Big Data Based Transfer Learning for Sentiment Classification with Multiple Source Domains

Qing Li, Gang Wang(✉), and Guowu Yang

School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China

{liqing,guowu}@uestc.edu.cn, carbite_wg@163.com

Abstract. Sentiment classification, served as a curial technology in natural language processing and computational linguistics, has drawn a lot of attentions from researchers. However, due to the high cost of manual labeling in the era of big data, conventional methods of sentiment classification are unqualified to be employed in a new domain directly. Hence, in this paper, we explore big data based transfer learning for sentiment classification with multiple source domains. To solve the problem of inherent domain gap, we propose a novel framework **Ad**versarial cross-domain sentiment classification with weighted domain-dependent f**EA**ture learning dubbed AdEA. Specifically, AdEA involves an individual domain-invariant feature extractor and several domain-dependent feature extractors. To obtain the domain-invariant feature, we use a reversed discriminator loss for these extractors. Furthermore, we propose a weighted learning module to reinforce the relationship of domain-dependent features between source and target domains. Integrated with these two domain-related features, AdEA is able to achieve better capability of cross-domain sentiment classification. Experimental results show the effectiveness of our proposed method.

Keywords: Big Data · Transfer Learning · Sentiment Classification

1 Introduction

In the era of big data, people can easily express their views on the internet, and thus there are flooded with massive amounts of user generated text [7, 13, 15]. To understand them, sentiment classification is a crucial technology in natural language processing, and has drawn a lot of attentions from researchers [11, 12, 19]. Conventional sentiment classification is generally limited in a single domain, which may be unpractical in real-world applications due to the high cost of labeling for different domains. Therefore, based on transfer learning that is an important technology in the area of big data, researchers recently pay their attentions to the task of big data based transfer learning for sentiment classification that aims to train a single model based on supervised source domains (labeled examples in source domain) and unsupervised target domains (unlabeled examples in target domain). How to utilize these domains that contain massive examples is a significant research problem in the area of big data. The core of UDA sentiment classification is to learn the shared sentiment knowledge between different domains. To

Z. Zhan et al. (Eds.): ICBDIE 2022, AHCS 5, pp. 256–265, 2023.
https://doi.org/10.2991/978-94-6463-034-3_26

achieve this aim, recent existing methods generally utilize adversarial learning to obtain the domain-invariant feature. However, they commonly overlook the domain-dependent features, which may affect the classification task of each domain. To solve this problem, recent methods explicitly divide them into two separate parts by using two independent networks. But, they generally ignore the relationship of domain-dependent feature extractors between source and target domains.

To solve this issue, in this paper, we propose a novel framework **Ad**versarial cross-domain sentiment classification with weighted independent f**EA**ture learning dubbed AdEA. Specifically, AdEA includes an individual domain-invariant feature extractor and several domain-dependent feature extractors. To obtain the domain-invariant feature, we use a reversed discriminator loss for these extractors. Furthermore, to reinforce the relationship of domain-dependent features between source and target domains, we propose a weighted learning module to promote learning these domain-dependent features. Experimental results show the effectiveness of our proposed method.

2 Research Background

2.1 Conventional Cross-Domain Sentiment Classification

Conventional methods of sentiment classification are mostly based on statistical machine learning [16, 21, 22]. The representative methods of this line are structure based and spectral clustering methods. For example, Blitzer et al. studied using the structural correspondence learning (SCL) to mine the correspondence between the features of various domains [1]; Pan et al. proposed a technology called spectral feature alignment (SFA) [17]. In SFA, it first aligned domain-specific words from different domains into a unified cluster, and then learned the model from the co-occurrence matrix generated by the mapping of pivot words and non-pivot words. By using pivot words as a bridge, the non-pivot words in the source domain and the target domain could be grouped to obtain several meaningful clusters. After that, it transformed the co-occurrence relationship between pivot words and non-pivot words into a bipartite graph between domains, and then spectral clustering algorithm [14] could be used on the bipartite graph to solve the domain mismatch problem.

2.2 Deep Cross-Domain Sentiment Classification

Recently, with the development of deep neural networks, deep learning based methods achieve better performance than traditional methods. In cross-domain sentiment classification, deep learning based methods aim to use deep neural networks to learn the cross-domain invariant feature representation. For example, Glorot et al. utilized Denoising Auto-Encoder (DAE) to learn domain-specific word representations [9], which accurately captured the meaning of domain specific words. Bolegala et al. proposed a cross domain word representation learning (CDWRL) method that could be used for cross-domain sentiment classification [2]. Furthermore, Generative Adversarial Network GAN [10] is a prominent generation model for various tasks. Based on it, Ganin and Lempitsky proposed adversarial training for domain adaptation [6]. Ziser and Reichart proposed Auto-Encoder based Structural Correspondence Learning (AE-SCL) [23], which encoded the non-pivot of data points into a low dimensional representation. Peng et al. proposed that the shared classifier and the independent classifier in the target domain could be trained respectively by using the information of the source domain and the target domain [18]. Besides, Chen et al. used the auto-encoder framework to learn transferable features [3]. Du et al. applied the pre-trained language model BERT to the unsupervised domain adaptation [5]. Ghosal et al. used the knowledge graph to introduce external common knowledge, and used the learned conceptual knowledge to fine-tune the UDA method [8]. Xue et al. introduced mutual learning into cross domain sentiment classification and designed a deep adversarial mutual learning method [20]. Dai et al. took advantage of the relationship between the target domain and multiple source domains by a weighted scheme based unsupervised domain adaptation framework (WS-UDA) [4].

3 Method

In the traditional solution of cross-domain sentiment classification, it is usually through the adversarial learning to obtain the domain-invariant feature, and then utilize the domain-invariant feature extractor for the target domain. After fine-tuning, it will be used as the final sentiment classification model of the target domain. However, it does not make use of the domain-dependent sentiment knowledge of each source domain, which may affect the classification task of each domain. To solve this issue, we propose a novel framework Adversarial cross-domain sentiment classification with weighted independent fEAture learning dubbed AdEA. Specifically, we introduce an individual domain-invariant feature extractor and several domain-dependent feature extractors. To obtain the domain-invariant feature, we use a reversed discriminator loss for these extractors. Furthermore, to reinforce the relationship of domain-dependent features between source and target domains, we propose a weighted learning module to promote learning these domain-dependent features. The framework of our proposed method is shown in Fig. 1.

In AdEA, k source domains have their own domain-dependent feature extractor $\{E_{pi}\}_1^k$, the target domain also has its own domain-dependent feature extractor E_t. All domains have one shared domain-invariant feature extractor E_s. Besides, there are domain discriminator network D and sentiment classification network C.

In the sentiment knowledge learning stage, the parameters of all feature extraction networks E, discriminator networks D and sentiment classifier networks C are updated and optimized. The main loss is the domain discriminator loss $Loss_D$, sentiment classification loss $Loss_C$, adversarial loss $Loss_{adv}$, and the loss of domain-dependent feature weighted learning $Loss_P$. Therefore, the total loss is shown in Eq. (1).

$$\begin{aligned} Loss_{tot} = Loss_D + \mu Loss_C + \lambda Loss_{adv} \\ + \beta Loss_P. \end{aligned} \tag{1}$$

Next, we will introduce each individual loss respectively. First, the domain discriminator loss $Loss_D$ is shown in Eq. (2). Specifically, the input samples are from all domains including the target domain. We instantiate $\mathcal{L}_D$ as the cross entropy loss.

$$Loss_D = \mathcal{L}_D(D(z), d), \tag{2}$$

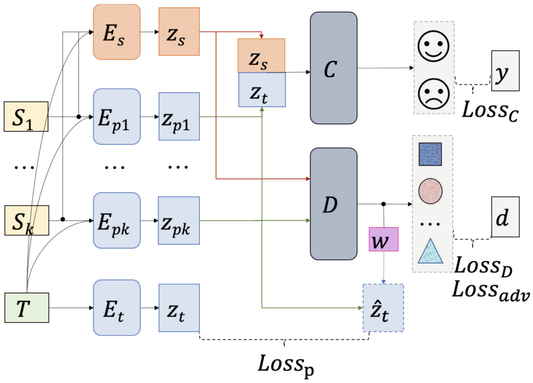

Fig. 1. The framework of our proposed method.

where z and d denote the features extracted by the sample and the number of domains.

Then, sentiment classification Loss of $Loss_C$ is shown in Eq. (3),

$$Loss_C = \mathcal{L}_C\big(C\big(z_s \oplus z_p\big), y\big), \tag{3}$$

where z_s and z_p are the shared feature and the domain-dependent feature, and y is the label of the sample, $\oplus$ indicates the concatenation operation.

To obtain the domain-invariant feature, we aim to make the discriminator network D unable to distinguish the shared domain feature. To achieve this aim, we use a reversed adversarial loss $Loss_{adv}$ shown in Eq. (4):

$$Loss_{adv} = -\mathcal{L}_D(D(z_s), d), \tag{4}$$

where z_s is the shared feature from all domains.

Now, we introduce a weighted learning module to promote learning these domain-dependent features. First, $\hat{z}_t$ can be obtained by weighted averaging k feature and vectors w. The calculation method of $\hat{z}_t$ is shown in Eq. (5).

$$\hat{z}_t = \sum_1^k z_{pi} \cdot w_i, \tag{5}$$

where w_i denotes the relationship between the target domain sample and the i-th source domain.

The loss $Loss_P$ is shown in Eq. (6). By optimizing the $Loss_P$, the E_t network can learn the domain-dependent sentiment knowledge of each source domain.

$$Loss_P = \mathcal{L}_P\big(z_t, \hat{z}_t\big), \tag{6}$$

where z_t is the feature extracted by the target domain sample on E_t.

As for the whole training procedure, we first apply the shared feature extractor E_s and the domain-dependent feature extractor E_t to the target domain. Then, the target domain samples need to be sent to the domain-dependent feature extractors of multiple source domains to finally obtain multiple sentiment classification results, and then the multiple sentiment classification results are weighted and averaged according to the relationship between the target domain samples and the multiple source domains to obtain the pseudo label of the sample. If the confidence of the pseudo label is greater than a certain value, the sample can be added to the training set $\widehat{T}$.

After that, the samples of the target domain are sent into the trained sentiment classifier of k source domains, and we can get k classification results, which are denoted as $\hat{y}_i$, The extracted k domain-dependent features are sent to the D network and the results are normalized to obtain the weight vector w. Finally, $\hat{y}_i$ is multiplied by the corresponding weight to get the pseudo label $\hat{y}$. The pseudo label calculation method is shown in Eq. (7). If the confidence of the pseudo label is greater than a certain value θ. The sample can be selected into the training set.

$$\hat{y} = \sum_{i}^{k} y_i \cdot w_i \tag{7}$$

In order to avoid the situation of over-confidence in each network after the training in the sentiment knowledge learning stage, we add a small random Gaussian noise to the initial sample of the target domain. If the confidence of multiple experiments is higher than the specified threshold and the corresponding pseudo label is the same, we select it into the training set for further training and fine-tuning.

4 Experiment

In this section, we mainly introduce our experimental setting and results. We first briefly introduce several groups of baseline models and explain the specific parameter settings in the experiment. Finally, the experimental results and analysis will be introduced to verify the effectiveness of our proposed method.

4.1 Dataset

We use the online_shopping_10cats dataset to carry out the experiment. There are 10 categories in the data set (books, tablets, mobile phones, fruits, shampoo, water heater, Mengniu, clothes, computers, hotels), with a total of more than 60000 comment data, and about 30000 positive and negative comments respectively. The detail of each category sample in the dataset is shown in Table 1. Furthermore, we use a series of preprocessing operations on the data of the other nine domains.

Since there are too few samples in the field of water heater and the positive and negative samples are extremely unbalanced, it is difficult to fairly regard it as a domain-dependent field. Therefore, we will not use the field of water heater to participate in the experiment, but only use the data of the other nine fields.

Besides, we take the comment data of each category as a single domain, and then use cross validation, i.e., using 8 domains as the source domain and the other one as the target domain in turn.

Table 1. online_shopping_10cats.

Category	*Positive*	*Negative*
book	2100	1751
tablets	5000	5000
mobile phones	1165	1158
fruits	5000	5000
shampoo	5000	5000
water heater	475	100
Mengniu	992	1041
clothes	5000	5000
computers	1996	1996
hotels	5000	5000

4.2 Baselines

1) Bi-Transferring AutoEncoder (BTAE): this method was proposed by Zhou et al. Bi-Transferring means that the self-encoder can transfer the data of the source domain to the target domain, and also transfer the data of the target domain to the source domain. Compared with the traditional Auto-Encoder, BTAE includes one encoder and two decoders acting on the source domain and the target domain respectively. The purpose of the encoder is to map the input samples of the source domain and the target domain to the same feature space to obtain the implicit feature representation with the same distribution. Different decoders try to map the implicit feature representation to the source domain and the target domain respectively.
2) Multi-source Domain Adversarial Network (MDAN): This method was Proposed by Zhao et al. Different from DANN that involves one single source domain, MDAN is the generalized version of DANN. MDAN includes a feature extraction network, a sentiment feature classification network and a group of domain identification networks. Feature extractor and affective feature classification network are responsible for learning sentiment knowledge and decision boundaries from source domain data with rich labels. The feature extractor and domain discriminate network form an antagonistic relationship. The feature extraction network attempts to extract domain-invariant features, while the domain discriminate network aims to distinguish the domain of the features. MDAN adopts gradient inversion layer to optimize parameters.
3) Weighting Scheme based Unsupervised Domain Adaptation (WS-UDA): WS-UDA is based on a "shared-independent feature" training framework to realize the separation of shared features and independent features. At the same time, the relationship between target domain samples and multiple source domains is established by using the discriminator network of adversarial training. Using this relationship and the sentiment classification model of multiple source domains, the sentiment classification task can be carried out for the target domain.

4) Two-Stage Training based Unsupervised Domain Adaptation(2ST-UDA): First, they use WS-UDA to obtain the pseudo labels of the target domain, and then use the domain-dependent feature extractor of the target domain to classify the features to obtain another group of pseudo labels, and select the samples with the same sentiment classification results of the two models into the training set. After that, the domain-dependent feature extractor in the target domain is trained again until the confidence is less than 0.5 or the sample is no longer significantly increased. Finally, all the accumulated pseudo label training sets are sent to the independent feature extraction network of the target domain, and the network is further trained. The final model is the sentiment classification model of the target domain.

4.3 Model Parameter Setting

We choose the word2vec model as CBOW, and the dimension of the word vector is 200. The parameters of all networks follow the normal distribution with mean value of 0 and standard deviation of 0.1. Bi-GRU network is adopted for the feature extractor network, and its output dimension is 100. Since the model is applied twice before and after the feature extractor network, the final dimension is 200. The activation function is ReLU, and the size of batch is 64. Feature classification network C adopts fully connected network, and domain discriminator network D adopts three-layer perceptron network. The parameter optimizer method of all networks adopts Adam method, and the initial value of learning rate is set to 0.001. Besides, this paper will use dropout technology for all networks and set the dropout parameter to 0.8 (80% neurons are reserved).

4.4 Experimental Result

The experimental result is shown in Table 2. In Table 2, each column denotes a target domain, and each row denotes different cross domain sentiment classification methods. DirT is direct transfer, i.e., the models trained in all other source domains are directly taken to the target domain for testing. The actual calculation method is as follows: first we train the sentiment classification tasks of 8 source domains separately, and then apply them to the test set of the target domain after these 8 tasks tend to converge. Finally, we average the accuracy of the 8 tests as the final accuracy of DirT. It can be seen that the accuracy of direct transfer in some domains is as high as 84%, such as "clothes". There is also one domain "books" where the accuracy of direct transfer is only 69%. We conjecture that sentiment characteristics of the clothing field are more general, and the performance of other models is better. On the contrary, the sentiment characteristics of the "book" field may be more unique, as it leads to the direct transfer of sentiment classification models in other fields, and the performance is not satisfactory.

From Table 2, we could observe that BTAE and MDAN have greatly improved compared with direct transfer (about 7% ~ 8% accuracy). Our proposed method is superior than all baseline methods, specially about 1% higher than WS-UDA and 2ST-UDA. Furthermore, in the six domains, i.e., tablet, mobile phone, fruit, shampoo, computer and hotel, our method is better than all comparison methods. On the other hand, in the domain of books, the performance of our method is lower than 2ST-UDA about 0.002,

Table 2. experiment result on online_shopping_10cats.

Target	***B***	***T***	***M***	***F***	***S***	***W***	***Me***	***Cl***	***Co***	***H***
DirT	.695	.792	.747	.784	.804	.703	.848	.718	.726	.758
BTAE	.782	.856	.821	.875	.881	.739	.914	.803	.787	.829
MDAN	.784	.868	.814	.873	.879	.738	.915	.794	.807	.830
WS-UDA	.777	.900	.870	.899	.917	.766	**.945**	.811	.829	.857
2ST-UDA	**.822**	.899	.827	.895	.914	**.803**	.941	.769	.824	.855
Ours	.820	**.906**	**.881**	**.902**	**.924**	.766	.944	**.821**	**.843**	**.868**

and similarly in the domains of clothes, it is lower than WS-UDA about 0.001. In Mengniu domain, the performance of our method is comparable to WS-UDA, but lower than 2ST-UDA with nearly 4%.

The experimental results show that compared with WS-UDA, the advantage of our method lies in domain-dependent sentiment knowledge with the weight learning module, and finally carries out additional training based on the generated pseudo labels. Compared with 2ST-UDA, the difference of our method lies in two aspects. The first is the domain-dependent feature extractor in the target domain is pre-trained in the sentiment knowledge learning stage. The second is that the generation of pseudo-label samples are different. The experimental results show that the performance of 2ST-UDA and WS-UDA are mostly the same, which is also true in the original paper. Due the pre-trained domain-dependent feature network in the target domain and the random noise method used in the generation of pseudo labels, our method is superior than these two methods. Besides, it verifies that our method can learn the domain-dependent sentiment knowledge of multiple source domains and the selection of pseudo label samples is more reasonable.

5 Conclusions

In this paper, we propose a novel framework to solve the problem of cross-domain sentiment classification. Specifically, it involves an individual domain-invariant feature extractor and several domain-dependent feature extractors. For the former, we use a reversed discriminator loss to obtain the domain-invariant feature. Furthermore, to reinforce the relationship of domain-dependent features between source and target domains, we propose a weighted learning module to limit these domain-dependent features. Experimental results show the effectiveness of our proposed method. It is also interested to further study how to better infuse domain-invariant and domain-dependent features.

Acknowledgements. This work was supported by the Natural Science Foundation of Guangxi under Grant No.2019GXNSFAA185033, and the Opening Project of Guangdong Province Key Laboratory of Cyber-Physical System (No. 2020B1212060069).

References

1. Blitzer, J., Crammer, K., Kulesza, A., Pereira, F., & Wortman, J. (2007). Learning bounds for domain adaptation. Advances in neural information processing systems, 20.
2. Bollegala, D., Maehara, T., & Kawarabayashi, K. I. (2015). Unsupervised cross-domain word representation learning. arXiv preprint arXiv:1505.07184.
3. Chen, Z., & Qian, T. (2020, July). Relation-aware collaborative learning for unified aspect-based sentiment analysis. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 3685–3694).
4. Dai, Y., Liu, J., Ren, X., & Xu, Z. (2020, April). Adversarial training based multi-source unsupervised domain adaptation for sentiment analysis. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 34, No. 05, pp. 7618–7625).
5. Du, C., Sun, H., Wang, J., Qi, Q., & Liao, J. (2020, July). Adversarial and domain-aware BERT for cross-domain sentiment analysis. In Proceedings of the 58th annual meeting of the Association for Computational Linguistics (pp. 4019–4028).
6. Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., ... & Lempitsky, V. (2016). Domain-adversarial training of neural networks. The journal of machine learning research, 17(1), 2096-2030.
7. Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmospheric environment, 32(14-15), 2627-2636.
8. Ghosal, D., Hazarika, D., Roy, A., Majumder, N., Mihalcea, R., & Poria, S. (2020). Kingdom: Knowledge-guided domain adaptation for sentiment analysis. arXiv preprint arXiv:2005.00791.
9. Glorot, X., & Bengio, Y. (2010, March). Understanding the difficulty of training deep feedforward neural networks. In Proceedings of the thirteenth international conference on artificial intelligence and statistics (pp. 249–256). JMLR Workshop and Conference Proceedings.
10. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.
11. Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2015, April). Long short term memory networks for anomaly detection in time series. In Proceedings (Vol. 89, pp. 89–94).
12. Medsker, L. R., & Jain, L. C. (2001). Recurrent neural networks. Design and Applications, 5, 64-67
13. Nasukawa, T., & Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. In Proceedings of the 2nd international conference on Knowledge capture (pp. 70–77).
14. Ng, A., Jordan, M., & Weiss, Y. (2001). On spectral clustering: Analysis and an algorithm. Advances in neural information processing systems, 14.
15. O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458.
16. Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10), 1345-1359.
17. Pan, S. J., Ni, X., Sun, J. T., Yang, Q., & Chen, Z. (2010, April). Cross-domain sentiment classification via spectral feature alignment. In Proceedings of the 19th international conference on World wide web (pp. 751–760).
18. Peng, H., Xu, L., Bing, L., Huang, F., Lu, W., & Si, L. (2020, April). Knowing what, how and why: A near complete solution for aspect-based sentiment analysis. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 34, No. 05, pp. 8600–8607).

19. Xu, G., Meng, Y., Qiu, X., Yu, Z., & Wu, X. (2019). Sentiment analysis of comment texts based on BiLSTM. Ieee Access, 7, 51522-51532.
20. Xue, Q., Zhang, W., & Zha, H. (2020, April). Improving domain-adapted sentiment classification by deep adversarial mutual learning. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 34, No. 05, pp. 9362–9369).
21. Yang, W., Cao, Z., Chen, Q., Yang, Y., & Yang, G. (2020, November). Confidence Calibration on Multiclass Classification in Medical Imaging. In 2020 IEEE International Conference on Data Mining (ICDM) (pp. 1364–1369). IEEE.
22. Yang, W., & Yang, Y. (2021, July). A Stabilized Dense Network Approach for High-Dimensional Prediction. In 2021 International Joint Conference on Neural Networks (IJCNN) (pp. 1–8). IEEE.
23. Ziser, Y., & Reichart, R. (2016). Neural structural correspondence learning for domain adaptation. arXiv preprint arXiv:1610.01588.