn the feature representation, and thus improve the sentiment classification of a given topic. Tang et al. [18] presented the target-dependent LSTM (TD-LSTM) and target connection LSTM (TC-LSTM). Considering the topic information, the two methods increase the accuracy of the sentiment classification for a given topic. TC-LSTM extracts topic word vectors by averaging word vectors contained in topic information. Wang et al. [19] proposed an attention mechanism for the sentiment analysis of a given topic. Facing different attention inputs, the mechanism focuses on different parts of the sentence, and excels in the sentiment classification of a given topic. Chen et al. [20] put forward a sentiment analysis approach for a given topic based on the recurrent attention memory network. Specifically, the RNN was combined nonlinearly with the attention mechanisms at multiple positions, such that the model could process more complex sentences. Tang et al. [21] created a sentiment analysis strategy for a given topic based on deep memory network, which integrates content attention with location attention. Their strategy can capture the important context words in the sentiment classification of a given topic, and realize outstanding processing speed and accuracy. Liang et al. [22] set up a multi-attention convolutional neural network (CNN) for the sentiment analysis of a given target. The multi-attention CNN fuses three types of attention mechanisms, and outperforms the ordinary CNN, single-attention CNN, and attention-based LSTM in sentiment classification. In addition, the attention-based deep LSTM model proposed by Baziotis et al. [23] achieved good results in the SemEval-2017 sentiment classification task. The above research works show that deep learning models that fuse given topic and attention have achieved great successes in given topic sentiment analysis. However, the existing models do not fully integrate the topic information after feature extraction, and, for the application of attention, fail to apply the multi-layer attention mechanism to sentiment analysis.

The graph neural network (GNN) [24, 25] attracts much attention in text analysis. In the GNN, each node collects information from other nodes and updates its form. This network has been applied to many NLP tasks, including text classification [26, 27], neural machine translation, and relationship reasoning [28]. Based on the graph convolutional network (GCN) [29], Yao et al. [26] proposed a text classification method called Text-GCN, which constructs a single text diagram for the entire corpus in the light of the word co-occurrence and document-word relationship. The disadvantages of the method include high memory occupation, incompatibility with online testing, and the neglection of word sequence. To address these problems, Huang et al. [28] presented a text-level GNN for text classification (H-GNN). Rather than build a map for the entire corpus, H-GNN develops a diagram for each input text. Nevertheless, H-GNN only considers the adjacent nodes in a small window, rather than the relationship between non-adjacent nodes, during the update of graph nodes. Thus, the model only focuses on local features, and has difficulty in learning the dependencies between remote nodes. As a result, H-GNN only emphasizes local features as the CNN.

The deep learning model that fuses the topic information improves the role of topic information in the text by taking the topic information as the sole input. But it cannot extract the key part of the text information that affects the topic. The attention-based deep learning model can obtain the key information from the text, but ignores the role of topic information. The deep learning model, which fuses the topic information and adopts a single layer of attention, e.g., the attention-based LSTM with aspect embedding (ATAE-LSTM) proposed by Wang et al. [19], connects the topic information with text information as input, and uses a single layer of attention. But it fails to fully utilize the topic information or the critical information in the text. This paper establishes a deep attention LSTM model that fuses the topic features. Both text information and topic information were taken as inputs. The weight-sharing bidirectional LSTM (BiLSTM) trains the topic word vectors and the text word vectors, and fuses the resulting topic features and text features. The fused features were processed by the deep attention mechanism, before effectively detecting the sentiment of different topics in the text.

2.2 LSTM

The RNN, as an improved version of the feedforward neural network, still faces exploding and vanishing gradients. To solve the problems, Hochreiter and Schmidhuber [30] proposed and realized the LSTM. The basic structure of the LSTM is shown in Fig. 1.

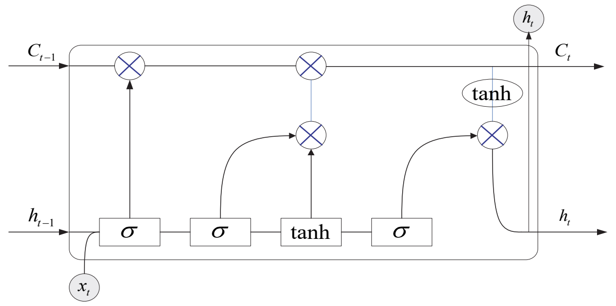

Figure 1 LSTM structure

Each LSTM model contains an input gate i_t, a forget gate f_t, an output gate o_t, and a memory cell c_t. The input series of word vectors can be denoted as $\{x_1, x_2, \ldots, x_n\}$, where x_t is an input to the LSTM unit, representing a word vector in the input series. The vector of the hidden layer can be denoted as h_t. Then, the three gates and the memory cell in each LSTM unit can be calculated by:

$$X = \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} \tag{1}$$

$$f_t = \sigma\left(W_f \cdot X + b_f\right) \tag{2}$$

$$i_t = \sigma\left(W_i \cdot X + b_i\right) \tag{3}$$

$$o_t = \sigma\left(W_o \cdot X + b_o\right) \tag{4}$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tan h\left(W_c \cdot X + b_c\right) \tag{5}$$

$$h_t = o_t \cdot \tan h\left(c_t\right) \tag{6}$$

where, W_i, W_f and $W_o \in R^{d\times 2d}$ are weight matrices; b_f, b_i and b_o are the biases learned in LSTM training; tan*h* is the activation function; *X* is the pointwise multiplication.

2.3 Attention Mechanism

The attention mechanism guides the model to focus on the important information in the text. The input of the attention mechanism can be constructed as $H \in R^{d\times n}$, according to the hidden layer features $H = \{h_1, h_2, \ldots, h_N\}$ of the LSTM. Note that d is the length of the hidden layer; *N* is the length of the input sentence. Fig. 2 shows the basic structure of the attention mechanism.

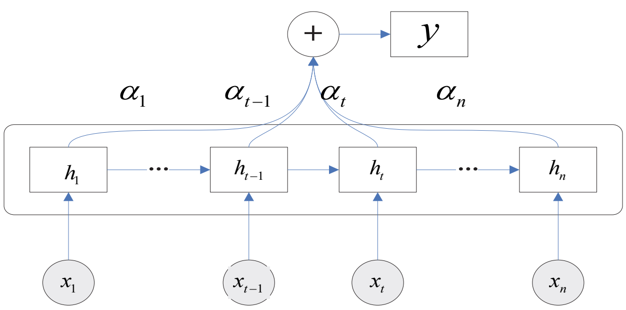

Figure 2 Basic structure of the attention mechanism

The attention mechanism generates an attention weight matrix α and a feature representation *v*:

$$u_t = \tan h\left(W_s h_i + b_s\right) \tag{7}$$

$$\alpha_i = \frac{\exp\left(u_i\right)}{\sum_i \exp\left(u_i\right)}, \sum_i \alpha_i = 1 \tag{8}$$

$$v = \sum_i \alpha_i h_i \tag{9}$$

3 AE-DATT-LSTM

To highlight more valuable information, this paper puts forward a deeper attention LSTM model with aspect embedding (AE-DATT-LSTM). Composed of an improved attention model and a two-layer BiLSTM, the proposed model can extract the topic information and sentiment information more effectively. As shown in Fig. 3, the AE-DATT-LSTM consists of five layers: a word vector layer, a BiLSTM layer, a pooling layer, a feature fusion layer, and an attention layer.

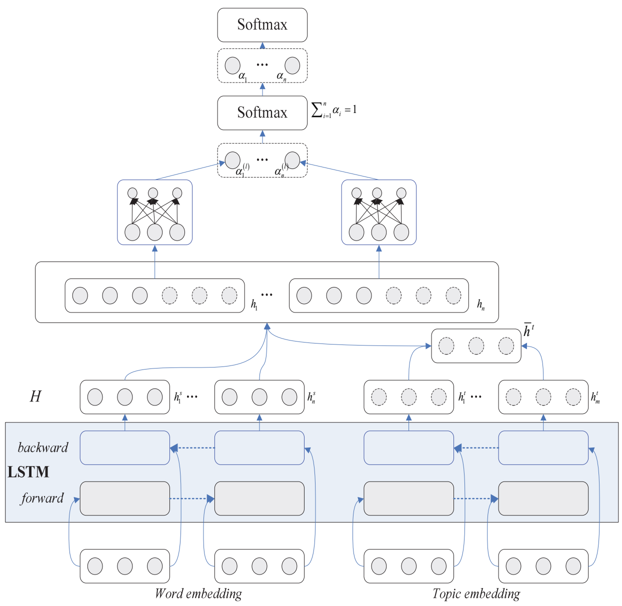

Figure 3 Structure of the AE-DATT-LSTM

3.1 Task Definition

For a sentence or text*s* = $\{\omega_1, \omega_2, ..., \omega_n\}$ composed of n words, ω_i is a topic word appearing in that text. The QoS sentiment analysis aims to determine the sentiment polarity of the topic word ω_i in text s. For example, the input text goes: "If you want a casual neighborhood bistro that has great food and excellent service, this is the place. " The topic "food" has a positive sentiment polarity, while the topic "service" has a negative sentiment polarity. During the processing of a text set, each word is mapped into a low-dimensional, continuous, and real value vector called the word vector. All word vectors constitute a word vector matrix $L \in R^{d\times|V|}$, where d is the dimension of the word vector; $|V|$ is the size of the dictionary. Let $W_o \in R^{d\times 2d}$ and $W_i^s \in \mathrm{R}^{d\times 1}$ be the vectors of a word ω_i and a topic word ω_j

in the text, respectively. The word ω_i and the topic word ω_j form a column of the word vector matrix $\boldsymbol{L}$.

3.2 Word Vector Input Layer

The AE-DATT-LSTM has two inputs. The input series of the words in the text are converted into a series of word vectors $w^s = \{w_1^s, w_2^s, \cdots, w_n^s\}$, where n is the number of words in the text. Meanwhile, the topic words of the text are converted into a series of topic word vectors $w^t = \{w_1^t, w_2^t, \cdots, w_m^t\}$, where m is the number of topic words in the text.

3.3 BiLSTM

To make meaningful comparison between the text word vectors and topic word vectors, this paper uses the weight-sharing BiLSTM to map the text and topics to the same vector space. On the text level, the BiLSTM generates a feature series $H^s = \{h_1^s, h_2^s, \cdots, h_n^s\}$. On the topic level, the network produces another feature series $H^t = \{h_1^t, h_2^t, \cdots, h_m^t\}$. Each eigenvector is generated from forward and reverse LSTM connections: $h_i^j = \overrightarrow{h_i^j} \parallel \overleftarrow{h_i^j}$, $h_i^j \in \mathrm{R}^{2L}$, $j \in \{s, t\}$, where || is the connection operation; L is the size of each LSTM.

3.4 Pooling Layer

This paper uses a pooling layer to aggregate the topic features into a single feature. The pooling layer adopts mean over time to generate topic features $h^s = \frac{1}{m}\sum_1^m h_i^t$.

3.5 Feature Fusion Layer

The text features produced by the LSTM are connected to the topic features aggregated by the pooling layer. In this way, a fused feature is generated for each word $h_i^{'} = h_i^s \parallel h^t$, $h_i^s \in \mathrm{R}^{4L}$.

3.6 Attention Layer

During the sentiment analysis on QoS in the text, more attention needs to be paid to the topic information and emotional words in the sentence. Therefore, this paper introduces the deep attention mechanism of multilayer perceptron (MLP) to extract the salient topic words and emotional words from the sentence. To enhance the contribution rate of emotional words to each topic in the sentence, a deep perceptual attention mechanism is employed for modeling based on the MLP. Through the deep semantic attention of a layer of MLP, the attention weight α_i can be obtained:

$$\alpha_i^{(1)} = \mathrm{ReLU}\left(W^{(1)} h_i^{'} + b^{(1)}\right) \tag{10}$$

$$\alpha_i^{(l-1)} = \mathrm{ReLU}\left(W^{(l-1)} \alpha_i^{(l-2)} + b^{(l-1)}\right) \tag{11}$$

$$\alpha_i^{(l)} = W^{(l)} \alpha_i^{(l-1)} + b^{(l)} \tag{12}$$

$$\alpha_i = soft\max\left(\alpha_1^{(l)}, \alpha_2^{(l)}, \cdots, \alpha_n^{(l)}\right), \sum_{i=1}^{i} a_i = 1 \tag{13}$$

$$h^* = \sum_{i=1}^{i} a_i h_i^{'}, r \in \mathrm{R}^{4L} \tag{14}$$

where, $W^{(1)}$, $W^{(2)}$, …, $W^{(l)}$, and $b^{(1)}$, $b^{(2)}$, …, $b^{(l)}$ are determined through training.

Taking h^* as the features of sentiment classification, the text is sentimentally classified by the softmax classifier:

$$p = \mathrm{softmax}\left(W_h h^* + b_h\right) \tag{15}$$

3.7 Model Training

Our model is trained through end-to-end backpropagation, with the cross-entropy cost function as the objective function. Let γ be the known sentiment class of each topic in the text; $\hat{\gamma}$ be the predicted sentiment class of each topic in the text. The goal of model training is to minimize the cross-entropy between the known and predicted sentiment classes in all the sentences:

$$\mathrm{loss} = -\sum_i \sum_j \gamma_i^j \log \hat{\gamma}_i^j + \lambda \|\theta\| \tag{16}$$

where, i is the serial number of sentences; j is the serial number of sentiment classes (there are a total of three sentiments: positive, neutral, and negative); λ is L_2 regularization, a penalty term of the cost function; θ is the parameter to be configured.

Table 1 Sample relations and topic classes

Serial number of topics	Type of relation	Number of samples	Serial number of topics	Type of relation	Number of samples
1	Cause-Effect	1331	6	Member-Collection	923
2	Component-Whole	1253	7	Message-Topic	895
3	Entity-Destination	1137	8	Content-Container	732
4	Entity-Origin	974	9	Instrument-Agency	660
5	Product-Producer	948	10	Other	1864

4 EXPERIMENTS AND RESULTS ANALYSIS

The proposed AE-DATT-LSTM was applied to classify the sentiments over the QoS. During our experiments, the pretrained 300 dimensional Glove word vectors proposed by Pennington et al. [31] were initialized to obtain a 1.9 MB dictionary. Any word not contained in the dictionary was randomly initialized under the uniform distribution $U(-\varepsilon, \varepsilon)$, with $\varepsilon = 0.01$. For a given text "The restaurant was too expensive", the topic word "price" has a negative sentiment polarity.

4.1 Datasets

The acquisition of relations was performed on the public corpus SemEval-2010 task 8 [23], which contains 10717 samples with 10 kinds of relations. Among them, 8000 were adopted for training. Tab. 1 shows the sample relations and topic classes.

The QoS evaluation was carried out on the following datasets: SemEval-2014 task4 [14], SemEval-2017 task4 [15], and Comment Dataset Library Service Quality (CDLSQ). Specifically, SemEval provides the datasets for semantic evaluation competitions. There are two subsets in SemEval-2014 task4: laptop and restaurant. Only the restaurant subset involves topic-based sentiment polarities. The consumers' review texts in the subset cover three sentiments (positive, neutral, and negative), and five topics {food, price, service, ambience, and anecdotes /miscellaneous}.

The CDLSQ is a QoS evaluation sample set built by our team. The 8176 comments in the dataset cover five sentiment indices unacceptable, Unpleasant, fair (neutral), agreeable, and favorable. For consistency, the five types of sentiment indices were reorganized into three emotional tendencies (negative, neutral, and positive).

SemEval-2017 task4 contains the reviews on Twitter. The topics of each review are extracted from the corresponding text. The topics in the dataset involve two sentiment polarities: positive and negative. Each review covers a series of topics and their sentiment polarities.

This paper attempts to differentiate the sentiment polarities (positive and negative) of different topics in the text. The experimental data are summarized in Tab. 2.

Table 2 List of datasets

Data source	Dataset	Negative	Neutral	Positive
restaurant	train	2179	500	839
	test	657	94	222
Twitter	train	14897	-	3997
	test	2463	-	3722
CDLSQ	train	1061	2299	2771
	test	354	767	924

4.2 Parameter Setting

The dimension of word vectors, topic words, and the synthetized output features of the LSTM was set to 300. The attention weight dimension was set as equal to the length of the text. The model was trained on batches of 64 samples. The Adam learning rate, penalty term of the cost function, and the number of deep attention layers were configured as 0.001, 0.001, and 4, respectively. In addition, the Dropout was set to 0.5, the number of LSTM layers to 64, and the number of BiLSTM layers to 128.

4.3 Comparative Analysis

The proposed AE-DATT-LSTM was compared with each of the following six approaches on two different datasets:

(a) LSTM:

The standard LSTM model [29] cannot infer any topic information in the text. It always produces the same sentiment polarity, despite being given different topics.

(b) TD-LSTM:

TD-LSTM, proposed by Tang et al. [18], extracts information before and after each topic word by forward and backward LSTMs, respectively. In the last step, the TD-LSTM adopts the LSTM output as the predicted class for the content. The model improves the sentiment classification by treating each topic feature as a target. Nevertheless, the lack of attention mechanism makes it impossible to obtain the keywords of a given topic from the text.

(c)TC-LSTM:

TC-LSTM [18] is extended from the TD-LSTM by introducing the topic vector to the feature representation of the sentence. With the new topic connection module, the TC-LSTM can better utilize the topic word and every other word in the text, and link them up into a feature representation of the text.

(d)AT-LSTM:

The standard LSTM cannot detect the important information in the text to assist the topic-based sentiment classification. To solve the issue, Wang et al. [19] introduced the attention mechanism into the LSTM, producing the AT-LSTM. The new model can pinpoint the key parts of the text for a given topic.

(e) ATAE-LSTM:

The AE-LSTM utilizes the topic information, and gives full play to topic word vectors in the attention weight mechanism. To better utilize the information, Chen et al. [20] improve the AE-LSTM into the ATAE-LSTM, which connects the topic word vectors to the input vector of each word.

(f) AE-ATT-BLSTM:

As a single-layer attention LSTM with topic fusion, the AE-ATT-BLSTM mainly adds an attention module with a single-layer MLP after the feature fusion layer.

Table 3 Sentiment classification accuracy of different methods on the three datasets (%)

Model	Restaurant		Twitter		CDLSQ	
	neu/neg/pos	neu/pos	neu/neg/pos	neu/pos	neu/neg/pos	neu/pos
LSTM	79.1	82.5	-	76.1	72.6	75.9
TD-LSTM	77.6	83	-	80.9	68.7	73.7
TC-LSTM	77.2	84.3	-	83.5	69.4	72.5
AT-LSTM	81.5	84.5	-	82.7	68.5	78.4
ATAR-LSTM	80.6	85.6	-	85.7	69.1	75.5
AE-ATT-LSTM	84	87.4	-	85.08	79.7	78.6
AE-DATT-LSTM	85.5	88.7	-	86.24	80.9	79.9

4.4 Results Analysis

To confirm whether our approach can effectively classify the sentiments on different datasets, the training samples of the three datasets in Tab. 2 were adopted to train and cross-validate our model, and the test samples were used to test the model performance. The results are shown in Tab. 3.

Tab. 3 reports the classification accuracy of the six models concerning the three sentiment polarities (positive, negative, and neutral) or two sentiment polarities (negative and positive) on three datasets, namely, restaurant, Twitter, and CDLSQ. Note that the Twitter dataset does not define the neutral class. That is why the classification accuracy column of that class on that dataset is blank in Tab. 3.

Overall, our AE-DATT-LSTM achieved better classification accuracy than any other model. The models with topic features outshined those without topic features. The models with the attention mechanism outperformed those without the mechanism. Therefore, the introduction of topic features and the attention mechanism greatly enhance the classification of QoS. The LSTM with deep attention and fused topic features does better than the models that only include the attention mechanism or only adopt the topic features.

The training process of our AE-DATT-LSTM is illustrated in Fig. 4. The training and test results of our AE-DATT-LSTM from epoch = 0 to epoch = 50 were analyzed, against those of the contrastive models. The trends of loss, accuracy, precision, and recall were plotted.

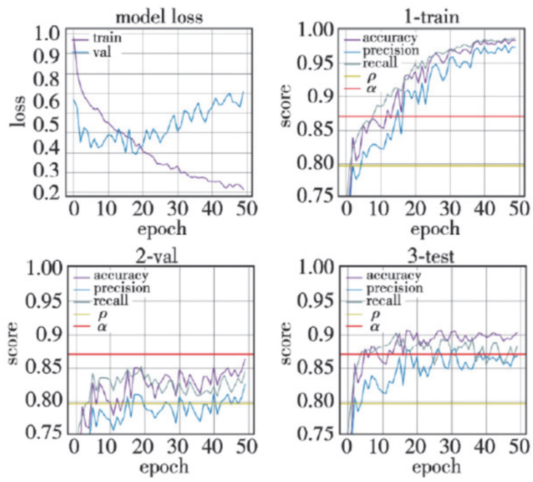

Figure 4 Training process of our AE-DATT-LSTM

During the testing process, the classification accuracy peaked at epoch = 18, when the F-score was 87.64.

Fig. 5 compares the accuracies of the single-layer attention LSTM with fused features, and the deep-layer attention LSTM with fused features. During the training from epoch = 0 to epoch = 50, the latter LSTM achieved better results than the former. Fig. 6 compares the F1 scores of the two models above. It can be seen that the latter generally outperformed the former.

Concerning the topic relation determination and classification, the topic relations in SemEval-2010 task 8 and CDLSQ were tested separately. The results are shown in Tab. 4.

Combining the results of topic analysis and QoS sentiment analysis, the evaluation results of QoS for different topics were summarized (Tab. 5).

As shown in Tab. 5, on the CDLSQ dataset, the sentiment distribution varied greatly from topic to topic. Overall, the sentiment evaluations were very positive. However, negative evaluations were obtained on (No. 2) Component-Whole, (No. 5) Product-Producer, and (No. 7) Message-Topic. In particular, the positive evaluations only took up 5% in Message-Topic. The services on these topics should be improved.

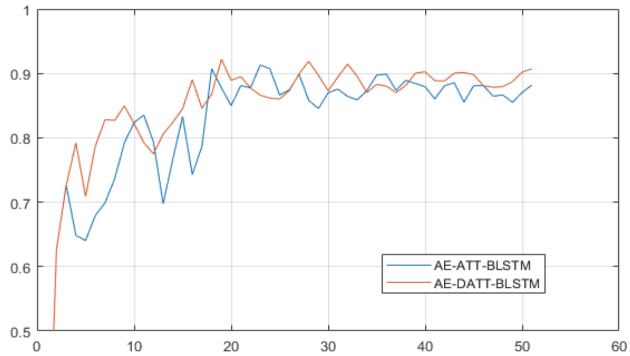

Figure 5 Accuracies of the single-layer attention LSTM with fused features, and the deep-layer attention LSTM with fused features

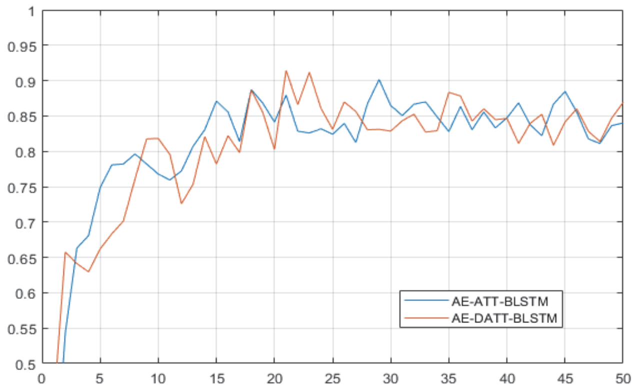

Figure 6 F1-scores of the single-layer attention LSTM with fused features, and the deep-layer attention LSTM with fused features

Table 4 Test results on topic relations (%)

Serial number of topics	SemEval-2010 task 8			CDLSQ		
	Precision	Recall	F	Precision	Recall	F
1	93.6	96.6	95.0	86.5	85.6	86.1
2	91.9	92.6	92.3	88.8	90.5	89.7
3	90.1	92.5	91.3	86.9	87.0	87.0
4	89.4	89.6	89.6	86.8	80.0	83.3
5	89.1	87.0	98.3	86.0	82.3	84.1
6	86.2	85.4	86.6	76.5	85.2	80.6
7	83.1	88.7	84.3	75.6	83.6	79.4
8	79.6	83.3	83.9	74.4	73.1	73.7
9	82.7	81.9	83.0	80.6	80.7	80.6
10	79.3	88.4	80.6	70.1	78.2	73.9

Table 5 QoS evaluation results of different topics (%)

Serial number of topics	Negative	Neutral	Positive	Topic No.	Negative	Neutral	Positive
1	17	41	42	6	11	18	71
2	21	48	31	7	25	70	5
3	13	18	69	8	10	27	63
4	19	28	53	9	15	41	44
5	23	42	35	10	19	42	39

5 CONCLUSIONS

Most sentiment classification approaches for QoS do not effectively combine topic information with sentiment information. This paper proposes a deeper attention LSTM model with fused topic features. The topic word vectors and text word vectors were trained by the weight-sharing BiLSTM, thereby completing feature fusion. After the processing of the deep attention mechanism, the information of different topics in the text can be obtained by the proposed method. Our AE-DATT-LSTM was compared with six other approaches on different datasets. The results show that our approach boasts better accuracy and stability than the other methods, providing a better tool for QoS sentiment analysis.

The main limitation of this paper is the failure to consider the part of speech (POS). The next work will try to introduce the POS to the QoS sentiment analysis, and make a deeper analysis on different entities (e.g., service personnel, service facilities, and tools) in negative evaluations.

Acknowledgment

This project was funded by the Humanities and Social Science Research Project of Ministry of Education of China (Grant No.: 20YJCZH081), and Scientific Research Project of Education Department of Hubei Province (Grant No.: D20212701, D20202701, B2020385).

6 REFERENCES

[1] Wang, Z., Cheng, J., Wang, H., & Wen, J. (2016). Short text understanding: A survey. *Journal of Computer Research and Development*, *53*(2), 262-269.

[2] Zhang, C., Li, Q. X., & Cheng, X. (2020). Text sentiment classification based on feature fusion. *Revue d'Intelligence Artificielle*, *34*(4), 515-520. https://doi.org/10.18280/ria.340418

[3] Amaria, S., Guidedi, K., & Lazarre, W. (2022). A Survey on Multimedia Ontologies for a Semantic Annotation of Cinematographic Resources for the Web of Data. *Acadlore Transactions on AI and Machine Learning*, *1*(1), 2-10. https://doi.org/10.56578/ataiml010102

[4] Wang, Z., Ho, S. B., & Cambria, E. (2020). A review of emotion sensing: categorization models and algorithms. *Multimedia Tools and Applications*, 79(47), 35553-35582. https://doi.org/10.1007/s11042-019-08328-z

[5] Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, *5*(1), 1-167. https://doi.org/10.2200/S00416ED1V01Y201204HLT016

[6] Socher, R., Perelygin, A., Wu, J., et al. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, 1631-1642.

[7] Tang, D., Qin, B., & Liu, T. (2015). Document modeling with gated recurrent neural network for sentiment classification. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 1422-1432. https://doi.org/10.18653/v1/D15-1167

[8] Tai, K. S., Socher, R., & Manning, C. D. (2015). Improved semantic representations from tree-structured long short-term memory networks. *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguisticsand the 7th International Joint Conference on Natural Language Processing,* Beijing, 1556-1566. https://doi.org/10.3115/v1/P15-1150

[9] Lin, L., Luo, H., Huang, R., & Ye, M. (2019). Recurrent models of visual co-attention for person re-identification. *IEEE Access*, 7, 8865-8875. https://doi.org/10.1109/ACCESS.2018.2890394

[10] Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint arXiv:1409.0473*.

[11] Rush, A. M., Chopra, S., & Weston, J. (2015). A neural attention model for abstractive sentence summarization. *arXiv preprint arXiv:1509.00685*. https://doi.org/10.18653/v1/D15-1044

[12] Zichao, Y., Diyi, Y., Chris, D., Xiaodong, H., Alex, S., & Eduard, H. (2016). Hierarchical attention networks for document classification. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 1480-1489. https://doi.org/10.18653/v1/N16-1174

[13] Pavlopoulos, J., Malakasiotis, P., & Androutsopoulos, I. (2017). Deeper attention to abusive user content moderation. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 1125-1135. https://doi.org/10.18653/v1/D17-1117

[14] Maria, P., Dimitris, G., John, P., Harris, P., Ion, A., Suresh, M. (2014). SemEval-2014 task 4: Aspect based sentiment analysis. *Proceedings of International Workshop on Semantic Evaluation (SemEval 2014)*, 27-35. https://doi.org/10.3115/v1/S14-2004

[15] Rosenthal, S., Farra, N., & Nakov, P. (2017). SemEval-2017 task 4: Sentiment analysis in Twitter. *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, 502-518. https://doi.org/10.18653/v1/S17-2088

[16] Kiritchenko, S., Zhu, X., Cherry, C., & Mohammad, S. (2014). Nrc-canada-2014: Detecting aspects and sentiment in customer reviews. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, 437-442. https://doi.org/10.3115/v1/S14-2076

[17] Nguyen, T. H. & Shirai, K. (2015). Phrasernn: Phrase recursive neural network for aspect-based sentiment analysis. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 2509-2514. https://doi.org/10.18653/v1/D15-1298

[18] Tang, D., Qin, B., Feng, X., & Liu, T. (2015). Effective LSTMs for target-dependent sentiment classification. *arXiv preprint arXiv:1512.01100*.

[19] Wang, Y., Huang, M., Zhu, X., &Zhao, L. (2016). Attention-based LSTM for aspect-level sentiment classification. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 606-615. https://doi.org/10.18653/v1/D16-1058

[20] Chen, P., Sun, Z., Bing, L., & Yang, W. (2017). Recurrent attention network on memory for aspect sentiment analysis. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 452-461.

https://doi.org/10.18653/v1/D17-1047
[21] Tang, D., Qin, B., & Liu, T. (2016). Aspect level sentiment classification with deep memory network. *arXiv preprint arXiv:1605.08900*. https://doi.org/10.18653/v1/D16-1021
[22] Liang, B., Liu, Q., Xu, J., et al. (2017). Aspect-based sentiment analysis based on multi-attention CNN. *Journal of Computer Research and Development*, *54*(8), 1724-1735.
[23] Baziotis, C., Pelekis, N., & Doulkeridis, C. (2017). Datastories at semeval-2017 task 4: Deep LSTM with attention for message-level and topic-based sentiment analysis. *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, 747-754. https://doi.org/10.18653/v1/S17-2126
[24] Franco, S., Marco, G., Ah Chung, T., Markus, H., & Gabriele, M. (2008). The graph neural network model. *IEEE Transactions on Neural Networks*, *20*(1), 61-80. https://doi.org/10.1109/TNN.2008.2005605
[25] Chen, W., Zheng, X., Zhou, H. J., & Li, Z. (2021). Evaluation of logistics service quality: Sentiment analysis of comment text based on multi-level graph neural network. *Traitement du Signal*, *38*(6), 1853-1860. https://doi.org/10.18280/ts.380630
[26] Yao, L., Mao, C., & Luo, Y. (2019). Graph convolutional networks for text classification. *Proceedings of the AAAI Conference on Artificial Intelligence*, *33*(1), 7370-7377. https://doi.org/10.1609/aaai.v33i01.33017370
[27] Kumar, A., Singh, K. U., Swarup, C., Singh, T., Raja, L., & Kumar, A. (2022). Detection of copy-move forgery using Euclidean distance and texture features. *Traitement du Signal*, *39*(3), 781-788. https://doi.org/10.18280/ts.390302
[28] Lianzhe, H., Dehong, M., Sujian, L., Xiaodong, Z., Houfeng, W. (2019). Text level graph neural network for text classification. *arXiv preprint arXiv:1910.02356*. https://doi.org/10.18653/v1/D19-1345
[29] Battaglia, P., Pascanu, R., Lai, M., Jimenez Rezende, D. (2016). Interaction networks for learning about objects, relations and physics. *30th Conference on Neural Information Processing Systems (NIPS 2016)*, 4509-4517.
[30] Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735
[31] Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532-1543. https://doi.org/10.3115/v1/D14-1162

Contact information:

Wenjing XUAN
College of Technology,
Hubei engineering university,
Xiaogan 432000, China
Email: wjxuan1982@163.com

Min DENG
(Corresponding author)
School of Computer and Information Science,
Hubei Engineering University,
Xiaogan 432000, China
Email: dm2010@whu.edu.cn