Dependency Parsing and bidirectional LSTM-CRF for Aspect-level Sentiment Analysis of Chinese

Huichao Xiong[1], Hua Yan[1], Zhixian Zeng[1] and Binhui Wang[1*]

[1]Nankai University, Tianjin, China
{2120160362,1511466, 2120170549}@mail.nankai.edu.cn
wangbh@nankai.edu.cn

Abstract. Aspect-level sentiment analysis of Chinese is to extract, aggregate a21 apply fine-grained aspect-level sentiment information from text for sentiment understanding, and it is useful in various application domains. It usually first extracts aspect terms and sentiment words simultaneously, then pairs the aspect terms with sentiment words, and lastly classifies the aspect-level sentiment. In this sense, we formulate this problem as a pipeline of aspect terms and sentiment words extraction through sequence labeling, aspect-sentiment word pairing, and aspect-level sentiment classification. In this paper, we use a bidirectional LSTM-CRF model to extract aspect terms and sentiment words, some syntax rules based on dependency parsing to pair aspect and sentiment words, and mainstream classifiers to determine the sentiment polarities.

Keywords: Aspect-level sentiment analysis, dependency parsing, bidirectional LSTM-CRF.

1 Introduction

Aspect-level sentiment analysis (ABSA) has been used in various application domains, including online marketing, corporate public opinion monitoring, and government opinion survey. Previous work used topic models for ABSA, while these methods learned topics are overly abstract and their work focuses on document-level sentiment classification [1, 2]. In addition, most research focused on the English language, only few work aims at ABSA of Chinese [3]. To understand sentiments at the level of aspects, there are two fundamental tasks, i.e., aspect term extraction and aspect-based sentiment classification. Since it is difficult to combine them in one step, in this paper we first extract and pair the aspect terms with sentiment words, and then classify for the aspect-level sentiment.

We first extract aspect terms and sentiment words using a BI-LSTM-CRF (Bidirectional Long-Short Term Memory Conditional Random Field) model that utilizes both the contextual information in bidirectional pathways and the sentence-level tags through a CRF layer. We then pair the aspects with sentiments using some syntax rules. Lastly, we use classifiers to classify sentiment polarity of aspect terms.

*Corresponding author

2 Approach

Our approach includes three parts: aspect/sentiment extraction by BI-LSTM-CRF, pairing aspect and sentiment words by syntax rules and gaining sentiment polarity of aspects through classifiers.

2.1 BI-LSTM-CRF

BI-LSTM-CRF model includes bidirectional LSTM and CRF (shown in Figure 1). It uses both the preceding input features and the future input features.

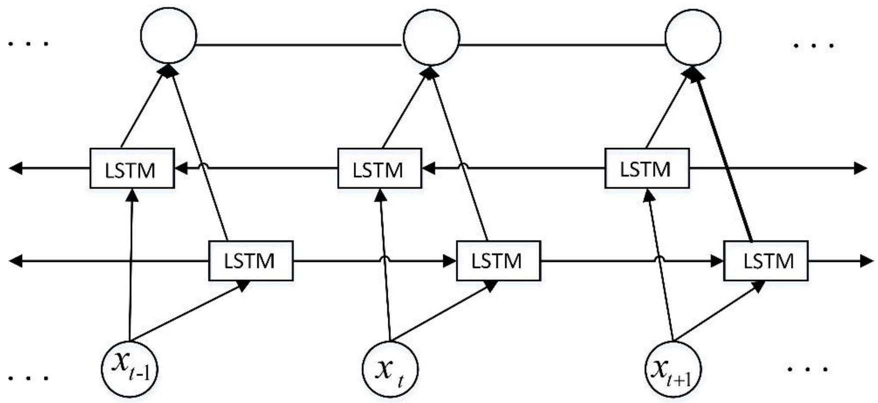

Fig. 1. A BI-LSTM-CRF Model [4]

2.2 Syntax Rules

After aspect terms and sentiment words are extracted from reviews, some syntax rules can help find the relation between the aspect terms and their corresponding sentiment words, to generate the aspect-sentiment pairs. Considering there may have some sentiment words without any toward aspect terms, but we need get these pairs if the sentiment words exist. So, we should first detect aspects and sentiment words then gain the dependency structure shows which words depend on (modify or are arguments of) which other words. In order to get the relation between aspects and sentiment words and find the actual aspect terms lastly, we use the Chinese NLP tool jieba (https://github.com/fxsjy/jieba) for word segmentation and LTP (https://github.com/HIT-SCIR/ltp) for dependency parsing. As illustrated in Table 1, their relations usually are ATT (attribute), SBV (subject-verb) and IOB (indirect-object). As the aspect usually appear in front of sentiment, their relation of position in a sentence can be used.

Table 1. Examples of Relation between Aspects and Sentiment Words

Sentences	Aspects	Sentiment words	Relation
这是一款漂亮的衬衫。 (This is a beautiful shirt.)	衬衫 (shirt)	漂亮 (beautiful)	ATT
iphone销量提升。 (The sales of iphone rises.)	销量 (sales)	提升 (rise)	SBV
抵制苹果品牌来支持Nokia。 (Resist the selling of iphone to support Nokia.)	销售 (selling)	抵制 (resist)	IOB

2.3 Classifiers

We form a lexicon from training set and use it to prepare the features for both training and testing sets. Due to the known "no free lunch theorem" in supervised learning, we use traditional classifiers, including Naïve Bayes (NB), Logistic Regression (LR) and Support Vector Machines (SVM) to determine the polarity of sentiment words that doesn't appear in the lexicon.

3 Experiments

To evaluate the performances of the models on the two sub-tasks of ABSA: 1) aspect term extraction 2) aspect-based sentiment classification. The metrics for the two sub-tasks are precision (P), recall (R), and F1-score (F1). We compare the performance across multiple models in Table2. We find that the models with CRF components perform better than without, and that bidirectional LSTM works better than LSTM in one direction.

Table 2. Comparison of ABSA Performance for Various Models

Models	Aspect Extraction			Sentiment Classification		
	P	R	F1	P	R	F1
LSTM+LR	46.36	52.75	49.34	59.35	65.74	62.38
LSTM+NB				58.48	65.40	61.75
LSTM+SVM				48.60	61.10	54.14
BI-LSTM+LR	76.87	77.10	76.99	71.08	83.95	76.98
BI-LSTM+NB				67.05	83.15	74.24
BI-LSTM+SVM				51.91	79.25	62.73
CRF+LR	83.20	69.74	75.88	69.79	**83.99**	76.23
CRF+NB				65.55	83.13	73.30
CRF+SVM				54.42	80.36	64.89
LSTM-CRF+LR				64.87	80.78	71.95

LSTM-CRF+NB	82.16	80.01	81.07	61.51	79.94	69.53
LSTM-CRF+SVM				53.66	77.66	63.47
BI-LSTM-CRF+LR	**83.55**	**80.24**	**81.86**	**71.64**	83.68	**77.19**
BI-LSTM-CRF+NB				67.85	82.92	74.63
BI-LSTM-CRF+SVM				56.95	80.30	66.64

4 Conclusion

This paper reviews and applies the bidirectional LSTM-CRF model in the product reviews in Chinese and systematically compares the performances across multiple state-of-the-art models that cast aspect term and sentiment word extraction to sequence labeling. Then, we use mainstream classifiers to gain the corresponding sentiment polarity of aspect terms. Experimental results show that the BI-LSTM-CRF with LR outperforms other counterparts. As one future work, we will try to incorporate convolutional neural networks into the BI-LSTM model with some attention mechanism [5] for CRF input to discover the implicit aspects.

Acknowledgments. This work is supported by Tianjin Social Science Planning Topics TJXC15-002.

References

1. Lin C., He Y.: Joint sentiment/topic model for sentiment analysis. Proceedings of the 18th ACM conference on Information and knowledge management. 375-384(2009).
2. Titov I., McDonald R.: Modeling online reviews with multi-grain topic models. Proceedings of the 17th international conference on World Wide Web. 111-120(2008).
3. Lipenkova, J.: A system for fine-grained aspect-based sentiment analysis of Chinese. Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics. 55–60(2015).
4. Huang Z., Xu W., Yu K.: Bidirectional LSTM-CRF models for sequence tagging. arXiv preprint arXiv:1508.01991(2015).
5. Bahdanau D., Cho K., Bengio Y.: Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473(2014).