Features Extraction Based on Neural Network for Cross-Domain Sentiment Classification

Endong Zhu[1], Guoyan Huang[1], Biyun Mo[1], and Qingyuan Wu[2](✉)

[1] Sun Yat-sen University, Guangzhou, China
{zhuend,huanggy7,moby5}@mail2.sysu.edu.cn
[2] Beijing Normal University, Zhuhai, China
wuqingyuan@bnuz.edu.cn

Abstract. Sentiment analysis is important to develop marketing strategies, enhance sales and optimize supply chain for electronic commerce. Many supervised and unsupervised algorithms have been applied to build the sentiment analysis model, which assume that the distributions of the labeled and unlabeled data are identical. In this paper, we aim to deal with the issue of a classifier trained for use in one domain might not perform as well in a different one, especially when the distribution of the labeled data is different with that of the unlabeled data. To tackle this problem, we incorporate feature extraction methods into the neural network model for cross-domain sentiment classification. These methods are applied to simplify the structure of the neural network and improve the accuracy. Experiments on two real-world datasets validate the effectiveness of our methods for cross-domain sentiment classification.

Keywords: Cross-domain sentiment classification · Neural network · Feature extraction · Transfer learning

1 Introduction

Sentiment analysis aims at classifying sentimental data into polarities (e.g., positive and negative) [1] or multiple emotional labels (e.g., happiness, sadness and surprise) [2] primarily. Traditional classification algorithms have been used to train sentiment classifiers [3]. However, these approaches are domain dependent. A classifier trained from one domain (source domain) might not work well when directly applied to another domain (target domain) [4]. The reason is that different domains focus on different sets of topics, which leads to various distributions of features.

In light of these considerations, we propose effective feature extraction methods based on the neural network model for cross-domain sentiment classification. Neural networks are widely employed in natural language processing, but their training process are time-consuming [5] and the risk of overfitting may be high when the data size is large. One strategy to tackle this problem is to apply

The authors contributed equally to this work.

H. Gao et al. (Eds.): DASFAA 2016 Workshops, LNCS 9645, pp. 81–88, 2016.
DOI: 10.1007/978-3-319-32055-7_7

pre-processing on data based on feature extraction methods, which simplify the structure of the neural network, reduce the cost of training process and further improve the accuracy of sentiment classification.

The rest of the paper is organized as follows. The next section reviews some related work. We present our feature extraction methods based on the neural network model for cross-domain sentiment classification in Sect. 3. The datasets and experimental results are discussed in Sect. 4. The last section draws conclusions and points out our future work.

2 Related Work

2.1 Features Extraction

Singular value decomposition (SVD) is a matrix decomposition method in linear algebra, and it has been widely used in the extraction of features. Francesca et al. [6] used SVD to select important features for their probabilistic model. Harb et al. [7] introduced an algorithm *SLRS* based on SVD, which also applied SVD to the document matrix. SVD method can extract important features and reduce the computing complexity by filtering many noise features, i.e., reducing the dimension of the document matrix.

2.2 Cross-Domain Sentiment Classification

Sentiment classification has been extensively studied for reviews [3] and news articles [8–10]. However, these methods focused on detecting emotions specific to one context primarily.

To tackle the issue of different feature distributions in cross-domain datasets, Pan et al. [1] proposed to divide words into two categories for sentiment classification. The first category contains domain-independent (DI) words, and the second one contains domain-specific (DS) words. DI words are those which occur frequently in both the source domain and the target domain, and DS words are those which occur frequently only in one specific domain. Dai et al. [11] introduced a classification model based on co-clustering by the following two processes. Firstly, in-domain documents were used to generate the class structure and propagate the label information. Secondly, co-clustering [12] was applied to out-of-domain data to obtain document and word clusters. Rao [13] developed a contextual sentiment topic model to distinguish context-independent topics from both a background theme and a contextual theme. The limitation of these studies, however, was that they were designed for normal documents which contain sufficient words. In this work, we focus on cross-domain sentiment classification over both normal and short documents.

3 Features Extraction Algorithms

In this section, we describe four effective methods to extract features and apply them to neural networks for cross-domain sentiment classification.

3.1 Naive Methods

The naive methods extract features based on their occurrences in the source domain and the target domain, which are totally unsupervised.

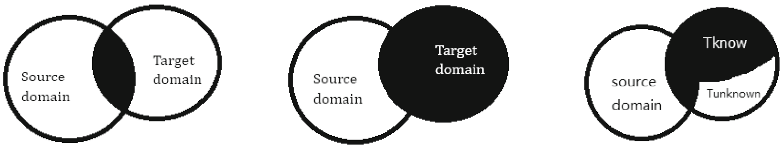

Fig. 1. Algorithm description: (a) words exist in both the source domain and the target domain, (b) words only exist in the target domain, (c) word selection based on limited labeled data in the target domain

Two simple ways are used to extract important features for cross-domain sentiment classification. The first one selects words which exist in both the source domain and the target domain, as shown in Fig. 1(a). The second one selects words which only exist in the target domain, as shown in Fig. 1(b).

3.2 Using Limited Labeled Data in Target Domain

This method extracts features by exploiting a small part of labeled data in the target domain, as shown in Fig. 1(c). The objective is to use as less labeled data in the target domain as possible for training.

We divide the target domain data into two parts with different proportions: the labeled part T_{know} in the target domain is used together with all labeled data in the source domain for training; and the rest unlabeled part T_{unknow} in the target domain is used for testing. Two kinds of features are selected to cross-domain sentiment classification: (1) all features in T_{know}. The reason is that features in T_{know} may be more important than those in the source domain; and (2) features of the source domain which occur in the target domain.

3.3 SVD-Based Method

To filter noise features, this method applies SVD to the $m \times n$ document matrix M, as follows:

$$M_{m\times n} = U_{m\times m}\Sigma_{m\times n}V^T_{n\times n}, \tag{1}$$

where U and V are orthogonal matrices, and Σ is a diagonal matrix where the diagonal elements are eigenvalues of the matrix M. Since the eigenvalues in Σ are sorted in the decreasing order, we can use r large eigenvalues to represent the matrix M approximately. According to the properties of orthogonal matrices, we obtain a new matrix A. The mathematical processes are described as follows:

$$M_{m\times n} \approx \bar{M}_{m\times n} = U'_{m\times r}\Sigma_{r\times r}V^T_{r\times n}, \tag{2}$$

$$M_{m\times n}V'_{n\times r} \approx U'_{m\times r}\Sigma'_{r\times r}\left(V'^T_{r\times n}V'_{n\times r}\right) = U'_{m\times r}\Sigma'_{r\times r}, \tag{3}$$

$$A_{m\times r} = M_{m\times n}V'_{n\times r}, \tag{4}$$

where A is a $m \times r$ matrix. The matrix V' is used to compress M into A, and this compression process can be considered as a feature extraction process if we represent the columns of M as features.

3.4 Feature Alignment Method

Sentiment classification is very domain-specific because users can use domain-specific words to convey emotions in different contexts. For instance, the word "compact" is often used to express a positive attitude in the domain of electronic product reviews. However, the word "interesting" expresses a primarily positive attitude in the books domain. Inspired by the previous work [1], we also evaluate the feature alignment method on neural networks for cross-domain sentiment classification. Algorithm 1 presents the process of selecting domain-independent (DI) and domain-specific (DS) words, where ε is a threshold. Note that all stop words have been deleted beforehand.

```
for each word w in the set of all words do
    Compute the frequency f_s in the source domain ;
    Compute the frequency f_t in the target domain ;
    if f_s > ε and f_t > ε then
        Add w to the set of DI words;
    end
    else
        Add w to the set of DS words;
    end
end
```

Algorithm 1. Selecting DI/DS words

In our neural network models, we use all DI words as the first part of features directly. To improve the generalization of models, we also apply the clustering algorithm to DS words and use the clusters of DS words as the second part of features. Algorithm 2 presents the process of clustering DS words. Firstly, all DI words are used as the "bridge" to construct an adjacent matrix $wordMat$. Secondly, the dimension-reduction method is conducted on the high-dimensional and sparse adjacent matrix. Finally, the unsupervised method such as spectral clustering and K-means is employed to generate the clusters of DS words.

for *each word* w_1 *in the set of DS words* **do**
  for *each word* w_2 *in the set of DI words* **do**
    if w_1 *and* w_2 *occur in the same document* **then**
      $wordMat[w_1][w_2]$ += 1 ;
    end
  end
end
Conduct dimension-reduction on $wordMat$;
Conduct clustering on $wordMat$;

Algorithm 2. Clustering DS words

4 Experiments

To evaluate our feature extraction methods, experiments are designed to cross-domain sentiment classification via neural networks. The neural network using all words as features is implemented for comparison.

4.1 Datasets

SemEval is an English dataset used in the 14th task of the 4th International Workshop on Semantic Evaluations (SemEval-2007) [14]. It contains news headlines extracted from Google news, CNN, and many others. We use the development set (i.e., the source domain) for training and the testing set (i.e., the target domain) for evaluation. The news headlines in the source domain and the target domain are very different in terms of words and topics.

SinaNews is a Chinese dataset which contains news articles extracted from Sina website [13]. We use a total of 2,343 documents published from January to February 2012 for training, and the rest 2,228 documents published from March to April 2012 for testing. Due to that the publication dates vary, the feature distributions in the source and target domains are also quite different.

4.2 Experimental Design

In our experiments, we use the back-propagation neural network (BP neural network) [15] as the classification model to compare the performance of different feature extraction methods. The reasons of employing BP neural network to evaluate the effectiveness of our algorithms are summarized as follows: First, we can use the control rules by non-linear mapping and self-learning of neural networks. Second, feature extraction is quite important for BP neural network to reduce the number of units in the input layer.

The structure of BP neural network is one input layer, one hidden layer and one output layer for simpleness. The number of units in the input layer equals the amount of features, and the number of units in the output layer equals the amount of emotion labels. We use the following formula to determine the number units in the hidden layer [16]:

$$N_{hidden} = \sqrt{N_{input} + N_{output}} + m \quad (-5 \leq m \leq 5)\,, \tag{5}$$

where the value of m is set to 0.

The accuracy at top 1 is employed as the indicator of performance [9], which is essentially the micro-averaged F1 that equally weights precision and recall.

4.3 Results and Analysis

The accuracies of BP neural network using all words as features are 37.0 % and 53.59 % on *SemEval* and *SinaNews*, respectively. This method is used as the baseline model because it outperforms most state-of-the-art models in the previous studies on these two datasets [8,13].

Table 1. Results of naive methods

Dataset	# of features	Accuracy
SemEval	535	35.5 %
	2380	38.2 %
SinaNews	19774	57.18 %
	21310	56.95 %

The results of naive methods are presented in Table 1. Compared to the 2,749 and 22,946 features in the baseline model on *SemEval* and *SinaNews*, the numbers of features are reduced by the naive methods as presented in Fig. 1(a) and (b). In terms of the accuracy, the performance of models using the naive methods to extract features is worse on *SemEval* when using too few features, and better on *SinaNews* than the baseline model. The results indicate that the effect of the naive methods relies on the scale of the dataset.

The accuracies of using several proportions of labeled data in the target domain for training are shown in Table 2, from which we can observe that the accuracy increases as more labeled data in the target domain are available. However, the gap between the source and target domains may have a negative influence on the performance when using too many features.

Table 2. Accuracies of using limited labeled data in the target domain

Proportion	*SemEval*	*SinaNews*
50 %	42.50 %	57.99 %
25 %	41.60 %	57.49 %
12.5 %	44.29 %	57.48 %
6.25 %	41.47 %	57.01 %
3.125 %	39.93 %	57.29 %

Table 3. Results of SVD-based method

Dataset	# of features	Accuracy
SemEval	2749	37.0 %
	1000	37.9 %
	800	37.8 %
	500	37.1 %
SinaNews	5000	56.86 %
	2000	56.91 %
	1000	56.82 %
	500	57.54 %

Table 4. Results of feature alignment method

Dataset	# of DI words	# of DS word clusters	Accuracy
SemEval	526	50	38.80 %
	526	100	38.40 %
	526	150	39.30 %
	526	200	40.10 %
SinaNews	2168	500	52.69 %
	2168	1000	58.25 %
	2168	1500	58.48 %
	2168	2000	59.57 %

The result of SVD-based method is presented in Table 3, which indicates that the accuracy depends on the degree of compressing the document matrix.

The results of cross-domain sentiment classification based on the feature alignment method are shown in Table 4, from which we can observe that the performance improves as the number of DS word clusters increases.

5 Conclusion

In this paper, we incorporated four effective features extraction methods into BP neural network to reduce the dimension and also improve the accuracy of cross-domain sentiment classification. Extensive evaluations using two real-world datasets validate the effectiveness of the proposed features extraction algorithms. As a trade-off strategy, a good dimension-reduction method should reduce the dimension of features as much as possible with the guarantee of accuracies.

For future work, we plan to develop methods of choosing the optimal parameters for our features extraction algorithms, in addition to combine other methods to tackle the problems in cross-domain sentiment classification.

Acknowledgements. This research has been substantially supported by a grant from the Soft Science Research Project of Guangdong Province (Grant No. 2014A030304013).

References

1. Pan, S.J., Ni, X., Sun, J.-T., Yang, Q., Chen, Z.: Cross-domain sentiment classification via spectral feature alignment. In: Proceedings of the 19th International Conference on World Wide Web, pp. 751–760 (2010)
2. Zhang, Y., Zhang, N., Si, L., Lu, Y., Wang, Q., Yuan, X.: Cross-domain and cross-category emotion tagging for comments of online news. In: Proceedings of the 37th International ACM SIGIR Conference, pp. 627–636 (2014)
3. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up? sentiment classification using machine learning techniques. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 79–86 (2002)
4. Nigam, K., McCallum, A.K., Thrun, S., Mitchell, T.: Text classification from labeled and unlabeled documents using em. Mach. Learn. **39**(2–3), 103–134 (2000)
5. Morchid, M., Dufour, R., Linares, G.: Topic-space based setup of a neural network for theme identification of highly imperfect transcriptions. In: IEEE Automatic Speech Recognition and Understanding Workshopp (2015)
6. Francesca, F., Zanzotto, F.M.: SVD feature selection for probabilistic taxonomy learning. In: Proceedings of the Workshop on Geometrical Models of Natural Language Semantics, pp. 66–73 (2009)
7. Dasgupta, A., Drineas, P., Harb, B., Josifovski, V., Mahoney. M.W.: Feature selection methods for text classification. In: Proceedings of the 13th ACM SIGKDD International Conference, pp. 230–239 (2007)
8. Rao, Y., Lei, J., Liu, W., Li, Q., Chen, M.: Building emotional dictionary for sentiment analysis of online news. World Wide Web J. **17**, 723–742 (2014)
9. Rao, Y., Li, Q., Liu, W., Wu, Q., Quan, X.: Affective topic model for social emotion detection. Neural Netw. **58**, 29–37 (2014)
10. Rao, Y., Li, Q., Mao, X., Liu, W.: Sentiment topic models for social emotion mining. Inf. Sci. **266**, 90–100 (2014)
11. Dai, W., Xue, G.-R., Yang, Q., Yu, Y.: Co-clustering based classification for out-of-domain documents. In: Proceedings of the 13th ACM SIGKDD International Conference, pp. 210–219 (2007)
12. Dhillon, I.S., Mallela, S., Modha, D.S.: Information-theoretic co-clustering. In: Proceedings of the 9th ACM SIGKDD International Conference, pp. 89–98 (2003)
13. Rao, Y.: Contextual sentiment topic model for adaptive social emotion classification. IEEE Intell. Syst. **31**(1), 41–47 (2016)
14. Strapparava, C., Mihalcea, R.: Semeval- task 14: affective text. In: Proceedings of the 4th International Workshop on Semantic Evaluations, pp. 70–74 (2007)
15. Hecht-Nielsen, R.: Theory of the backpropagation neural network. In: International Joint Conference on Neural Networks, pp. 593–605 (1989)
16. Kurkova, V., Kainen, P.C., Kreinovich, V.: Estimates of the number of hidden units and variation with respect to half-spaces. Neural Netw. **10**(6), 1061–1068 (1997)