Towards Aspect-level Sentiment Modification without Parallel Data

Qingnan Jiang
Chinese Academy of Sciences

Lei Chen
Chinese Academy of Sciences

Wei Zhao
Technische Universität Darmstadt

Min Yang
Chinese Academy of Sciences

***Abstract*—This work takes the lead to study aspect-level sentiment modification (ALSM) without parallel data. Given a sentence, the task of ALSM needs to reverse the sentiment with respect to the given aspect while preserving other content. The main challenge is reversing the sentiment of the given aspect without affecting the sentiments of other aspects in the sentences. To handle this problem, we propose a joint aspect-level sentiment modification (JASM) model. JASM is a multi-task system, which jointly trains two coupled modules: aspect-specific sentiment words extraction and aspect-level sentiment transformation. Besides, we propose a novel memory mechanism to learn aspect-aware sentiment representation and a gating mechanism to dynamically select aspect-aware sentiment information or content information for generating the next words. Experiments show that the proposed model substantially outperforms the compared methods in both aspect-level sentiment transformation and content preservation. For applications, we conduct data augmentation for aspect-based sentiment analysis (ABSA) through generating plausible training data with the ALSM model. Experiments show that data augmentation with generated data boosts the performance of a broad range of ABSA models.**

SENTIMENT MODIFICATION of natural language texts has attracted increasing attention recently since it can facilitate many NLP applications such as the conversion of review attitude. Existing studies convert the sentiment for a whole piece of text, such as a sentence/document [1], [2]. However, in the real world, people may mention several target entities in one sentence/document [3]. It is essential to study the methods of sentiment modification with respect to a given aspect, i.e., aspect-level sentiment modification (ALSM).

Several challenges exist in aspect-level text sentiment modification. First, the parallel data with the same content but different aspect-level sentiment polarities is usually not available since it is time-consuming and costly to annotate a parallel dataset for ALSM task. Second, since a sentence could contain multiple sentiment-aspect pairs, differentiating sentiments towards different aspects and conversing the corresponding sentiment is challenging. Third, it is difficult to identify the aspect-irrelevant content in implicit ways, which makes the previous models tend to generate input-irrelevant content and lead to poor content preservation.

To alleviate the aforementioned challenges, we propose a **j**oint **a**spect-level **s**entiment **m**odification (JASM) model. Our model simultaneously trains two coupled modules: an aspect-specific sentiment words extraction module and an aspect-level sentiment transformation module, which share the same sentence modeling component so as to identify the aspect-specific sentiment words better. Specifically, the aspect-specific sentiment words extraction module is a sequence labeling model, which is supposed to detect aspect-specific sentiment words and produce a revised sentence with sentiment words being masked with the corresponding sentiment labels. The aspect-level sentiment transformation module is a sequence-to-sequence (seq2seq) model equipped with a sentiment memory to learn aspect-aware sentiment representation and a gating mechanism that dynamically selects sentiment information or content information in the decoding stage. Conditioning on the revised sentence, the aspect-level sentiment transformation module is used to generate the final output sentence. Finally, we jointly optimize these two modules in the training procedure to improve the performance of detecting aspect-specific sentiment words.

Our main contributions are summarized as follows. (1) To the best of our knowledge, this is the first study to deal with the target-specific text sentiment modification. (2) We propose an effective model that significantly outperforms compared methods in both target-specific sentiment transformation and content preservation. (3) We use data generated by target-specific sentiment modification models to boost the performance of aspect-based sentiment analysis models.

RELATED WORK

Sentiment analysis is a suitcase research problem that encapsulates many NLP tasks [4], [5]. The ALSM task, in particular, integrates both the methods of ABSA and sentiment modification. ABSA includes a few subtasks, such as aspect extraction and aspect-based sentiment classification. Aspect extraction aims to identify opinion targets in opinionated text. [6] applied a 7-layer convolutional neural network to classify each word in the input sentence as either aspect or non-aspect word. [7] proposed a novel LSTM-based deep multi-task learning framework to extract aspect terms from user review sentences. To identify the sentiment polarity towards the given aspect, there have been many studies attempting to learn aspect-based sentiment classifiers with deep neural networks and attention mechanisms [3], [8]. For example, [3] proposed an interactive attention mechanism to learn the representations of context and aspect words interactively.

Recently, great efforts have been made to study the text sentiment modification task that aims to generate a sentence with similar content but opposite sentiment with the original sentences [1], [2], [9]. For example, [1] proposed a cross-aligned auto-encoder model for unpaired text style transfer.

[2] proposed to learn sentiment memories for sentiment modification. [10] proposed a method that combines the strength of the retrieval model and the generative model for text attribute modification. [9] proposed an adversarial regularization auto-encoder for text style transfer.

METHODOLOGY

Given a sentence $X = \{x_1, \ldots, x_n\}$, an aspect term $A = \{a_1, \ldots, a_k\}$ and the sentiment label S of the given aspect, the goal of ALSM is to generate a new sentence $Y = \{y_1, \ldots, y_m\}$ which reverses the sentiment towards the given aspect while preserving other content. Here, n, k and m indicate the lengths of X, A and Y, respectively. As illustrated in Figure 1, our model contains two modules: an aspect-specific

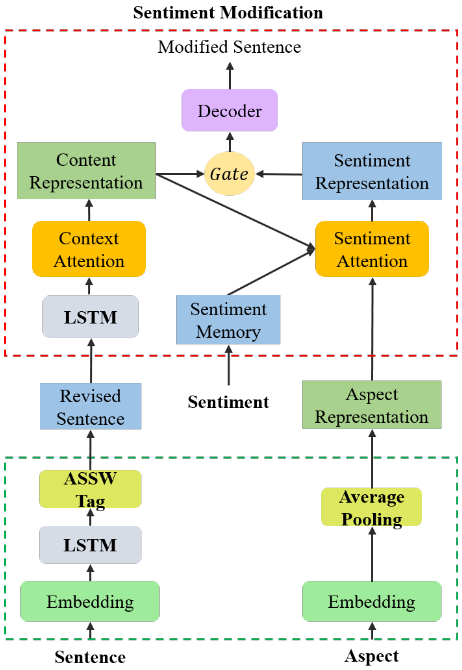

Figure 1: The overview of the proposed ALSM model.

sentiment words extraction module and an aspect-level sentiment transformation module. Next, we will introduce these two components in detail.

Aspect-Specific Sentiment Words Extraction Module

The aspect-specific sentiment words extraction module is a BiLSTM [11] based sequence tagger enhanced with the positional embedding. Let q_i denote whether x_i is an aspect word.

$$q_i = \begin{cases} 1 & \text{if } x_i \text{ is an aspect word} \\ 0 & \text{if } x_i \text{ is not an aspect word} \end{cases} \quad (1)$$

Let $e(x_i)$ and $e_p(q_i)$ denote the embeddings of x_i and q_i. We obtain the positional-aware embedding by concatenating the two above embeddings:

$$e_{pa}(x_i) = [e(x_i); e_p(q_i)] \quad (2)$$

Then, we compute the forward and backward context-aware representations for each word using LSTM:

$$\mathbf{u}_i^f = \text{LSTM}(\mathbf{u}_{i-1}, e_{pa}(x_i)) \quad (3)$$
$$\mathbf{u}_i^b = \text{LSTM}(\mathbf{u}_{i+1}, e_{pa}(x_i)) \quad (4)$$

We get the tagging features for each word by concatenating the forward and backward context-aware representations.

$$\mathbf{r}_i = [\mathbf{u}_i^f; \mathbf{u}_i^b] \quad (5)$$

The tagging features $\mathbf{r}_i$ are then fed into a softmax classifier to predict if x_i is an aspect-specific sentiment word:

$$P(T_i|X, A) = \text{softmax}(W_s \mathbf{r}_i + \mathbf{b}_s) \quad (6)$$

where W_s and $\mathbf{b}_s$ are parameters to be learned. We train this aspect-specific sentiment words tagger by minimizing the cross-entropy loss:

$$L_1 = -\frac{1}{N} \sum_{i=1}^{N} \frac{1}{n} \sum_{j=1}^{n} P(T_{ij}|X_i, A_i) \quad (7)$$

where N is the number of training instances.

Aspect-level Sentiment Transformation Module

To perform ALSM, we first replace the extracted aspect-specific sentiment words with corresponding sentiment masks (i.e., negative or positive) in a soft way.

We scale the logits by an annealing coefficient τ. The probability p_i that the i-th word is an aspect-specific sentiment word is used as the soft mask. We perform the replacing operation through the linear transformation.

$$p_i = \text{softmax}(\text{logit}_i / \tau) \quad (8)$$
$$e'(x_i) = (1 - p_i) e(x_i) + p_i e(S) \quad (9)$$

where $e(S)$ is the embedding of sentiment label S for aspect A in sentence X, $e'(x_i)$ is the embedding of i-th word in the revised sentence.

Encoder The ALSM module is a sequence-to-sequence (seq2seq) based generation model. The encoder of the seq2seq model is an LSTM network that encodes the revised sentence into hidden states. Formally, the hidden state $\mathbf{h}_i$ at time step i can be updated based on hidden state $\mathbf{h}_{i-1}$ and current input $e'(x_i)$:

$$\mathbf{h}_i = \text{LSTM}(\mathbf{h}_{i-1}, e'(x_i)) \quad (10)$$

Hence, we can obtain the hidden states of the revised sentence as $H = [\mathbf{h}_1, \ldots, \mathbf{h}_n]$.

Sentiment	Training	Validation	Test
Positive	144010	3613	3595
Negative	30802	1387	1405

Table 1: Statistics of training, validation and test data.

Decoder The decoder is an LSTM network equipped with a sentiment memory to store sentiment information and a gating mechanism. Let $M \in R^{C\times\gamma\times d}$ denote the sentiment memory, where C is the number of sentiment categories, γ is the memory size, and d is the dimension of the hidden states. At each decoding step t, we first compute a content vector $\mathbf{c}_t$ using the context attention as:

$$\mathbf{c}_t = \sum_{i=1}^{n} \alpha_i^c \mathbf{h}_i \quad (11)$$

$$\alpha_i^c = \text{softmax}(\mathbf{h}_i W_c \mathbf{s}_t) \quad (12)$$

where W_c is a learnable parameter in bilinear term, $\mathbf{s}_t$ is the decoder hidden state at t-th time step.

Then, we extract contextualized sentiment representation $\mathbf{z}_t$ from the sentiment memory through sentiment attention:

$$\alpha_i^z = \text{softmax}(M_{S,i} W_z [\mathbf{c}_t; \mathbf{s}_t; \mathbf{a}]) \quad (13)$$

$$\mathbf{z}_t = \sum_{i=1}^{\gamma} \alpha_i^z M_{S,i} \quad (14)$$

where W_z is a parameter, $\mathbf{a}$ is the aspect representation, $M_{S,i}$ is the i-th row of sentiment memory with label S. We design a gate g_t that dynamically chooses content information or sentiment information to predict the word at time step t:

$$g_t = \text{sigmoid}(W_g[\mathbf{c}_t; \mathbf{z}_t; \mathbf{s}_t]) \quad (15)$$

$$\mathbf{o}_t = W_o[\mathbf{s}_t; (1-g_t)\mathbf{c}_t + g_t \mathbf{z}_t] \quad (16)$$

where W_g, W_o are learnable parameters. The output vector $\mathbf{o}_t$ is used to predict the word at time step t and update the decoder hidden state $\mathbf{s}_t$ at time step $t+1$:

$$\mathbf{s}_{t+1} = \text{LSTM}(\mathbf{s}_t, [e(y_t); \mathbf{o}_t]) \quad (17)$$

$$P(y_t|Y_{<t}; X, A, S) = \text{softmax}(W_p \mathbf{o}_t + \mathbf{b}_p) \quad (18)$$

where W_p and $\mathbf{b}_p$ are parameters to be learned.

Since there is no parallel data, we train the model to reconstruct the original sentences, similar to previous work [1]. We optimize the model by minimizing the the cross-entropy loss:

$$L_2 = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{m} P(y_{i,j}|Y_{i,<j}, A_i, S_i) \quad (19)$$

where $y_{i,j}$ is the j-th word of the i-th sentence in the training set.

Method	BLEU	APD	ASTA
CAAE	1.30	12.89	11.30
SMAE	4.18	18.10	11.65
ARAE	3.20	18.92	18.76
DAR	35.28	54.50	11.14
ReALSM	10.11	31.12	28.44
Seq2Seq	14.07	67.40	43.34
JASM	**39.82**	**91.34**	**59.24**

Table 2: Automatic evaluation results.

Method	Trans	Cont	Fluency
CAAE	1.96	1.87	4.02
SMAE	2.11	2.05	3.13
ARAE	2.32	1.96	4.26
DAR	2.02	3.35	3.56
ReALSM	2.39	3.17	**4.78**
Seq2Seq	3.24	2.93	4.07
JASM	**3.65**	**3.89**	4.16

Table 3: Human evaluation results.

Joint Learning Overall, our model consists of two subtasks: aspect-specific sentiment words extraction (L_1) and ALSM (L_2), each has a training objective. We pre-train the aspect-specific sentiment words extraction module to give it the initial extraction ability. Then, we train these two related tasks simultaneously. The joint objective function is minimized as:

$$L = \lambda_1 L_1 + \lambda_2 L_2 \quad (20)$$

where λ_1 and λ_2 are hyper-parameters. We empirically show that setting $\lambda_1 = 0.1$ and $\lambda_2 = 0.9$ achieves the best performance of our model.

Sentiment Modification At the testing stage, the detected sentiment words in the sentence will be replaced by opposite sentiment masks, and the sentiment information will be extracted from the opposite sentiment memory to generate the target sentence with reversed sentiment for the given aspect.

EXPERIMENTS

Data Collection

We train the aspect-specific sentiment words extraction model using the TOWE dataset [12].

Sentence: the food is amazing, the ambience is fantastic, the service is attentive, and the prices are outstanding.	
Aspect: ambience	**Sentiment polarity:** positive
CAAE: the food is small and the service is n't worth the price.	
SMAE: the food lacks [UNK] is cold and the service is slow and the food is only [UNK].	
ARAE: the food is average as well as the food , service and prices are for the prices.	
DAR: the food is excellent, the ambience is very disappointed that , the atmosphere is not his [UNK].	
ReALSM: on top of that , the ambience is non - existent , the service is gruff , and the place is always packed.	
AspectSeq2Seq: the food is completely , the ambience is terrible , the service is atmosphere , and the staff are ok.	
JASM: the food is amazing , the ambience is horrible , the service is attentive , and the prices are outstanding.	

Table 4: Case study of generated sentences.

Method	Acc w/o DA	Acc w/ DA
IAN	78.99	79.90 (+0.91)
GCAE	81.67	81.98 (+0.31)
AOA_LSTM	79.90	80.25 (+0.35)
AT_LSTM	80.88	81.51 (+0.63)
ATAE_LSTM	80.75	80.90 (+0.15)
BiLSTM_Attn	78.87	80.56 (+1.69)
MemNet	77.24	78.52 (+1.28)

Table 5: Results of data augmentation experiments.

Since there is no adequate aspect-level sentiment classification data to train our model, we construct a large-scale dataset automatically with machine learning methods. First, we pre-train an aspect extraction model and two aspect-level sentiment classification models using the benchmark datasets from the SemEval ABSA challenge [13]. Then, we use these models to annotate the Citysearch New York Restautant Review Dataset [14] to construct experimental data.

Specifically, for aspect term extraction, we train a state-of-the-art aspect term extraction model, i.e., DECNN [15] to extract aspect terms in the input sentence. DECNN achieves the F1-score of 0.8832 on average. Then, two aspect-level sentiment classification models (GCAE [16] and IAN [3]) are trained on the SemEval corpus to classify the sentiment polarity (i.e., positive, negative or neutral) of the input sentence towards the given targets. We only keep the instances when GCAE and IAN have the same prediction. Since we transform the aspect-based sentiment polarity of the input sentence from positive (or negative) to negative (or positive), the instances with neutral sentiment are removed.

We use the auto-annotated data for training and manually annotate 5,000 samples for validation and testing, respectively. The statistics of the training, validation, testing data are reported in Table 1.

Baseline Methods

Since this paper is the first work to deal with ALSM, and we can not find previous work for the same task. Thus, we choose several sentence-level sentiment modification methods: CAAE [1], SMAE [2], ARAE [9], DAR [10] as baselines. For fair comparisons, we also implement two baselines directly for ALSM: a seq2seq model [17] with aspect and sentiment embedding as additional input (called Seq2Seq), and a retrieval model (called ReALSM) that retrieves the sentence with the same aspect and reversed sentiment from the training set.

Implementation Details

We use 300-dimension word vectors pre-trained by GloVe to initialize the word embeddings. Other parameters are initialized with Xavier uniform initializer. The dimension for position embedding is set to be 200. The annealing temperature τ is set to 0.001. The size of hidden states d is set to 600. We employ Adam optimizer to train our model with a learning rate η of 0.0001. The batch size is set to 64.

EXPERIMENTAL RESULTS

Automatic Evaluation

Following [2], we use BLEU to evaluate the content preservation degree. For the ALSM task, we additionally define two metrics: aspect-term preservation degree (denoted as APD) and aspect-level sentiment transformation accuracy (denoted as ASTA). Since the aspect term is the most important content in the ALSM task, we define a metric, i.e., aspect-term preservation degree (denoted as APD) to evaluate the performance of preserving aspect information. We compute APD as the word-level recall of the generated sentence against the original aspect term.

The ASTA metric calculates the ratio of the cases that the models successfully modify the aspect-level sentiment of the given sentences and aspects. Only when the APD score is larger than a threshold (we set the threshold to 0.3), the generated sentence with opposite sentiment towards the given aspect can be considered as a successful sentiment modification case. The experimental results are reported in Table 2. JASM significantly outperforms all baselines on all the three automatic evaluation metrics, which verifies the effectiveness of our model in ALSM.

Human Evaluation

We also use human annotations to evaluate the generated sentences from aspect sentiment transformation, content preservation, and fluency perspectives, similar to [2]. We randomly sample 300 instances from the test set and invite three human annotators to assign each generated sentence a score of 1 (bad), 2 (poor), 3 (not bad), 4 (satisfactory), 5 (good) for sentiment transformation (Trans.), content preservation (Cont.), and fluency, respectively.

The human evaluation results are summarized in Table 3. We observe that our model performs best in content preserving and aspect sentiment transformation.

Since ReALSM is a retrieval method, it has the best fluency. However, ReALSM cannot preserve the content of the input sentence.

Case Study

To evaluate the proposed model qualitatively, the generated sentences by all models given the original sentence from the test set are demonstrated in Table 4. From the results, we can observe that our model generates better sentences than other models, especially in content preserving.

Data Augmentation for Aspect-Based Sentiment Analysis

For application, we use pre-trained JASM model to modify the aspect-specific sentiment polarities of the training data in SemEval-14 Restaurant dataset [13] and generate plausible ABSA data. To verify the effectiveness of data augmentation with the samples generated by the JASM model, we evaluate the performance of a broad range of ABSA models on both the original data (denoted as w/o DA) and the augmented data (w/ DA). The baseline methods for ABSA include IAN [3], GCAE [16], AOA_LSTM [18], AT_LSTM [19], ATAE_LSTM [19], BiLSTM_attn and MemNet [20]. Table 5 shows the results of data augmentation experiments. We can observe that data augmentation using plausible data generated by JASM can boost the performance of a broad range of ABSA models, which shows the practical value of ALSM models.

CONCLUSION

This work proposed the aspect-level sentiment modification which is a fine-grained sentiment modification task. We designed an effective model that jointly trained an aspect-specific sentiment words extraction module and an aspect-level sentiment transformation module. The experimental results demonstrated that our model performed significantly better than the compared methods from aspect-level sentiment transformation and content preserving perspectives. In addition, the data augmentation experiments on various ABSB models showed the practical value of our model.

ACKNOWLEDGEMENT

This work was also partially supported by the Natural Science Foundation of Guangdong Province of China (No. 2018A030313943, No. 2019A1515011705), the SIAT Innovation Program for Excellent Young Researchers.

REFERENCES

1. T. Shen et al., “Style transfer from non-parallel text by cross-alignment,” *NIPS*, 2017, pp. 6830–6841.
2. Y. Zhang et al., “Learning sentiment memories for sentiment modification without parallel data,” *EMNLP*, 2018, pp. 1103–1108.
3. D. Ma et al., “Interactive attention networks for aspect-level sentiment classification,” *IJCAI*, 2017, pp. 4068–4074.
4. E. Cambria et al., “The CLSA Model: A Novel Framework for Concept-Level Sentiment Analysis,” LNCS, vol. 9042Springer, 2015, pp. 3–22.
5. E. Cambria et al., “Sentiment Analysis is a Big Suitcase,” *IEEE Intelligent Systems*, vol. 32, no. 6, 2017, pp. 74–80.

6. S. Poria, E. Cambria, and A. Gelbukh, “Aspect extraction for opinion mining with a deep convolutional neural network,” *Knowledge-Based Systems*, vol. 108, 2016, pp. 42–49.
7. X. Li and W. Lam, “Deep multi-task learning for aspect term extraction with memory interaction,” *EMNLP*, 2017, pp. 2886–2892.
8. Y. Ma, H. Peng, and E. Cambria, “Targeted aspect-based sentiment analysis via embedding commonsense knowledge into an attentive LSTM,” *AAAI*, 2018, pp. 5876–5883.
9. J. Zhao et al., “Adversarially regularized autoencoders,” *ICML*, 2018, pp. 5897–5906.
10. J. Li et al., “Delete, retrieve, generate: A simple approach to sentiment and style transfer,” *NAACL-HLT*, 2018, pp. 1865–1874.
11. S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Computation*, vol. 9, 1997, pp. 1735–1780.
12. Z. Fan et al., “Target-oriented opinion words extraction with target-fused neural sequence labeling,” *NAACL-HLT*, 2019, pp. 2509–2518.
13. M. Pontiki et al., “SemEval-2014 Task 4: Aspect Based Sentiment Analysis,” *COLING*, 2014, pp. 27–35.
14. G. Ganu, N. Elhadad, and A. Marian, “Beyond the stars: improving rating predictions using review text content,” *WebDB*, 2009, pp. 1–6.
15. H. Xu et al., “Double Embeddings and CNN-based Sequence Labeling for Aspect Extraction,” *ACL*, 2018, pp. 592–598.
16. W. Xue and T. Li, “Aspect based sentiment analysis with gated convolutional networks,” *ACL*, 2018, pp. 2514–2523.
17. I. Sutskever, O. Vinyals, and Q. V. Le, “Sequence to sequence learning with neural networks,” *NIPS*, 2014, pp. 3104–3112.
18. B. Huang, Y. Ou, and K. M. Carley, “Aspect level sentiment classification with attention-over-attention neural networks,” *SBP-BRiMS*, 2018, pp. 197–206.
19. Y. Wang et al., “Attention-based LSTM for aspect-level sentiment classification,” *EMNLP*, 2016, pp. 606–615.
20. D. Tang, B. Qin, and T. Liu, “Aspect level sentiment classification with deep memory network,” *EMNLP*, 2016, pp. 214–224.

Qingnan Jiang is currently a research assistant at Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. He received the B.S. degree in computer science from China University of Mining and Technology in 2019. His research interests include natural language processing and information retrieval. Contact him at qn.jiang@siat.ac.cn.

Lei Chen is currently a master student at Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. He received the B.S. degree in computer science from China University of Mining and Technology in 2019. His research interests include natural language processing and information retrieval. Contact him at lei.chen@siat.ac.cn.

Wei Zhao is currently a second-year PhD student at Technische Universität Darmstadt, Germany. Prior to that, He received master degree in Computer Science at University of Chinese Academy of Sciences in 2018. He was a visiting student at Nanyang Technological University in 2018, Singapore. His research interests include Natural Language Processing and Computational Linguistics. Contact him at zhao@aiphes.tu-darmstadt.de.

Min Yang is the corresponding author and an associate professor at Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. She received her Ph.D. degree from the University of Hong Kong in 2017. Her current research interests include machine learning and natural language processing. Contact her at min.yang@siat.ac.cn.