A Study on Subjectivity-oriented Polarity Classification

Na Dong Kiyoaki Shirai
Japan Advanced Institute of Science and Technology
{s2010417,kshirai}@jaist.ac.jp

Abstract

This paper presents two methods of document-level polarity classification that consider the subjectivity of sentences, which is based on the idea that subjective sentences in a document play a more important role than objective sentences. The first method determines the polarity of a document by weighted voting of the polarity of sentences in it, where the weights are defined as the intensity of the subjectivity. The second method utilizes pre-trained language models such as BERT and XLNet after objective sentences are filtered out. The effectiveness of these two methods is empirically evaluated on two datasets.

1 Introduction

Polarity classification, which is a kind of sentiment analysis (SA), is a task to classify a given text into polarity, i.e. to judge whether a text expresses positive or negative opinions [1]. A text for polarity classification can be a document, sentence, or aspect, but this study focuses on document-level polarity classification. A typical example is the classification of the polarity of user reviews.

In general, there are two kinds of sentences in user reviews. One is a subjective sentence that expresses the writer's emotion or opinion, the other is an objective sentence that refers to objective facts. Intuitively, subjective sentences play a more important role than objective sentences in polarity classification. It would be helpful to filter out objective sentences or put more priority on subjective sentences in order to improve the performance of polarity classification. Subjectivity is also considered in past studies of SA, such as a subjectivity classification task, but it is not paid much attention in polarity classification.

The goal of this study is to propose two methods of document-level polarity classification that heavily considers the subjectivity of the sentences. One is an approach to determine the polarity of an overall document by voting of the polarity of sentences in it where the polarity of the subjective sentences are highly weighted. The other is an approach to use pre-trained language models such as BERT[2] and XLNet[3] with filtering of the objective sentences.

2 Related work

The subjectivity of the sentences has been considered in polarity classification in a few previous studies. Pang and Lee proposed a method to extract only subjective sentences from a document by using a minimum cut framework, then classify the polarity of the extracted document by naive Bayes model [4]. Their results of experiments demonstrated that the document consisting of only subjective sentences was not only shorter but also more effective than the original document, namely the accuracy of the polarity classification was significantly improved. Sindhu et al. proposed a similar method consisting of subjectivity classification and polarity classification [5]. The subjectivity classification was performed to filter out objective sentences, then the polarity classification was carried out using only subjective sentences. The subjectivity was also considered in a wide variety of sentiment analysis, e.g. extraction of aspect and opinion words. Kamal proposed a method to extract pairs of features (aspects) and opinions by combining supervised machine learning and rule-based approaches [6]. The subjectivity of sentences in texts is classified first, then feature-opinion pairs are mined from only the subjective sentences by a rule-based method.

This study shared the basic idea with Pang's and Sindhu's methods: subjective sentences are more important in polarity classification. Although the objective sentences are just ignored in their methods, we suppose that the objective sentences have less but also useful information for polarity classification. Therefore, this study investigates the way to use both subjective and objective sentences with priority on the former.

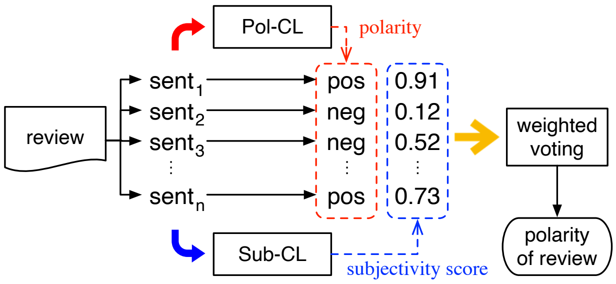

Figure 1 Overview of subjectivity weighted voting

Recently, BERT has been applied for many tasks of natural language processing (NLP) and often achieved state-of-the-art results. Inspired by the success of pre-trained language models, Pota et al. proposed a two-step approach for polarity classification of tweets posted on Twitter [7]. First, tweets including special tokens of Twitter were converted into plain texts using a language-independent or easily adaptable process for different languages. Second, the polarity of the converted tweets was classified using the language model BERT. This paper also explores how to use pre-trained language models. Furthermore, we investigate how subjective and objective sentences can be handled to improve the performance of polarity classification by language models.

3 Proposed method

This section explains the details of two proposed subjectivity-oriented polarity classification methods.

3.1 Polarity Classification by Subjectivity Weighted Voting

Our first method classifies the polarity of a given document by voting of the polarity of each sentence where the subjective sentences are more heavily considered. The overview of this method is shown in Figure1. It is implemented by the following steps.

1. Data preprocessing. Tokenization, lemmatization, and removing stop words are performed as preprocessing.
2. Polarity classification of each sentence. For each sentence in a review, the polarity of it is classified by the supervised machine learning model.
3. Subjectivity classification of each sentence. For each sentence, a subjectivity score, i.e. intensity of the subjectivity, is calculated by the supervised machine learning model.
4. Voting. The polarity of the overall review is determined by the weighted voting of the polarity classes of the sentences, where the weight is defined as the subjectivity score of each sentence.

Since our goal is document-level polarity classification, we suppose that a collection of documents (reviews) annotated with their polarity is available as the training data. In the step 2, to train the sentence-level polarity classifier, we automatically assign the same polarity label for each sentence as that of the review. That is, all sentences in a positive (or negative) review are classified as positive (or negative) sentences. Then the sentence-level polarity classifier ("Pol-CL" in Figure 1) is trained from this dataset.

In the step 3, we suppose that a dataset of sentences annotated with the subjectivity labels ("subjective" or "objective") is available. The subjectivity classifier ("Sub-CL" in Figure 1) is trained from the subjectivity dataset. Then, the probability of the classification predicted by the classifier is used as the subjectivity score of the sentence.

An illustrative example of this method is shown in Figure 2. The example review, whose gold label is "positive", consists of five sentences. The simple voting wrongly classifies it as "negative", since the number of negative sentences is more than that of the positive sentences. On the other hand, the proposed method classifies it as "positive", since the sum of the subjectivity score for positive sentences is greater than that of negative sentences.

review	polarity	score
① This a fantastic movie about three prisoners who become famous.	positive	0.647
② One of the actors is George Clooney and I'm not a fan but this role is not bad.	negative	0.265
③ Another good thing about the movie is the soundtrack (The man of constant sorrow).	negative	0.454
④ I recommend this movie to everybody.	positive	0.782
⑤ Greetings Bart.	negative	0.544

Simple voting:
$Count_{pos} = 2 < Count_{neg} = 3 \Rightarrow$ negative

Proposed method:
$Score_{pos} = 1.429 > Score_{neg} = 1.263 \Rightarrow$ positive

Figure 2 Example of polarity classification

3.2 Polarity Classification by Pre-trained Language Model with Subjectivity Filtering

Our second method mainly relies on the pre-trained language model. We use two common language models: BERT and XLNet. In addition, we incorporate a filtering mechanism to use only subjective sentences for polarity classification. The document-level polarity classifier is trained by the following steps.

1. Data preprocessing. It is the same as the one in Subsection 3.1.
2. Subjectivity filtering. Each sentence is classified whether it is subjective or objective. Then the objective sentences are discarded and only subjective sentences are retained. We call a set of remaining subjective sentences as "pseudo review".
3. Fine-tuning. Using the pseudo reviews as the training data, we fine-tune BERT or XLNet model. Although the pseudo review consists of several sentences, we treat it as a single sentence.

In the test phase, an input pseudo review is obtained by the same preprocessing and filtering, or the original review is just used as an input. Then, the polarity is classified by the fine-tuned BERT or XLNet.

4 Evaluation

This section reports several experiments for the evaluation of our proposed methods.

4.1 Dataset

The following two datasets are used for the experiments.

IMDB Review Dataset IMDB Review Dataset [8] is a collection of 50,000 movie reviews. Each review is annotated with binary polarity labels: positive or negative. In this experiment, 25K reviews are used as the training data, while the rest of 25K reviews are used as the test data.

Amazon review dataset Amazon review dataset [9] is a collection of user reviews posted to the EC website Amazon. There are 35 million reviews for 18 years, up to March 2013. Each data includes a product name, user information, a rating, and a user review as plain text. The reviews with rating 4 or 5 are used as positive reviews, while 1 or 2 as negative reviews. We use the same number of reviews as the IMDB dataset, i.e. we randomly choose 25K reviews as the training data and another 25K reviews as the test data.

Besides, the following dataset is used to train the subjectivity classifier.

Subjectivity datasets Subjectivity datasets [10] include 5,000 subjective sentences excerpted from the movie review website and 5,000 objective sentences excerpted from IMDB plot summaries. Each line in the file of this dataset corresponds to a single sentence or snippet, and all sentences or snippets are downcased. Only sentences or snippets containing at least 10 tokens were included. The subjectivity labels of the sentences and snippets were automatically assigned.

Note that the subjectivity classifier trained from this dataset is used for the polarity classification in the aforementioned two polarity datasets. The domain of the IMDB dataset is the same as the subjectivity dataset (i.e. the movie review), but that of the Amazon dataset is different. It may degrade the classification performance on the Amazon dataset.

4.2 Result

4.2.1 Result of Subjectivity Classification

First, we analyze the performance of subjectivity classification. Three classifiers are compared: Support Vector Machine (SVM) using bag-of-words features, BERT and XLNet. Table 1 shows the accuracy of these classifiers[1]. It is found that XLNet performs the best and its accuracy is 0.96. In the rest of the experiments, this XLNet model is used as the subjectivity classifier.

Table 1 Accuracy of subjectivity classification

Model	SVM	BERT	XLNet
Accuracy	0.76	0.94	**0.96**

4.2.2 Results of Subjectivity Weighted Voting

We evaluate the method using the subjectivity weighted voting explained in Subsection 3.1. Our method "Weighted Voting" is compared with two baselines: "Sub. Only Voting" where the objective sentences are removed and the polarity is determined by voting of the polarity of the subjec-

1) The subjectivity dataset is randomly divided into 50% training data and 50% test data.

tive sentences[2]), and "Simple Voting" where the polarity is determined by voting without considering the subjectivity scores. Table 2 reveals the accuracy on the IMDB dataset when BERT and XLNet are used as the base polarity classifier[3]).

Table 2 Accuracy of polarity classification by subjectivity weighted voting (IMDB dataset)

Method	BERT	XLNet
Sub. Only Voting	0.749	0.667
Simple Voting	0.813	0.829
Weighted Voting	0.816	**0.853**

"Sub. Only Voting" is obviously worse than other methods, indicating that it is not good to totally ignore objective sentences. Our method outperforms the "Simple Voting" baseline for both BERT and XLNet, however, the improvement of the BERT model is rather small. On the other hand, significant improvement is found for XLNet, and our method with XLNet achieves the highest accuracy 0.853.

4.2.3 Results of Language Model with Sub. Filtering

We evaluate the proposed method using the pre-trained language models with the subjectivity filtering explained in Subsection 3.2. In this method, only the subjective sentences are used for training the polarity classification. For comparison, we also evaluate the method using all (both subjective and objective) sentences and only objective sentences in each training and test data. Table 3 shows the accuracy on the IMDB and Amazon datasets. The system using S+O (subjective and objective sentences) as the training and test data is the baseline that simply applies BERT or XLNet without filtering ("BL" in Table 3), while the systems using S (subjective sentences) as the training data and S+O or S as the test data are our proposed systems ("PRO1" or "PRO2"). The best system among ones trained from the same training data is indicated in bold.

As for the IMDB dataset, BERT always achieves better accuracy than XLNet. When the settings (S+O, S, or O) of the training and test data are the same, the accuracy becomes the highest. It seems reasonable because the classifiers are fine-tuned using the training data obtained

2) It is a similar approach of previous work [4, 5] that filtered out objective sentences.

3) We did not evaluate the methods using the Amazon dataset, since the accuracy was much worse than our second method as reported in 4.2.3.

Table 3 Accuracy of polarity classification by language models

	Training	Test	IMDB		Amazon	
			BERT	XLNet	BERT	XLNet
(BL)	S+O	S+O	**0.997**	0.975	**0.939**	0.938
		S	0.749	0.701	0.920	0.928
		O	0.668	0.601	0.806	0.803
(PRO1)	S	S+O	0.886	0.819	**0.953**	0.924
(PRO2)		S	**0.980**	0.962	0.918	0.900
		O	0.663	0.616	0.800	0.764
	O	S+O	0.859	0.638	**0.924**	0.892
		S	0.743	0.669	0.894	0.868
		O	**0.963**	0.646	0.799	0.765

by the same filtering strategy as the test data. The baseline achieves the best accuracy, 0.997. Thus the filtering of the objective sentences is not effective in the IMDB dataset.

As for the Amazon dataset, BERT is slightly better than XLNet but they are almost comparable. Comparing the settings of the test data, the systems using subjective and objective sentences (S+O) are always the highest, following only S and only O. In the test data, the subjective sentences seem more effective than the objective sentences, but the latter also includes some useful information. The best system is one of our proposed methods where S is the training data and S+O is the test data. It indicates that the removal of the objective sentences from the training data is effective to improve the quality of the polarity classifier using BERT.

Finally, it is found that the methods using the pre-trained language model (Table 3) are much better than the voting methods (Table 2) on the IMDB dataset. Those results prove that the pre-trained language model is powerful and effective for the polarity classification as reported in many previous papers on various NLP tasks.

5 Conclusion

This paper presented the methods of document-level polarity classification that consider the subjectivity of the sentences. The experimental results demonstrated that the weighting or filtering of the subjectivity sentences could improve the performance in some cases. However, our method did not always achieve the best performance on the IMDB dataset. In the future, we will investigate the major reason for it and explore a robust subjectivity-oriented method that can be effective for any datasets.

References

[1] Valerio Basile, Nicole Novielli, Danilo Croce, Francesco Barbieri, Malvina Nissim, and Viviana Patti. Sentiment polarity classification at evalita: Lessons learned and open challenges. **IEEE Transactions on Affective Computing**, Vol. 12, No. 2, pp. 466–478, 2018.

[2] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In **Proceedings of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies**, pp. 4171–4186, 2019.

[3] Zhilin Yang, Zihang Dai, Yiming Yang, Jaime Carbonell, Russ R. Salakhutdinov, and Quoc V. Le. XLNet: Generalized autoregressive pretraining for language understanding. In **Advances in Neural Information Processing Systems**, Vol. 32. Curran Associates, Inc., 2019.

[4] Bo Pang and Lillian Lee. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. **arXiv preprint cs/0409058**, 2004.

[5] C Sindhu, Binoy Sasmal, Rahul Gupta, and J Prathipa. Subjectivity detection for sentiment analysis on Twitter data. In **Artificial intelligence techniques for advanced computing applications**, pp. 467–476. Springer, 2021.

[6] Ahmad Kamal. Subjectivity classification using machine learning techniques for mining feature-opinion pairs from web opinion sources. **arXiv preprint arXiv:1312.6962**, 2013.

[7] Marco Pota, Mirko Ventura, Rosario Catelli, and Massimo Esposito. An effective BERT-based pipeline for Twitter sentiment analysis: A case study in Italian. **Sensors**, Vol. 21, No. 1, p. 133, 2020.

[8] Andrew Maas, Raymond E Daly, Peter T Pham, Dan Huang, Andrew Y Ng, and Christopher Potts. Learning word vectors for sentiment analysis. In **Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies**, pp. 142–150, 2011. https://ai.stanford.edu/~amaas/data/sentiment/index.html.

[9] Julian McAuley and Jure Leskovec. Hidden factors and hidden topics: Understanding rating dimensions with review text. In **Proceedings of the 7th ACM Conference on Recommender Systems**, pp. 165–172, 2013.

[10] Subjectivity datasets, (Accessed on Dec. 2022). https://www.cs.cornell.edu/people/pabo/movie-review-data/.