

Aspect Based Sentiment Analysis for Online Reviews

Lamei Xu¹, Jin Liu¹, Lina Wang¹, and Chunyong Yin²(✉)

¹ College of Information Engineering, Shanghai Maritime University,
Shanghai, China

{lmxu, jinliu, nwang}@shmtu.com

² School of Computer and Software, Nanjing University of Information Science
and Technology, Nanjing 210044, China

yhzhen@nuist.edu.cn

Abstract. Learning good semantic vector representations for sentiment analysis in phrases, sentences and paragraphs is a challenging and ongoing area of natural language processing. In this paper, we propose a Convolution Neural Network for aspect level sentiment classification. Our model first builds a convolution neural network model to aspect extraction. Afterwards, we used a sequence labeling approach with Conditional Random Fields for the opinion target detection. Finally, we concatenate an aspect vector with every word embedding and apply a convolution neural network over it to determine the sentiment towards an aspect. Results of an experiment show that our method performs comparably well on Yelp reviews.

Keywords: Convolution neural network · Conditional random fields
Word2-ec · Aspect-based sentiment analysis

1 Introduction

Sentiment analysis aims to analysis people's sentiment or opinions according to their generated texts and plays a critical role in the area of data mining and natural language processing. Sentiment analysis has drawn a lot of attentions from both industry and academic communities due to continuous growing of user-generated content on the Internet. [1].

Aspect-based sentiment analysis refers to the identification of specific entities and their aspects (aspect terms, opinion targets) in text and to the classification of their polarity [2]. In aspect-based sentiment analysis, we assume that the general target of evaluation has several aspects and we attempt to identify user's opinions on these individual aspects. Unlike the more general task of sentiment analysis where the goal would be to classify the polarity of entire sentence, aspect level analysis performs a finer-grained sentiment analysis by addressing three sub-problems: extracting aspects from the reviews text, identifying the entity that is referred to by the aspect, and finally classifying the opinion polarity towards the aspect.

For the present challenge, we develop convolution neural network for aspect level sentiment classification, the remainder this paper is structured as follows. In Sect. 2

describes recent approaches to the aspect-based polarity analysis, Sect. 3 contains our new method to detect aspect and aspect-based sentiment analysis, Sect. 4 provides our preliminary test results.

2 Related Work

Aspect-based sentiment analysis has been a subject of some interesting works so far. The simple approach is to calculate a sentiment score of a give aspect as the weighted sum of opinion scores, which are defined by a sentiment lexicon, of all words in the sentence [3]. This method is further improved by identifying the aspect-opinion relations using tree kernel method [4].

Recently, deep learning-based approaches have demonstrated remarkable results for text classification and sentiment analysis [5]. Recursive Neural Network is a kind of deep neural network. Using distributed representations of words (aka word embedding) [6], RNN merges word representations to represent phrases or sentence. It is one of the best methods to predict sentiment labels for phrases [7].

Previously, the system by Kiritchenko [8] used various innovative linguistic features, publicly available sentiment lexicon corpora and automatically generated polarity lexicons, achieve the best performance in polarity classification. Tang [9] rely on use a target-dependent LSTM to determine sentiment towards a target word, while Nguyen and Shirai [4] used wide range of features such as Big-of-Words, negation words, bigram after negation, polarity inversion, polarized terms in last 5 tokens, publicly available lexicons etc. They used MALLET [10, 12] with Maximum Entropy classifier.

3 Aspect-based Sentiment analysis Model

We present the different components of our system, dedicated to linguistic feature ex-traction, sequence labeling classification.

3.1 Deep Convolution Neural Network Model

The model architecture we use is an extension of the CNN-rand structure used Kim. The model takes as input a sentence, and we represent the sentence as a concatenation of its word embedding which can be described using Eq. (1).

$$x_{1:n} = x_1 \oplus x_2 \oplus \dots \oplus x_n \quad (1)$$

Where that $x_i \in R^k$ is a k dimensional word vector for the i^{th} word in sentence which have n words. \oplus is the concatenation operator. Generally, a convolution layer filter with weights $w \in R^{hk}$ of h words and generates a new feature c_i .

$$c_i = f(w \cdot x_{x:i+h-1} + b). \quad (2)$$

Where $b \in R$ is a bias term and f is a non-linear function, ReLU [13]. This filter is applied to each possible window of h words in the sentence to generate a feature map.

$$c = [c_1, c_2, \dots, c_{n-h+1}] \quad (3)$$

We use a max-pooling operation [11] over the feature map and take the maximum value $\hat{c} = \max\{c\}$ as the corresponding to this filter.

A softmax layer takes the concatenation of the maximum values of the feature maps produced by all filters and computes probability distribution over the possible categories.

Please check that the lines in line drawings are not interrupted and have a constant width. Grids and details within the figures must be clearly legible and may not be

3.2 Hyper-parameters

We use the following hyper-parameters, which are similar to Kim [3]: the filter lengths of 3, 4 and 5, and 4, 5 and 6 for aspect extraction and sentiment analysis, word embedding size is 200, mini-batch size of 10, maximum sentence length of 100 tokens, set the dropout rate to 0.5 and the l_2 maximum value to 3.

Word embedding is initialized with 300-dimensional word2vec [12] trained on 100 billion words from Google News. The vectors have dimensionality of 300 and were trained using the continuous bag-of-words architecture. Words not present in the set of pre-trained words are initialized randomly.

3.3 Aspect Category Detection

For this task, the goal is to identify entity and attribute pairs expressed by given review sentence. To choose the threshold, define an aspect a 's probability p given a sentence s as $p(a/s) = 1/n$ if a appears in s and a contains $p(a/s) = 0$. We define a threshold f and choose all aspect with $p(a/s) > f$, after training, we found out that a threshold value of 5 performed best. Thus, we take that all aspects occur less than 5 times to *other* aspect.

3.4 Opinion Target Expression

For this task, we are asked to extract the exact expressions or words in the sentence, in which an opinion is expressed. We developed a system based on CRFs, using CRF++ tool [14, 15] and the training file provided for building the model. A training file is needed to build as input for the CRF, whose structure is as follows. In the first column, all the words for every sentence are written, then in the second column, the corresponding lemma. The third column represents the tag and the last one represents if the word is an aspect or not or if it is included in a multiword aspect. Then for creating the model we take into account all these features, as well as all the possible bigrams in each sentence. In the output, if no target is found, no opinion is returned for that sentence.

3.5 Sentiment Analysis

Opinion has to be classified with the two following polarities: positive and negative. We applied a similar strategy as for the aspect categories classification, i. e. We use the same pipeline as previously but in this case, we associate the highest polarity probability to the term or sentence, ignoring the few cases presenting a mixed polarity (i.e. both positive and negative). Features are extracted the same way, but we add the aspect category detected previously as a feature for polarity classification.

We embed the tokens of all aspects in the same embedding space as word tokens to find the semantics of embedding. Then, look up to the embedding of every token and average them to retrieve the aspect vector [16]. For this, the model can learn aspect sharing the same entity. Finally, the results aspects vector is concatenated with each word vector, then aspect vector together with word vector to produce a sentence matrix. Max-pooling and softmax layer in convolutional neural network are applied to this matrix.

4 Experiments and Results

We describe experimental settings and report empirical results in this section. We conduct experiments on two datasets from Yelp Dataset [17], one from laptop domain and another from restaurant domain. Statistics of the datasets are given in Table 1. Evaluation metric is classification accuracy.

Table 1. Statistics of the dataset.

Dataset	Positive sentences	Negative sentences
Restaurant-Train	870	807
Restaurant-Test	728	196
Laptop-Train	994	870
Laptop-Test	341	128

Experimental results are given in Table 2 and Fig. 1. We can find that feature-based SVM is an extremely strong performer and substantially outperforms other baseline methods, which demonstrates the importance of powerful representation for aspect level sentiment classification. One of the major reasons why the polarity method did not perform better is that we adapted a method that was designed for identifying two categories (positive and negative), which has three categories (positive, negative and neutral).

Table 2. Model performance on training dataset.

Model	Restaurant	Laptop
Feature +SVM	72.10	80.89
LSTM	66.45	74.28
RNN	63.00	64.56
Our model	68.34	70.90

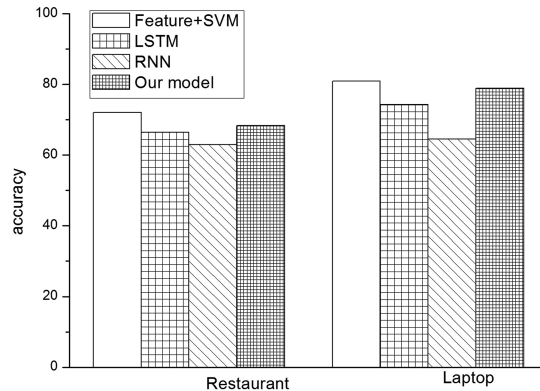


Fig. 1. Accuracy of model on test set.

5 Conclusions

In this paper, we reported our work on the task of aspect-based sentiment analysis, which covers three subtasks: aspect identification, opinion target extraction and sentiment polarity classification. We have presented a deep learning-based approach to aspect-based sentiment analysis, which employs a convolutional neural network for aspect extraction and sentiment analysis and CRF for opinion target expression. To advance our model, further work is needed on better identification of neutral cases. We will also explore how our CNN systems can be further enhanced.

References

1. Liu, B.: Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers, San Rafael (2012)
2. Hu, M., Liu, B.: Mining and summarizing customer reviews. In: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2004, pp. 168–177. ACM, New York (2004)
3. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Found. Trends Inf. Retrieval. **2**(1–2), 1–135 (2008)
4. Nguyen, T.H., Shirai, K.: PhraseRNN: phrase recursive neural network for aspect-based sentiment analysis. In: Proceedings of the 2015 Conference on Empirical Method Shirai (2015)
5. Kim, Y.: Convolutional neural networks for sentence classification. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014), pp. 1746–1751. <http://emnlp2014.org/papers/pdf/EMNLP2014181.pdf>.
6. Hinton, G.E.: Learning distributed representations of concepts. In: Proceedings of the Eighth Annual Conference of the Cognitive Science Society, pp. 1–12. Erlbaum, Hillsdale, NJ (1986)
7. Socher, R., Chen, D., Manning, C.D., Ng, A.Y.: Reasoning with neural tensor networks for knowledge base completion. In: Proceedings of the Advances in Neural Information Processing Systems, NIPS 2013, vol. 26, pp. 1–10 (2013)

8. Kalchbrenner, N., Grefenstette, E., Blunsom, P.: A convolutional neural network for modelling sentences. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, 655–665 (2014)
9. Tang, D., Qin, B., Feng, X., Liu, T.: Target-Dependent Sentiment Classification with Long Short Term Memory. arXiv preprint [arXiv:1512.01100](https://arxiv.org/abs/1512.01100) (2015)
10. Machine Learning for Language Toolkit
11. Mikolov, T., Sutskever, I., Chen, K., Corrado, G., Dean, J.: Distributed Representations of Words and Phrases and their Compositionality (2013)
12. Socher, R., Perelygin, A., Wu, J.Y., Chuang, J.: Recursive deep models for semantic compositionality over a sentiment treebank. In: EMNLP (2013)
13. Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R.: Improving neural networks by preventing coadaptation of feature detectors. arXiv: 1207.0580, pp. 1–18 (2012). Nair and Hinton, 2010. <http://doi.org/arXiv:1207.0580>
14. Kudo, T.: CRF++: Yet another CRF toolkit. Software (2005). <http://crfpp.sourceforge.net>
15. http://www.yelp.com/dataset_challenge
16. Go, A., Bhayani, R., Huang, L.: Twitter sentiment classification using distant supervision. CS224 N Project Report, Stanford, pp. 1–12 (2009)
17. Zhang, Min-ling, Zhou, Zhi-hua, Member, Senior: Multilabel neural networks with applications to functional genomics and text categorization. IEEE Trans. Knowl. Data Eng. **18**(10), 1338–1351 (2006)