A Study on Subjectivity-oriented Polarity Classification

2010417 Dong Na

Polarity classification is a task to identify sentiment or polarity expressed in a text. It is a kind of classification problems where classification labels are defined as positive, neutral, or negative. Accurate polarity classification is crucial for many Natural Language Processing (NLP) applications such as recommendation systems, content creation, and marketing. In addition, user reviews are often considered as a text for polarity classification. Because nowadays a lot of people write reviews on various things such as movies, products, restaurants etc. and they are valuable sources for opinion mining.

Numerous studies have been devoted on polarity classification. In the most of current approaches, however, the subjectivity of sentences is often ignored, which can lead to inaccurate or inconsistent results. This is because the sentiment of a sentence may vary greatly depending on the perspective of the person expressing it. It is worth investigating how to consider the subjectivity in the polarity classification to improve the performance of it.

The goal of this thesis is to propose two methods for document-level polarity classification that take into account the subjectivity of sentences to a large extent. Our approach is based on the idea that the polarity of a document is strongly influenced by its level of the subjectivity. Sentences with strong subjectivity are more likely to express personal opinions, and such sentences play an important role in the polarity classification, while objective sentences are more likely to provide objective information and tend to be less important. By considering the intensity of the subjectivity of sentences, we aim to improve the accuracy of polarity classification.

Two subjectivity-oriented polarity classification methods are proposed. The first one is called Polarity Classification by Subjectivity Weighted Voting (PCBWV). First, after preprocessing including tokenization, lemmatization, and removing stop words, we split a review into sentences. Then, the polarity of each sentence is identified by a polarity classifier. BERT or XLNet is used as the polarity classifier, which is the state-of-the-art language model. The pre-trained model, bert-base-uncased for BERT or xlnet-base-cased for XLNet, is fine-tuned using the polarity labeled data. Next, the subjectivity (subjective or objective) of each sentence is identified. The subjectivity classifier is constructed by fine-tuning the pre-trained XLNet (xlnet-base-cased) using the subjectivity dataset. The subjectivity classifier also provides the prediction scores for each subjectivity class. They are normalized by the soft max function, then the subjectivity score of each sentence, which expresses the intensity of the subjectivity, is obtained. Finally, the polarity label of

the overall review is determined by the weighted voting of the number of the positive or negative sentences, where the subjectivity score of each sentence is used as the weight.

The second method is called Polarity Classification by Pre-trained Language Model with Subjectivity Filtering (PCPLM-SF). First, the review is split into sentences after preprocessing in the same way in the first model PCSWV. Next, the subjectivity of each sentence in the review is determined by the subjectivity classifier, which is the same XLNet model used in PCSWV. If a sentence is classified as objective, it is filtered out, and only sentences classified as subjective are retained. The remained subjective sentences are concatenated in the order they appear in the original review to form a pseudo review. Finally, the polarity of the overall review is determined by BERT or XLNet that accepts the pseudo review as an input. The pre-trained model bert-base-uncased and xlnet-base-cased are used again and fine-tuned using a set of pseudo reviews consisting of only subjective sentences. An variation of PCPLM-SF uses the original review including both subjective and objective sentences as an input in a test phase, though the polarity classification model is fine-tuned using the subjective sentence only.

Several experiments are conducted to evaluate our proposed methods. Three datasets are used. The subjectivity dataset consists of 5,000 subjective sentences from the Rotten Tomatoes website and 5,000 objective sentences from IMDb plot summaries. The IMDb dataset consists of 50,000 movie reviews including 25,000 positive and 25,000 negative reviews. Each half of it is used as the training or test data. The Amazon dataset is a collection of user reviews posted on the Amazon website. We randomly select 25,000 for training and 25,000 for test. The first dataset is used for training the subjectivity classifier, while the second and third datasets are used to train the polarity classifier and measure the accuracy of the proposed methods and baselines.

As for the evaluation of the PCBWV method, the best accuracy, 85.3%, is obtained by our method, where XLNet is used as the sentence-level polarity classifier module. It is better than the accuracy of the baseline, 82.9%, which simply counts positive and negative sentences and chooses the most frequent polarity class without considering the subjectivity. In addition, it is found that XLNet performs better than BERT in the PCBWV.

As for the evaluation of the PCPLM-SF method, three sets of sentences are used as the training and test data: S (subjective sentences), O (objective sentences), and S+O (both subjective and objective sentences). Thus, the polarity classifiers of nine (3×3) settings are compared. Our proposed methods are named as PCPLM-FS-1 using S as both the training and test data as well as PCPLM-FS-2 using S as the training data and S+O as the test

data. The results indicate that the PCPLM-SF-1 using BERT achieves the highest accuracy of 95.3% for the Amazon dataset. For the IMDb dataset, the accuracy of the PCPLM-SF-2 using BERT is 98.0%, which is higher than PCPLM-SF-1. However, it is worse than the baseline using S+O as both the training and test data (i.e., the ordinary BERT), whose accuracy is 99.7%. In this experiment, BERT outperforms XLNet overall. Comparing the IMDb and Amazon datasets, the accuracy on Amazon is better than that on IMDb, except for the case when both training and testing sets were S+O.

Finally, comparing our two proposed methods, the PCPLM-SF method performs much better than the PCBWV method. It indicates that the pre-trained language model is powerful and effective for the polarity classification as reported in many previous work on various NLP tasks. In addition, error analysis is carried out to reveal limitations of our method. It is found that objective sentences sometimes carry important information that facilitates polarity classification.

Despite the promising results, there are still some open challenges that need to be addressed. The main challenges are to apply our method to reviews written in not only English but also other languages, and to evaluate our methods on various domains, not only movies and products.