Sentence-level Sentiment Classification with Weak Supervision

Fangzhao Wu[1], Jia Zhang[1], Zhigang Yuan[1], Sixing Wu[1], Yongfeng Huang[1], Jun Yan[2]

[1]Tsinghua University, Beijing, China

[2]Microsoft Research Asia, Beijing, China

Presented in Conference Proceedings of SIGIR 2017

Outline of problem

- Sentence-level sentiment classification is important to understand users' fine-grained opinions.

- Existing methods for sentence-level sentiment classification are mainly based on supervised learning. However, it is difficult to obtain sentiment labels of sentences since manual annotation is expensive and time-consuming.

- Although a document may contain sentences with different sentiments, most of the sentences in an opinionated document usually have the same sentiment with this document.

Proposed solution

- In this paper, authors propose an approach for sentence-level sentiment classification without the need of sentence labels. More specifically, they propose a unified framework to incorporate two types of weak supervision, that is, document-level and word-level sentiment labels, to learn the sentence-level sentiment classifier.

- In addition, the contextual information of sentences and words extracted from unlabeled sentences is incorporated into their approach to enhance the learning of sentiment classifier.

- Since words are the basic elements to express sentiments in a sentence, the sentiment labels of words may also be useful to train sentence-level sentiment classifiers. Besides, although it is difficult to obtain the sentiment labels of sentences, the sentiment relations between sentences are relatively easy to infer in many cases.

Some related approaches

- The first category is training sentence-level sentiment classier only based on labeled documents without the need of labeled sentences.

- The second category is combining coarse-grained document labels with fine-grained sentence labels for sentence-level sentiment classification.

- The contexts of sentences have also been explored in several existing sentence-level sentiment classification methods.
  - However, adjacent sentences may have different even opposite sentiments.
  - The sentence similarity is measured by word sequence closeness.

Solution approach

- In many cases the sentiment relations between them are much easier to infer. Following this basis, the paper explores to extract the sentiment relations between sentences based on coordinating and adversative conjunctions.

- First, if two words have the same parts-of-speech tag, and they are connected by coordinating conjunction “and” or used to describe the same target in the same sentence, then we regard they convey the same sentiment. Second, if two words are connected by adversative conjunction “but” and have the same parts-of-speech tag, then they are assumed to have opposite sentiments.

Some examples -

- “It cleaned quickly and required no seasoning. Also, it’s a really pretty skillet.” Since these two sentences are connected by the coordinating conjunction “also”, they probably convey the same sentiment.
- “These dishes look very nice on your table, but they have many problems.” Since the two sentences are connected by the adversative conjunction “but”, we can infer that they may have opposite sentiments.

Solution approach

- The goal of proposed approach is to incorporate the document-level supervision, the word-level supervision, and the contextual information of sentences and words to train an accurate sentence-level sentiment classifier.
- Instead of directly using sentiment labels of documents for sentences, the solution constrains the average sentiment score of the sentences in a document which is consistent with the label of this document.
- If two unlabeled sentences have same-sentiment (or opposite-sentiment) relation, then the solution constrains the sentiment classifier and assigns the same (or opposite) sentiment label to them.

Datasets and Experiment Results

- In experiments the sentiment dataset built by Täckström and McDonald [1] was used. The sentiment labels of the sentences in this dataset were manually annotated. Three domains were involved in our experiments, i.e., Book, DVD, and Electronics.

- In addition, authors used the Amazon sentiment dataset crawled by Blitzer et al. [2] to obtain labeled documents in these domains. The sentiment labels of documents were automatically inferred from their ratings. The detailed statistics of these datasets are illustrated in Table 1.

- The word-level supervision was extracted from Bing Liu’s sentiment lexicon. The authors used half of the documents as labeled documents to provide document-level supervision, and used the others to extract the contextual sentiment relations between sentences and the sentiment similarities between words.

Table 1: The statistics of the datasets.

	Labeled Sentences			*#Document*
	#Positive	*#Negative*	*#Total*	
Book	160	165	355	975,194
DVD	164	264	428	124,438
Electronics	161	240	401	23,009

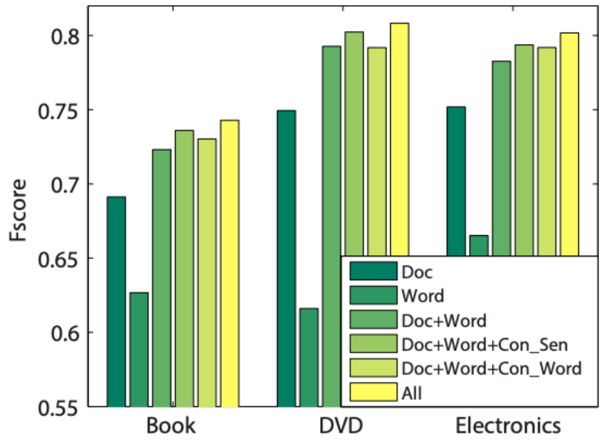

Figure 1: Performance of our approach with different combinations of sentiment information.

Doc and **Word** mean document-level and word-level supervision respectively. **Con_Sen** and **Con_Word** represent contextual information of sentences and words respectively. **All** means all information is incorporated.

Performance Evaluation

- In this section authors evaluate the performance of their approach by comparing it with several baseline methods, including:
- 1) SVM, LR, and LS, i.e., support vector machine, logistic regression, and least square method, which are trained on labeled sentences;
- 2) CNN, convolutional neural network for sentence classification;
- 3) ParaVec, the paragraph vector method;
- 4) HCRF, the hidden conditional random fields method;
- 5) SSLVM, the semi-supervised latent variable model;
- 6) MEM, the weakly supervised multiexperts model;
- 7) WDE, the weakly supervised deep embedding method;
- 8) PR, the sentence sentiment classification method with posterior regularization;
- 9) SSWS, our sentence-level sentiment classification method with weak supervision.

- For baseline methods which need fine-grained labels in model learning, authors used half of the labeled sentences for training and the others for test. Experimental results are shown in Table 2. The performance metric is macro-averaged Fscore.

Table 2: The performance of different methods.

	Book	*DVD*	*Electronics*
SVM	0.6580	0.7071	0.6717
LR	0.6694	0.7218	0.6684
LS	0.6560	0.7086	0.6668
CNN	0.6885	0.7689	0.6753
ParaVec	0.6204	0.7508	0.6585
HCRF	0.7021	0.7566	0.7615
SSLVM	0.7142	0.7821	0.7906
MEM	0.7207	0.7846	0.7865
WDE	0.7099	0.7629	0.7726
PR	0.7255	0.7931	0.7859
SSWS	0.7428	0.8082	0.8017

The paper's approach performs better than posterior regularization(PR) and semi-supervised latent variable model(SSLVM), Because it can exploit not only document-level but also word-level supervision for learning sentence-level sentiment classifiers.

Conclusion

- The solution in paper does not rely on the fine-grained sentiment labels of sentences, which are difficult to obtain. Instead, it can exploit both document-level and word-level supervision to learn sentence-level sentiment classifiers.

- In addition, the contextual information of sentences and words is incorporated into the solution to enhance the learning of sentiment classifier.

- Experiments on benchmark datasets show that the solution proposed in paper can effectively improve the performance of sentence-level sentiment classification.

References -

- [1] Oscar Täckström and Ryan T. McDonald. 2011. Discovering Fine-Grained Sentiment with Latent Variable Structured Prediction Models. In ECIR. 368–374

- [2] John Blitzer, Mark Dredze, Fernando Pereira, and others. 2007. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classi�cation. In ACL. 440–447.

- Fangzhao Wu, Jie Zhang, Zhigang Yuan, Sixing Wu, Yongfeng Huang, Jun Yan Proceeding SIGIR '17 Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval Pages 973-976