Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)

Cross-Domain Sentiment Classification via Topic-Related TrAdaBoost

Xingchang Huang,[1] Yanghui Rao,[1*] Haoran Xie,[2] Tak-Lam Wong,[2] Fu Lee Wang[3]
[1] School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China
[2] Department of Mathematics and Information Technology, The Education University of Hong Kong,
Tai Po, New Territories, Hong Kong
[3] Caritas Institute of Higher Education, Tseung Kwan O, Hong Kong
+86-13750023601, huangxch3@mail2.sysu.edu.cn
raoyangh@mail.sysu.edu.cn, hrxie2@gmail.com, tlwong@eduhk.hk, pwang@cihe.edu.hk
*The Corresponding Author

Abstract

Cross-domain sentiment classification aims to tag sentiments for a target domain by labeled data from a source domain. Due to the difference between domains, the accuracy of a trained classifier may be very low. In this paper, we propose a boosting-based learning framework named TR-TrAdaBoost for cross-domain sentiment classification. We firstly explore the topic distribution of documents, and then combine it with the unigram TrAdaBoost. The topic distribution captures the domain information of documents, which is valuable for cross-domain sentiment classification. Experimental results indicate that TR-TrAdaBoost represents documents well and boost the performance and robustness of TrAdaBoost.

Introduction

Detecting the sentimental category of documents can help understand the feeling of users towards products or events. Thus, sentiment classification is widely studied nowadays. However, a sentiment classifier trained in one domain may not perform as well in others. This is because users often use domain specific words to express sentiments in different domains, which may lead to the word distribution of documents differs between the source and the target domain.

Due to the problem caused by domain-specific words, the TrAdaBoost (Dai et al. 2007) is developed to improve the generalization ability of classifiers. Given a large number of labeled data from the source domain (Tr_d), and a small amount of labeled data from the target domain (Tr_s), the idea of TrAdaBoost is to use boosting method to adjust the weights of training instances from Tr_d and Tr_s. Compared to the target domain, feature distributions of Tr_d are often more different than Tr_s. Thus, instances of Tr_d should have smaller weights than Tr_s. However, the effectiveness of TrAdaBoost is highly dependent on Tr_s if applied to cross-domain sentiment classification. For instance, the performance of TrAdaBoost decreased when Tr_s mainly contain domain-independent words. This is because Tr_s can not well represent the target domain with a few domain-specific words. To tackle this issue, we propose a topic-related boosting model to represent and classify sentimental documents. Experimental results validate the effectiveness of our method on cross-domain sentiment classification.

Related Work

Cross-domain sentiment classification has drawn much attention in researchers. Zhang et al. (Zhang et al. 2014) developed a framework to adapt two domains with the same or different emotion categories using probabilistic models. Pan et al. (Pan et al. 2010) proposed the spectral feature alignment algorithm that uses some domain-independent words as a bridge to cross-domain sentiment classification. Li et al. (Li, Jin, and Long 2012) employed the topic correlation analysis method to group words into topics, and used these topics to bridge domains with a well-defined probabilistic topic model. However, the limitation of these approaches is that they require a sufficient collection of target domain samples to guarantee improvements.

Given a small amount of labeled data from the target domain (Tr_s), the TrAdaBoost (Dai et al. 2007) tries to decrease the weights of training instances that have different feature distributions with the target domain. However, the effectiveness of TrAdaBoost depends on the quality of Tr_s, and it can not deal with multiple different domains simultaneously. To this end, we extract the domain information for sentimental documents from both source and target domains by unsupervised topic modeling. Our method not only improves the performance of TrAdaBoost framework, but also adapts to multi-domain sentiment classification.

Topic-Related TrAdaBoost

To extract both domain-specific and shared topics between source and target domains, we employ the Latent Dirichlet Allocation (LDA) (Blei, Ng, and Jordan 2003) and generate topic distributions of documents. Given the document-topic matrix $\hat{\theta}_{it}$, we can get the topic-level representation of each document i as $\phi_i = [\hat{\theta}_{i1}, \hat{\theta}_{i2}, ..., \hat{\theta}_{it}, ..., \hat{\theta}_{iT}]$, where ϕ_i represents the topic distribution of the $i\text{-}th$ document and $\hat{\theta}_{it}$ denotes the topic mixture proportion on topic t for document i. Then, we append the topical distribution ϕ_i to the unigram model to construct a new representation of document i. With this new representation, each document is transformed into a new feature space that is represented by both words and topics, which is more adaptive than unigram-based methods to the framework of TrAdaBoost for cross-domain sentiment classification tasks. Note that we can also apply the above feature space to other classifiers (e.g., SVM).

Algorithm 1 Topic-Related TrAdaBoost

Input: Training sets Tr_d and Tr_s; Testing set S;
Maximum number of iterations N;
Number of topics T;

1: Extracting the topic distribution ϕ_i for each document i with T topics.
2: Construct the new representation for each document i by append the ϕ_i to the unigram model.
3: Initialize the weight vector $\mathbf{w}^1 = (w_1^1, ..., w_{n+m}^1)$ as random numbers.
4: For $t = 1, ..., N$:
a. Train a basic learner using the combined data set T with different weights for instances according to w^t. Then, get the hypothesis $h_t : X \to Y$
b. Calculate the error rate:

$$\epsilon_t = \sum_{i=n+1}^{n+m} \frac{w_i^t \cdot |(h_t(x_i) + c(x_i))/2|}{\sum_{i=n+1}^{n+m} w_i^t}$$

c. Let $\beta_t = \epsilon_t/(1 - \epsilon_t)$ and $\beta = 1/(1 + \sqrt{2 \ln nN})$ and assume that β_t is less than $1/2$
d. Update the weight vector w_i^{t+1} and continue the loop:

$$w_i^{t+1} = \begin{cases} w_i^t \beta^{|(h_t(x_i)+c(x_i))/2|}, i = 1, ..., n \\ w_i^t \beta_t^{-|(h_t(x_i)+c(x_i))/2|}, i = n+1, ..., n+m \end{cases}$$

Output: Return the hypothesis $h_f(x)$:

$$h_f(x) = sign(\sum_{n=1}^{N} \beta_t h_t(x))$$

The framework of our TR-TrAdaBoost is given in Algorithm 1. For each iteration, we decrease the weight of a misclassified training instance, i.e., the instance of Tr_d that has different feature distributions with Tr_s ones, through multiplying its weight by $\beta^{|(h_t(x_i)+c(x_i))/2|}$ as $Y = \{-1, 1\}$. Therefore, the influence of the misclassified training instances will decrease in the next turn, and the instances of Tr_d that fit those of Tr_s better will have larger weights. The advantage of our method is that even if Tr_s contains little domain-specific information, it can captures the characteristics of the target domain well by mapping word-level features to the topic-level space.

Experiments

The multi-domain sentimental corpora (Blitzer, Dredze, and Pereira 2007) that contains a collection of product reviews from Amazon is employed for experiments. Those reviews are collected from four product domains: books (B), dvd (D), electronics (E) and kitchen (K). Each review is assigned a sentimental polarity of positive or negative.

The performance of different algorithms is presented in Table 1, where the symbol before and after each arrow denotes the source domain and the target domain, respectively. We run 10 independent tests and show the mean of accuracy. The number of topics is determined by cross-validation on the training set. As shown in Table 1, the proposed TR-TrAdaBoost outperformed other baselines for most tasks. We can also observe that our method of combining the word and topic features is also benefit to SVM, i.e., TR-SVM outperformed SVM in nearly all cases.

Table 1: The accuracy of different algorithms on 12 tasks

Tasks	SVM	TR-SVM	TrAdaBoost	TR-TrAdaBoost
B $\to$ D	0.7795	0.7820	0.7881	**0.7955**
B $\to$ E	0.7240	0.7270	0.7397	**0.7490**
B $\to$ K	0.7455	0.7490	0.7675	**0.7775**
D $\to$ B	0.7310	0.7325	0.7329	**0.7470**
D $\to$ E	0.7360	0.7430	0.7484	**0.7585**
D $\to$ K	0.7245	0.7340	0.7561	**0.7565**
E $\to$ B	0.6815	**0.6910**	0.6851	0.6905
E $\to$ D	0.7045	0.7115	0.7129	**0.7180**
E $\to$ K	0.8130	0.8180	0.8321	**0.8365**
K $\to$ B	0.6975	**0.7060**	0.7035	**0.7060**
K $\to$ D	0.7405	0.7400	0.7376	**0.7440**
K $\to$ E	0.8050	0.8055	0.8286	**0.8310**

Conclusion

In this work, we propose a topic-related boosting method, which makes TrAdaBoost more adaptive to cross-domain sentiment classification. The idea of our model is to capture the latent semantic structure by extracting the topic distribution of documents, so as to embed both domain-specific and shared information of documents. For future work, we plan to extend cross-validation method to select best T values for different domains.

Acknowledgments

The research described in this paper was supported by the Natural Science Foundation of China (no. 61502545), a grant from Research Grants Council of Hong Kong Special Administrative Region, China (UGC/FDS11/E06/14), and the Start-Up Research Grant (RG 37/2016-2017R) of The Education University of Hong Kong.

References

Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent dirichlet allocation. *J. Mach. Learn. Res.* 3:993–1022.

Blitzer, J.; Dredze, M.; and Pereira, F. 2007. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In *ACL*, 440–447.

Dai, W.; Yang, Q.; Xue, G.-R.; and Yu, Y. 2007. Boosting for transfer learning. In *ICML*, 193–200.

Li, L.; Jin, X.; and Long, M. 2012. Topic correlation analysis for cross-domain text classification. In *AAAI*, 998–1004.

Pan, S. J.; Ni, X.; Sun, J.-T.; Yang, Q.; and Chen, Z. 2010. Cross-domain sentiment classification via spectral feature alignment. In *WWW*, 751–760.

Zhang, Y.; Zhang, N.; Si, L.; Lu, Y.; Wang, Q.; and Yuan, X. 2014. Cross-domain and cross-category emotion tagging for comments of online news. In *SIGIR*, 627–636.