

MULTI-LAYER ATTENTION APPROACH FOR ASPECT BASED SENTIMENT ANALYSIS

Xinzhi Ai¹, Xiaoge Li¹, Feixiong Hu², Shuting Zhi¹ and Likun Hu²

¹School of Computing, Xi'an University of Posts and Telecommunications, Xi'an, China, 710121

²Tencent Technology (Shenzhen) Co., Ltd, China, 518057

ABSTRACT

Based on the aspect-level sentiment analysis is typical of fine-grained emotional classification that assigns sentiment polarity for each of the aspects in a review. For better handle the emotion classification task, this paper put forward a new model which apply Long Short-Term Memory network combine multiple attention with aspect context. Where multiple attention mechanism (i.e., location attention, content attention and class attention) refers to takes the factors of context location, content semantics and class balancing into consideration. Therefore, the proposed model can adaptively integrate location and semantic information between the aspect targets and their contexts into sentimental features, and overcome the model data variance introduced by the imbalanced training dataset. In addition, the aspect context is encoded on both sides of the aspect target, so as to enhance the ability of the model to capture semantic information. The Multi-Attention mechanism (MATT) and Aspect Context (AC) allow our model to perform better when facing reviews with more complicated structures. The result of this experiment indicate that the accuracy of the new model is up to 80.6% and 75.1% for two datasets in SemEval-2014 Task 4 respectively, While the accuracy of the data set on twitter 71.1%, and 81.6% for the Chinese automotive-domain dataset. Compared with some previous models for sentiment analysis, our model shows a higher accuracy.

KEYWORDS

Aspect-level sentiment analysis, Multiple attention mechanism, LSTM neural network

1. INTRODUCTION

Aspect sentiment analysis is one of the important parts of natural language processing (NLP), which provides fine-grained classification of emotions [1]. Such as, the following sentence “Large memory, but performance was poor”, where the polarity of the aspect “memory” shows positive, the polarity of another aspect “performance” reveals negative. It's complicated to judge the polarity of such a multi-aspect review if information about the objective of the aspect is ignored. This is a common challenge in general sentiment classification tasks.

The key of aspect sentiment analysis is to model appropriate context features for an aspect target in a comment. Traditional approaches mostly with manual features to train classifiers. In contrast, neural networks can learn feature representations from given data without manual feature engineering effectively. However, these models largely rely on information from additional analysis such as dependency parsing, which may accumulate errors during the sentiment analysis. It has been proposed that differentiated the both sides context of aspect target based on their locations relative to the aspect word and applied Bi-LSTM networks to encode the both side context respectively. However, the model proposed may not capture the sentiment information

David C. Wyld et al. (Eds): NLP, JSE, MLTEC, DMS, NeTIOT, ITCS, SIP, CST, ARIA - 2020

pp. 131-142, 2020. CS & IT - CSCP 2020

DOI: 10.5121/csit.2020.101410

that is far from the aspect target in the comment. Prior work has applied the advantage of the attention module to solve aspect sentiment analysis problems. The model automatically generates aspect-to-text and text-to-aspect bidirectional attention [7]. In general, prior models with attention module only consider the semantic feature for an aspect target or context, which fails to take into account other features affecting the accuracy of the aspect sentiment analysis, such as context location and class balancing.

The above deficiencies, a model named LSTM-MATT-AC [11] is proposed, which introduce Multiple attention combining mechanisms (MATT) and aspect context module(AC). Where multi-attention mechanism consists of three parts: content attention, location attention, and class attention. The content attention is responsible for considering the semantic information related to the aspect targets in the comments; the location attention mainly focuses on the relative location information between the aspect targets and their context; the class attention is introduced to overcome the classification model data variance caused by an imbalance in the training data. And adopted the aspect context module encodes the bidirectional information of the aspect targets into the sentiment features. Finally, the performance of the model is evaluated on four different language datasets, including Chinese and English. The evaluation results show that the model is widely applicable. In addition, the results verify that our model is insensitive to language.

The others of this paper is arranged as below, the second section introduces some related work; the third part gives an overview of LSTM-MATT-AC model and a detailed description of location, content and class attention; A large number of experiments are carried out in the fourth part, and the results prove the validity of the new model, and last section concludes the paper and presents direction for future improvement.

2. RELATED WORK

2.1. Aspect Sentiment Classification

Traditional methods used sentiment classifiers with expensive hand-crafted features. Most prior work consisted of two steps: (1) A sentiment lexicon was built. For example, use existing dependency relation triple parsers; the dependency parsing knowledge between sentiment words and aspect words; use WordNet's annotation information; Also constructing domain specific sentiment lexicon based on the space similarity of sentiment embedding that contain semantic information. (2) The appropriate classifier was selected. Which contains three ways: one is Naive Bayes, the another is Support Vector Machine (SVM), or combine these two classifiers. Recently, an increasing number of researchers have used neural networks for aspect sentiment analysis. Some typical models that are commonly used include the recursive neural network, and the tree-LSTMs neural network [5]. These models utilize the syntax structures of comments to encode the hierarchical grammatical information of the comments. This work suggests that word embedding and deep neural network could utilize syntactic and semantic structures in comments for aspect sentiment analysis without manual intervention.

2.2. Attention Mechanism

The essence of attention module is to retain useful information, filter out irrelevant information [7], and overcome the limitation of poor performance of recurrent neural network (RNN) in feature coding of long sequence texts. This method associates the target object with each feature in the text, and obtains the corresponding attention probability distribution. In 2014, successfully solved the problem of image classification by introducing the attention mechanism into the RNN [8]. Subsequently, neural networks with attention mechanisms became a heated research topic.

With the advancement of research, many subtasks of natural language processing have got good results using attention mechanisms, such as textual entailment recognition, text abstraction extraction, speech recognition, and machine reading comprehension. Prior work also validated the effectiveness of attention mechanisms in aspect sentiment analysis tasks.

3. MODEL

3.1. Task Definition

First, the sentiment analysis tasks, we need to give a fixed sentence of length n $s = \{w_1, \dots, w_l, \dots, w_r, \dots, w_n\}$ and an aspect target $a = \{w_l, \dots, w_r\}$. For each word w_i , we obtain an embedding $v_i \in R^{d_w}$ from $L \in R^{d_w \times |V|}$, where $|V|$ refers to size of the vocabulary, d_w is used to indicate dimension of embedding word.

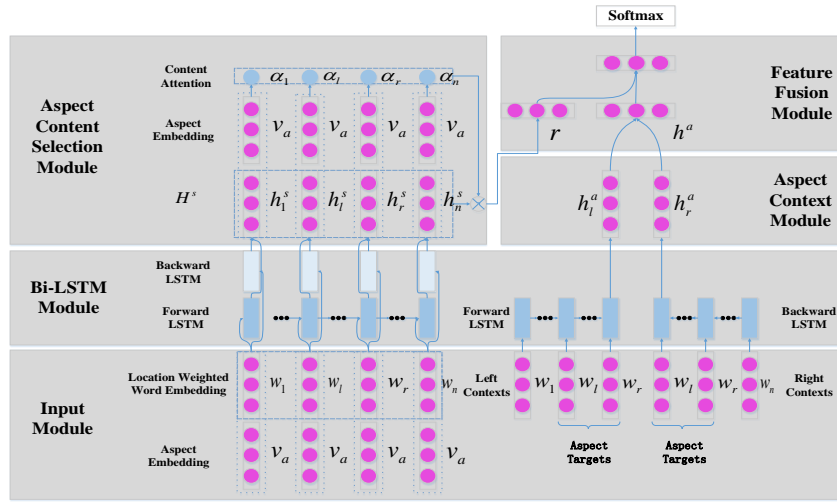


Figure 1. The architecture of LSTM-MATT-AC model

3.2. An Overview of Model

Input module: The first module is input embedding which includes aspect, location-weighted word and aspect context, in which aspect embedding need to map each aspect word into embedding $v_a \in R^{d_w}$, while location-weighted word embedding refers to integrating the location distribution, and generate corresponding weight embedding, and aspect context embedding splits sentence word embedding sequence of into two parts on both sides of the context.

Bi-LSTM module: The aspect embedding and the location-weighted word embedding are combined as input of the left Bi-LSTM module. We obtain the forward hidden states $\vec{H}^s \in R^{d_h \times n}$ and backward hidden states $\overleftarrow{H}^s \in R^{d_h \times n}$ from Bi-LSTM simultaneously. Concatenating \vec{H}^s and \overleftarrow{H}^s generate output hidden states $H^s \in R^{2d_h \times n}$, where d_h refers to the dimension of the hidden states.

$$H^s = \vec{H}^s \parallel \overleftarrow{H}^s \quad (1)$$

Then embed context of the aspect into the Bi-LSTM neural network to the right of the Bi-LSTM module. Similar to the left part, the hidden state in two directions will be generated.

$$\vec{H}^a = LSTM(\vec{[v_1, v_2, \dots, v_l, \dots, v_r]}) \quad (2)$$

$$\overleftarrow{H}^a = LSTM(\overleftarrow{[v_1, v_2, \dots, v_r, \dots, v_{n+1}, v_n]}) \quad (3)$$

Aspect content selection module: The module main task that hidden state is multiplied by the content attention distribution $\alpha \in R^{1 \times n}$, to achieve the feature embedding which is highly relevant to aspect targets.

$$r = H^s \alpha^T \quad (4)$$

Aspect context module: The feature embedding $h^a \in R^{2d_h}$ through concatenate the hidden states $\vec{h}^a \in R^{d_h}$ and $\overleftarrow{h}^a \in R^{d_h}$ is generated.

Feature fusion: By $r \in R^{2d_h}$ and $h^a \in R^{2d_h}$ are integrated with the expression with the following equation to come into being the final feature embedding $h^* \in R^{d_h}$.

$$h^* = \tanh(W_r r + W_a h^a) \quad (5)$$

3.3. Location Attention

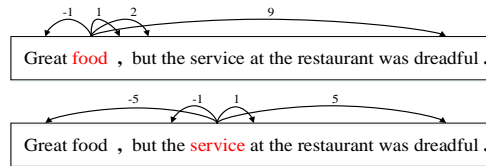


Figure 2. The relative distances between words and aspect targets in sentences

For sentiment analysis tasks, the emotional relative with the distance between the target and the context provides crucial information. In the sentence shown in Figure 2, “great” is the sentiment word for the aspect target “food”, and “great” is closer to the aspect target “food” than to “dreadful”. However, for another aspect “service”, the relative distance to the relevant sentiment word “dreadful” and the irrelevant sentiment word “great” is the same. This phenomenon poses a threat to the simple location weight calculation. In this paper, citing the punctuation-based algorithm proposed by Han [11] to be calculated.

3.4. Content Attention

According to the aspect target, the hidden states H^s concatenation with the aspect embedding v_a to produce content attention distribution α .

$$M = \tanh(W_h H^s; W_v v_a \otimes e_n) \quad (6)$$

$$\alpha = \text{soft max}(w^T M) \quad (7)$$

Where $M \in R^{(d_h+d_w) \times n}$, $\alpha \in R^{1 \times n}$, $W_h \in R^{d_h \times d_h}$, $W_v \in R^{d_w \times d_w}$ and $w \in R^{d_w+d_h}$ are projection parameters, $v_a \otimes e_n = [v_a, v_a, \dots, v_a]$ means that v_a is repeatedly concatenated n times, and e_n refers to column vector with n 1s.

3.5. Class Attention

To avoid the bias problem of the LSTM-MATT-AC model, this paper introduces the loss function [3] into the class attention, which penalizes the misclassification of underrepresentation more seriously. In addition, in the case of extremely unbalanced training data, smoothing factor φ is introduced to smooth the weights, which may lead to very high class attention weights.

$$\omega_i^c = \frac{\max(c)}{c_i + \varphi \times \max(c)} \quad (8)$$

Where C is the list containing the number of data points under each class, and C_i refers to the amount of data points under the i -th class. For the data from the Restaurant, Automobile and Laptop domains, we set the parameter $\varphi = 0$, for the Twitter domain data, set the parameter $\varphi = 0.1$.

3.6. Model Training

The end-to-end training model is established by using the back propagation, and the cross entropy loss is selected as the loss function. To avoid over fitting, we add L2 regularization to the loss function [10].

$$\text{loss} = - \sum_i \sum_j y_i^j \log \hat{y}_i^j + \lambda \|\theta\|^2 \quad (9)$$

Where y_i^j is the correct sentiment, \hat{y}_i^j indicates the prediction result of the model, i refers to index of review data, j means index of class of the final sentiment, λ as L2 regularization items, and θ present parameter set.

4. EXPERIMENTS

4.1. Dataset

New proposed model LSTM-MATT-AC on four data sets were evaluated, as shown in Table 2. Firstly, we evaluated our model on three more English datasets from SemEval-2014 Task 4 and Twitter. For different languages have considerably different expressions for sentiments. After that, a dataset from the Chinese Automobile domain were used for comparison.

Table 2. Details of the experimental datasets

Data Set	Train			Test			Total
	Positive	Neutral	Negative	Positive	Neutral	Negative	
Restaurant	2148	631	894	720	194	209	4796
Automobile	953	57	528	295	48	144	2025
Laptop	989	468	868	337	167	131	2960
Twitter	1561	3127	1560	173	346	1743	8510

4.2. Hyper parameters Setting

In our evaluation, we selected the training data of 20% randomly as the testing data to tune a set of hyper parameters. Word embedding is initialized by the 300-dimensional size Glove vectors [7]. We use $U(-0.01, 0.01)$ to uniformly distribute random initialization of off-vocabulary words. The LSTM hidden state dimension is set as 150. All parameters of the random initialization, equal distribution of $U(-0.05, 0.05)$. Our model for the batch of 25 samples of training.

4.3. Model Comparison

We have four domain-specific datasets will be in the field of the new proposed model and the following baseline model are compared.

- TD-LSTM (Target-dependent LSTM) using two LSTM network context is modelled as a target. To connect two LSTM network finally hide status for the sentiment prediction.
- TC-LSTM [12] (Target-connection LSTM) extends the TD-LSTM, which combines aspect embedding and the embedding of each word.
- AT-LSTM (Attention-based LSTM) Based on the attention of the LSTM (attention - -based LSTM) by LSTM network context modelling to target first, and then combining embedded LSTM hide status and ways, to generate attention to weight.
- ATAE-LSTM [14] is an aspect-oriented embedded attention-based LSTM developed on the basis of attention-LSTM.
- AOA-LSTM [6] (attention-over-attention LSTM) First, aspects and text are modelled simultaneously through Bi-LSTM. Then, the fusion between the aspect target and the text representation produces bidirectional attention.

4.4. Result and Discussion

Next apply accuracy metric to confirm results and performance analysis of the model classifier, which defined as follows:

$$acc = \frac{num_{cor}}{num_{all}} \quad (10)$$

Where num_{cor} represents the amount of correctly predicted samples, num_{all} represents the total amount of samples and acc measures the percentage of the correctly predicted samples in all sample data [11].

4.4.1. 3-class Sentimental Classification

Table 3. Comparison results: accuracy of the 3-class sentimental classification the Best performances in bold

Model	Restaurant	Laptop	Twitter	Automobile
TD-LSTM	0.756	0.681	0.646	0.786
TC-LSTM	0.763	0.710	0.680	0.792
AT-LSTM	0.762	0.689	0.672	0.788
ATAE-LSTM	0.772	0.687	0.681	0.795
AOA-LSTM	0.812	0.745	—	—
LSTM-MATT-AC	0.806	0.751	0.711	0.816

Compared with the unidirectional LSTM used by the ATAE-LSTM, the LSTM-MATT-AC utilizes the Bi-LSTM to better encode the context semantic features of the comments. The aspect content attention is appended to the output layer of the Bi-LSTM to generate the weight distribution of the semantic correlation between the aspect and context, and then the weighted total of the hidden states from the Bi-LSTM generates feature embedding that is closely related with aspect target. To address the model data variance caused by the class imbalance in the training data, we introduce the class attention mechanism into the loss function of the model for capture deeper sentiment features from the class with less data.

For aspect phrases, different words contribute to the aspect expression differently. Attention mechanism that phrase is an important part of the model is able to focus on, so as to promote the accuracy of sentiment classification [11]. In general, our model achieves better results on datasets of four different domains and two different languages than other models. The high performance of our model shows that it has a good domain migration ability and is language-insensitive.

4.4.2.2-class sentimental classification

For further compare the performance of new model with other baseline models in aspect sentiment analysis, we conducted another experiment with only the data of "Positive" and "Negative" in four datasets. The experimental results are shown in the following table.

Table 4. Comparison results: accuracy of 2-class sentimental classification. the Best performances in bold

Model	Restaurant	Laptop	Twitter	Automobile
ASGCN	0.806	0.755	0.721	—
AEN-BERT[15]	0.831	0.799	—	—
TD-LSTM	0.892	0.868	0.830	0.869
TC-LSTM	0.894	0.853	0.858	0.878
AT-LSTM	0.892	0.874	0.858	0.873
ATAE-LSTM	0.909	0.876	0.867	0.882
LSTM-MATT-AC	0.910	0.870	0.868	0.897

Table 4 reveals that the accuracy of the models were significantly improved after the "Neutral" data was removed. We investigated the "Neutral" reviews in the four datasets, and found that most of the "Neutral" comments are objective statements about aspect targets, and do not contain sentiment expressions from reviewers. For example, the review "It takes about 2 hours to be served our 2 courses." does not express reviewer's sentiment tendency. Another finding is that

some "Positive" and "Negative" comments are too implicit about the sentiment of aspect targets and the syntactic structures are relatively complex, resulting in the model not able to accurately capture the sentiment features in the training phase, thus misclassifying the reviews into "Neutral" class. In conclusion, the objectivity of "Neutral" review expression and the ambiguity of "Neutral" sentiment class itself lead to "Neutral" data greatly affecting the accuracy of model. The accuracies of the LSTM-MATT-AC on the Restaurant, Twitter and Automobile datasets are higher than that of other models, reaching 91%, 86.8%, and 89.7% respectively. Our experiments show that the LSTM-MATT-AC can better deal with aspect sentiment classification of different domains and languages.

We conducted another experiment to proof the validity of the multi-attention module, such as the below models to compare: (1) the LSTM-NL-AC model which has no location attention compared with model presented in this paper; (2) the LSTM-NC-AC model which has no class attention compared with the new proposed model. For this experiment, we just kept "Positive" and "Negative" part data from the four domain-specific datasets, as a binary sentiment classification experiment. The evaluation results are shown in Figure 3. The overview of the four datasets is shown in Figure 4.

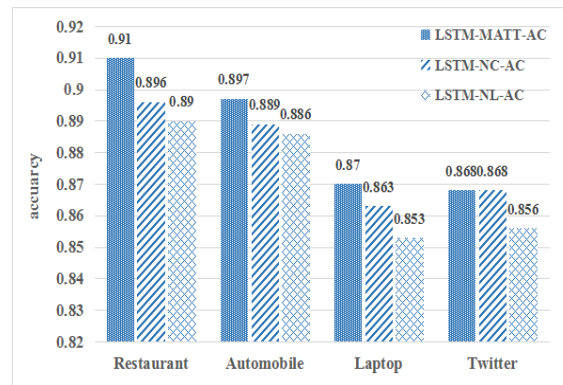


Figure 3. The impact of the multi-attention module on the classification accuracy

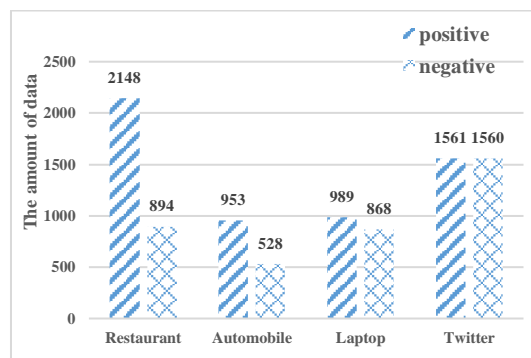


Figure 4. Data overview of the two-classification tasks

From figure 3, the LSTM-MATT-AC model reaches optimal performance. After the addition of location attention, the model at least 1.1% improvement in the datasets compared with the no location attention model; With the class attention, other accuracy rates have improved, except that twitter remained consistent. Figure 4 shows that class distributions of Restaurant, Laptop and Automobile datasets are imbalanced compared with Twitter dataset. Furthermore, we observed that, compared with the location attention, class attention to improve model performance

contribution. To study the impact of aspect context module the accuracy of aspect sentiment classification, we constructed two models named LSTM-MATT-1 and LSTM-MATT-2, replacing the aspect context module with and respectively. Then the three LSTM-MATT models (i.e., LSTM-MATT-1, LSTM-MATT2, and LSTM-MATT-AC) were used to conduct another aspect sentiment analysis experiment on four domain-specific datasets. The evaluation results are shown in Figure 5.

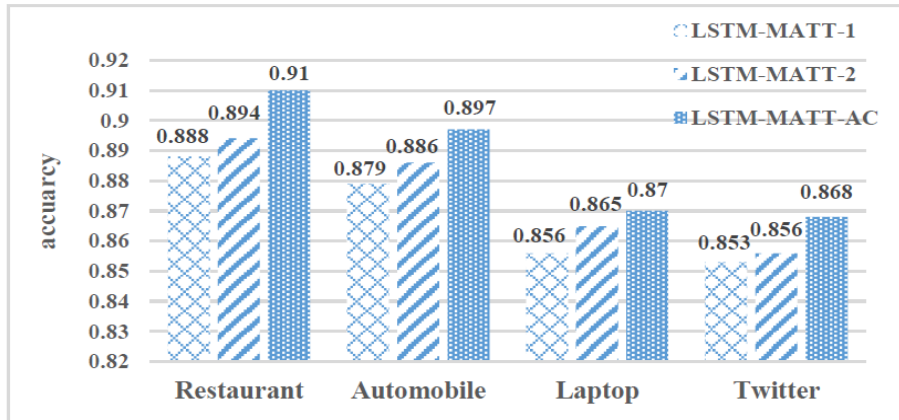


Figure 5. The impact of different aspect context processing methods on classification accuracy

Figure 5 shows that the classification accuracy of the LSTM-MATT-2 is generally higher than that of the LSTM-MATT-1 on four domain-specific datasets. This indicates that encodes richer aspect context information than, and also verifies that the LSTM with three threshold mechanisms can effectively learn long-distance semantic dependencies in reviews. The accuracy of the LSTM-MATT-AC is higher than that of the LSTM-MATT-2, with improvements of 1.79%, 1.24%, 0.58%, and 1.4% in the Restaurant, Automobile, Laptop, and Twitter domains respectively. The result shows that aspect context module, which takes an aspect target as a target object and encodes the left and right context of aspect word with Bi-LSTM network, can make the LSTM-MATT-AC generate more accurate sentiment features with the aspect, and ultimately raise the accuracy of aspect sentiment classification.

Table 5. Attention weights of the word sequence in different models

Model	great	food	,	but	the	<u>service</u>	was	dreadful	!
LSTM-NP-AC	0.2880	0.0311	0.0311	0.0237	0.0269	0.0643	0.0463	0.4606	0.0280
LSTM-NC-AC	0.1783	0.1705	0.0885	0.0769	0.0771	0.0755	0.0874	0.2000	0.0458
LSTM-MATT-AC	0.0043	0.0021	0.0007	0.0012	0.0004	0.1250	0.0012	0.8691	0.0006

To further study the effect of the location and class attention on the aspect sentiment classification, we extracted a sample from the Restaurant dataset. We then compared the attention weights of word sequences among LSTM-MATT-AC, LSTM-NL-AC and LSTM-NC-AC. Table 5 shows that for the aspect target "service", the LSTM-NL-AC without location attention assigns the sentiment words "great" and "dreadful" a higher weight, while the LSTM-MATT-AC can accurately identify the correct sentiment word "dreadful" and increase its weight from 0.4606 to 0.8691. This weight fluctuation shows that the model with location attention can effectively capture the influence of words on the aspect target through the relative location information

between an aspect target and each word in a review, enhance the model's content attention layer to recognize correct sentiment words, and thus generate more accurate aspect sentiment features through more reasonable weight allocation. On the other hand, because the number of positive reviews from the Restaurant dataset is about 2.4 times the number of negative reviews, the class is imbalanced. For the aspect target "service" whose sentiment polarity is negative in the instance, we find that the LSTM-NC-AC assigns weights to each word more evenly in the clause "but the service was dreadful!", and the sum of the weights allocated to the clause is 0.5169, which is much lower than the sum of the 0.9949 assigned to the clause by the LSTM-MATT-AC. After analysing the attention weights of word sequence, we found that the model will likely be biased when the training data is imbalanced. From the perspective of attention distribution, we find that the model reduces the attention to the class clauses with less data, which is mainly manifested in the lower weight sum given to clauses and the more average weight assigned to each word in the clauses. The above demonstrates the validity of the new model that introduces location attention and class attention in aspect sentiment analysis.

4.5. Visualizing Attention Mechanism

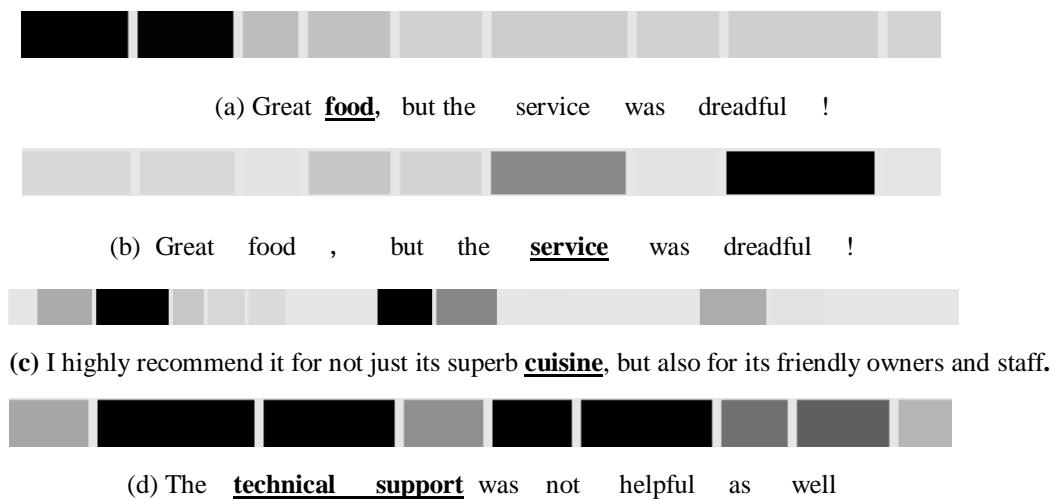


Figure 6. Attention visualizations. the color darkness reveals the importance degree

In Figure 6, we list four examples from the Restaurant and Laptop datasets, and visualizing the weight of each word's attention to determine which has the greatest effect on the emotional polarity of the aspect. The darkness of the colour indicates the grade of the attention weight [20]. The darker the colour is, the more important the word is. In the first two examples, the comment "Great food, but the service was dreadful!" contains aspect of the "food" and another aspect "service". However, the emotional polarity of the aspect "food" reveals positive. However, the polarity of the "service" represents opposite, as two clauses with opposite sentiment polarity are connected by the word "but". The results of the attention visualizations in Figure 6 (a) and Figure 6 (b) show that our model accurately finds the corresponding sentiment words "great" and "dreadful" of aspect words "food" and "service" under the premise of identifying the turning sentences and thus predicts the correct sentiment polarity of the two aspect targets. In the third example, the correct polarity of the aspect word "cuisine" is positive in comment "I highly recommend it for not just its superb cuisine, but also for its friendly owners and staff". Although there is a negative word "not" in front of the word "cuisine", our model recognizes that the "not" does not mean negative in this review and therefore accurately classifies the sentiment of the aspect target "cuisine". The last example, the polarity of the aspect "technical support" is negative in comment "The technical support was not helpful as well". Compared to the third example, our

model recognizes that "not" is a negation of the sentiment word "helpful" that is positive and therefore accurately predicts the correct emotional polarity of the aspect.

5. CONCLUSION

This paper proposed a new model LSTM-MATT-AC to tackle challenges in aspect sentiment analysis. The model leveraged the multi-attention mechanism and aspect context, where multiple attention makes the new model to concentrate the relative between aspect targets and their context words from different perspectives. And accurately represent the impact of each context words on the aspect targets. Meanwhile, because different aspect targets have different contexts in a review, the aspect context module of the LSTM-MATT-AC independently encodes the both sides context of the aspect targets through Bi-LSTM. Evaluations on four domain-specific datasets indicate that new proposed model has improved from the accuracy compared to models from prior work, validating the high-performance of the multi-attention mechanism and aspect context on aspect sentiment analysis.

As our model averages the each word embedding in an aspect phrase, the features of important words in the aspect phrase may not be captured effectively. We plan to study how the introduction of the aspect of future work about the attention mechanism.

ACKNOWLEDGEMENTS

The paper was supported by the Shaanxi science and technology innovation foundation under Grant NO. 2019PT-12, the Graduate Innovation Foundation Project of Xi'an University of Posts & Telecommunications (CXJLY2019094) and Tencent Technology (Shenzhen)Co., Ltd.

REFERENCES

- [1] Fan X, Li X, Du F, et al. Apply word vectors for sentiment analysis of APP reviews[C]// International Conference on Systems and Informatics. IEEE, 2017:1062-1066.
- [2] Ilmania, Abdurrahman, S. Cahyawijaya and A. Purwarianti, "Aspect Detection and Sentiment Classification Using Deep Neural Network for Indonesian Aspect-Based Sentiment Analysis," 2018 International Conference on Asian Language Processing (IALP), Bandung, Indonesia, 2018, pp. 62-67.
- [3] X. Fang and J. Tao, "A Transfer Learning based Approach for Aspect Based Sentiment Analysis," 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS), Granada, Spain, 2019, pp. 478-483.
- [4] R. S. Silva, R. R. M. U. A. Rathnayaka, P. M. A. K. Wijeratne, M. T. N. Deshani, Y. H. P. P. Priyadarshana and D. Kasthurirathna, "Ontological Approach for Aspect-Based Sentiment Analysis," 2019 IEEE 5th International Conference on Computer and Communications (ICCC), Chengdu, China, 2019, pp. 106-111.
- [5] Ruder S, Ghaffari P, Breslin J G. A Hierarchical Model of Reviews for Aspect-based Sentiment Analysis[J]. 2016.
- [6] Rocktäschel T, Grefenstette E, Hermann K M, et al. Reasoning about Entailment with Neural Attention[J]. 2015.
- [7] Ma D, Li S, Zhang X, et al. Interactive Attention Networks for Aspect-Level Sentiment Classification[J]. 2017.
- [8] Liang B, Liu Q, Xu J, et al. Aspect-based sentiment analysis based on multi-attention CNN[J]. Journal of Computer Research and Development, 2017, 54(08):1724-1735.
- [9] Tang D, Qin B, Liu T. Aspect Level Sentiment Classification with Deep Memory Network[J]. 2016.
- [10] Baziotis C, Pelekis N, Doukeridis C. DataStories at SemEval-2017 Task 4: Deep LSTM with Attention for Message-level and Topic-based Sentiment Analysis[C]// International Workshop on Semantic Evaluation. 2017:747-754.

- [11] Han H, Li X, Zhi S, et al. Multi-Attention Network for Aspect Sentiment Analysis[C]// Proceedings of the 2019 8th International Conference on Software and Computer Applications. 2019:22-26.
- [12] Lingling Song, Yazhou Zhang, Yuexian Hou. Convolutional neural network with pair-wise pure dependence for sentence classification[C]// 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), 2018:78-95.
- [13] Zhai Penghua, Zhang Dingyi. Bidirectional-GRU Based on Attention Mechanism for Aspect-level Sentiment Analysis[C]// Proceedings of the 2019 11th International Conference on Machine Learning and Computing - ICMLC, 2019:56-64.
- [14] Zhang, C., Li, Q., Song, D. (2019). Aspect-based Sentiment Classification with Aspect-specific Graph Convolutional Networks. The 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP2019). 3-7 November 2019, Hong Kong, China.
- [15] Song, Y. Wang J, Jiang T, Liu Z and Rao Y. 2019. Attentional encoder network for targeted sentiment classification. arXiv preprint arXiv:1902.09314.
- [16] Huang, Binxuan & Carley, Kathleen M. (2018). Parameterized Convolutional Neural Networks for Aspect Level Sentiment Classification. In Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (EMNLP), ACL.

AUTHORS

Xinzhai Ai received B.Eng. in electrical and electronic engineering from Hao Jing College of Shaanxi University of Science & Technology in 2018, she is pursuing a master's degree at Xi'an University of Posts and Telecommunications. As a research assistant at text mining Lab., Her research interests include deep learning, data mining and aspects sentimental analysis.



Xiaoge Li is a Professor of Computer Science at Xi'an University of Posts and Telecommunications. In 2000, He was with Cymfony, NY., USA engaging in research and development of information Extraction and Question & answering system. His currently research include natural language processing, Machine learning and knowledge Graphing.



Feixiong Hu received his B.Eng. in Computer Engineering from Xi'an University of Posts and Telecommunications in 2012. He is currently the leader of Artificial intelligence for IT operations intelligent operation (AIOps) of Shenzhen Tencent Computer System Co., Ltd. His current research interests are big data and machine learning.



Shuting Zhi received her B.Eng. (2016) and M.Eng. (2018) in Computer Engineering from Xi'an University of Posts and Telecommunications. Since 2018, she has joined Xiaomi Corporation, China. She has been working on the projects of xiaoai knowledge platform, word sense ambiguity and sentimental analysis.



Likun Hu received the master's degree in computer system structure from Northeast Forestry University in 2017. Currently, He is working in Shenzhen Tencent Computer System Co., Ltd., as a senior operations engineer of Shenzhen Tencent Computer System Co., Ltd., China. He has being engaged in development of big data analysis and log Anomaly Detection.

