Neurocomputing 494 (2022) 56–66

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Domain adaptation with a shrinkable discrepancy strategy for cross-domain sentiment classification

Yanping Fu [a,b], Yun Liu [a,b,*]

[a] School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China
[b] Key Laboratory of Communication and Information Systems, Beijing Municipal Commission of Education, Beijing, China

ARTICLE INFO

Article history:
Received 6 November 2021
Revised 2 April 2022
Accepted 17 April 2022
Available online 20 April 2022
Communicated by Zidong Wang

Keywords:
Domain adaptation
Sentiment classification
Prototypical information
Cross-domain

ABSTRACT

Cross-domain sentiment classification (CDSC) is used to predict the sentiment polarity of a text in an unlabeled target domain by analyzing the reviews in the labeled source domain. Domain adaptive approaches have become the preferred solution in recent years to the unsupervised domain migration problem. Among them, adversarial learning aligns the sample distribution of the two domains through domain confusion to transfer sentiment across domains. However, traditional adversarial learning often roughly measures domain discrepancy. Although scholars have attempted to adjust the decision boundary of different categories to eliminate the domain shift, such as maximum classifier discrepancy model, there are still two problems with this approach. First, it ignores the intra-domain structure, which causes the samples distributed on the decision boundary to be easily misclassified. Second, it only realizes coarse-grained sentiment migration and lacks a refined evaluation of the transferable information in the inter-domain, which causes a negative transfer. To solve these problems, we propose domain adaptation with a shrinkable discrepancy strategy (DA-SDS) for the task of CDSC. Specifically, we propose to shrink the category subspace in the intra-domain while building the decision boundary of classifiers, which reduces the misclassification by clustering samples to the category center. We also propose to measure the weighted domain discrepancy in the inter-domain, which mitigates the negative transfer through the refined assessment of domain discrepancy. Extensive evaluations showed that DA-SDS outperformed state-of-the-art methods on the Amazon Review dataset.

1. Introduction

As a trending issue in natural language processing tasks, sentiment classification [1] aims to analyze the emotional polarity (positive and negative) of opinions. However, due to the domain discrepancy, the same language may indicate different emotional tendencies in different domains, i.e., "it runs so fast" indicates negative for the *battery* in the electronics domain while positive for the *playmobile* in the toy domain. Thus, a sentiment classifier trained in a specific domain (i.e., source domain) cannot be transferred to a new domain (i.e., target domain). Cross-domain sentiment classification (CDSC) [2,3] mainly predicts the sentiment of unlabeled target data by analyzing labeled source data. Because labeling data is time-consuming and labor-intensive, the CDSC task is a promising research direction.

Domain adaptation [4] mainly aligns the marginal distribution between both domains to bridge the domain gap, which is the classic method for solving unsupervised domain migration problems. In the task of CDSC, unsupervised domain adaptive approaches generally include two categories: pseudo-labeling techniques [5,6] and discrepancy-based methods [7,8]. Among them, domain adversarial training is a discrepancy-based method and obtains excellent performance with simple components in most cross-domain tasks. Therefore, based on the superiority of the algorithm, many studies [9,10] apply adversarial training to eliminate the domain shift between both domains. Its specific design is to construct a sentiment classifier and a domain discriminator to learn the domain-shared features, so that the sentiment classifier trained from the source domain can be transferred to the target domain.

Adversarial adaptation [11] applies domain-adversarial training to confuse the distribution between both domains by maximizing the difference between the domains while minimizing the sentiment classification error. The representative work includes a domain-adversarial neural network (DANN) [12], which proposes a new gradient reversal layer (GRL) to establish a reverse

* Corresponding author at: School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China.
E-mail addresses: 17111012@bjtu.edu.cn (Y. Fu), liuyun@bjtu.edu.cn (Y. Liu).

https://doi.org/10.1016/j.neucom.2022.04.092
0925-2312/

backpropagation that can be applied to various neural network structures to generate the domain-invariant features between both domains. Based on the DANN's innovative work, many studies have adopted adversarial training to build the relationship between the sentiment classifier and domain discriminator. For example, zhang et al. [13] introduced interactive attention transfer network to jointly train domain classifier and sentiment classifier, and used GRL in domain confusion. Zhao et al. [14] proposed parameter transferring and attention sharing mechanisms to transfer sentiment across domain and bridged source and target domains by reducing the cosine distance of attentional weights. Yao et al. [15] introduced DANN-based approach to achieve the domain-adaptation, and extracted domain-relevant information to predict the sentiment polarity of target domain.

Although general domain-adversarial networks can largely reduce the domain discrepancy, they are defective when matching the feature distribution of the source domain with that of the target domain. As shown in Fig. 1 (b), their sentiment classifier can not distinguish samples around the decision boundary. The reasons can be summarized as follows: 1. they only trained the sentiment classifier using labeled source samples but ignored the influence of target samples on the decision boundary of the classifier; and 2. their domain discriminator only vaguely confuses the domain distribution without considering the category of the samples so that they cannot extract discriminative features. To overcome this disadvantage, Saito et al. [16] proposed a novel adaptive method that uses two classifiers to divide task-specific decision boundaries, as shown in Fig. 1 (c). Their approach focuses on the relationship between the target samples and the decision boundary, and combines the information of the target domain to obtain the discriminative features; however, it still faces some problems. For example, it ignores the distribution of samples in the intra-domain, in which samples of different categories may be scattered and overlapped, resulting in misclassification of samples at the decision boundary of the classifier. Moreover, it also lacks a refined assessment method for transferable information between domains, which leads to the negative transfer problem across domain.

To resolve these problems, we propose a domain adaptation with a shrinkable discrepancy strategy (DA-SDS) for CDSC tasks. Based on the prototype network [17], we adopt the prototypical information of the feature to shrink each category subspace in the intra-domain while building the decision boundary of the classifiers. In addition, we also explored discriminative feature learning by dynamically measuring the sample-level domain discrepancy in the inter-domain. As shown in Fig. 1 (d), our proposed shrinkable discrepancy strategy (SDS) not only tightened the sample distribution of the same category to reduce misclassification in the intra-domain, but also refined transferable information in the inter-domain to reduce the negative transfer across domain.

The main highlights of this study are summarized as follows:

(1) We propose a shrinking category subspace strategy to cluster samples to the category center in the intra-domain, which reduces the misclassification caused by samples distributed around the boundary of the classification decision.
(2) We propose a fine-tuning domain discrepancy strategy to automatically capture the discriminative domain-shared features, which mitigates the negative transfer through the refined assessment of domain discrepancy.
(3) We conduct extensive experiments on the Amazon review dataset, and the results demonstrate that our DA-SDS framework can achieve significant performance on unsupervised CDSC tasks.

The remainder of this paper is organized as follows: Section 2 presents the related work; Section 3 introduces the proposed methodology and the details of the model structure; and Section 4 describes the experiment, and in Section 5, we summarize the conclusions and future work.

2. Related Works

In this section, the related work on CDSC approaches and domain adaptation is discussed, and these methods with the proposed DA-SDS model are compared.

CDSC approaches

The CDSC task studies the problem of unsupervised cross-domain sentiment transfer. CDSC approaches can generally be summarized into three categories: word embedding-based methods [18], pivot and non-pivot based methods [19], and domain adaptation-based methods [20].

The method based on word vectors mainly studies the different word representations across domain, and applies the emotional information carried by words to classify the sentiment of reviews. Bollegala et al. [21] proposed a two-step unsupervised approach including SVD smoothing and PLSR prediction to obtain the distribution of words across domains, which also applied prior knowledge with a sentiment lexicon. Hao et al. [22] proposed CrossWord to obtain the connection between domain-specific

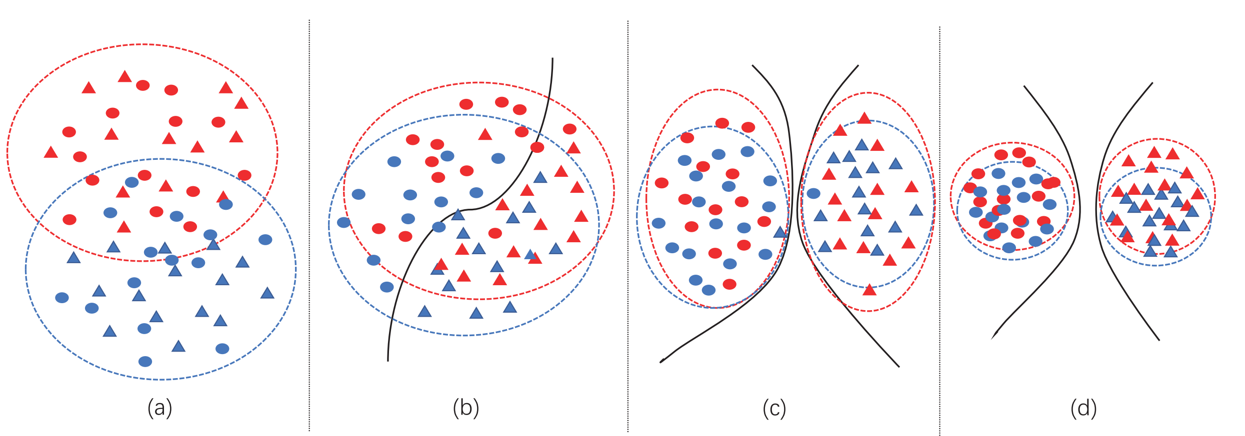


Fig. 1. Comparison of previous and the proposed domain adaptation approach. (Red shapes represent the source data and blue shapes represent the target data. Different shapes indicate different classes. The black line represents the decision boundary of the sentiment classifier. The dotted line represents the range of samples, not the decision boundary.) (a) Sample distribution before domain adaptation. (b) The sample distribution after classic domain adaptation. (c) Saito's domain adaptation, which exploits two classifiers with maximum classifier discrepancy strategy. (d) DA-SDS, our approach obtains intra-class compactness and inter-class separability.

words and pivots by computing their probabilistic similarity. Liu et al. [23] combined sentiment information and domain characteristics to generate domain-aware sentiment word embeddings.

The method based on pivots and non-pivots mainly selects the transferable sentiment words (pivots) and domain-special sentiment words (non-pivots) to generate the domain-shared and domain-private features for the sentiment transfer across domains. Blitzer et al. applied the term frequencies of two domains to generate the pivot [24], and then added mutual information to features and combined supervision information to select more efficient pivots[25]. Recently, deep neutral networks have promoted pivot and non-pivot selection strategies, such as marginalized stacked denoising autoencoder (MSDA) [26], hierarchical attention transfer network (HATN) [27], transferable pivot transformer (TPT) [28] and pre-trained deep contextualized embedding model [29].

In addition to the previous two methods, the domain adaptive method has attracted the attention of scholars. Because cross-domain problems mainly need to bridge the domain gap, domain adaptation is the most studied method in various cross-domain applications. In this study, we adopt the domain-adaptive method to study the sentiment transfer of cross-domain texts. Therefore, we introduce existing domain-adaptive approaches in the following section.

Domain adaptation

Unsupervised domain adaptation (UDA) aims to acquire transferable information by eliminating domain shift, which is widely used in various cross-domain scenarios. According to existing research, UDA is divided into pseudo-labeling techniques and discrepancy-based methods, in which, discrepancy -based methods include maximum mean discrepancy (MMD) [30] and domain adversarial training [31].

Pseudo-labeling techniques mainly apply self-training [32] or tri-training [33] methods to obtain pseudo-labels for unlabeled target data. Saito et al. [34] introduced the generation of pseudo-labels using an asymmetric tri-training method. Rotman et al. [35] explored contextualized word representations to acquire pseudo-labels. Ye et al. proposed self-distillation to reduce the produced label noise by combining self-training with XLMR [36]. The MMD method reduces the domain shift by minimizing the discrepancy in distance between both domains. Since the measurement of divergence determines the quantification of domain shift, most studies focus on the measurement methods of domain distance, such as the Wasserstein distance [37], Jensen-Shannon divergence[38] and Kullback–Leibler (KL) divergency [39].

Domain adversarial adaptation learns domain-shared features by training a sentiment classifier and a domain discriminator in an adversarial manner. Ganin et al. [31] proposed a DANN to augment a classification network with an additional domain classifier. By playing a minimax game [40], adversarial training aims to learn domain-invariant features. Liu et al. [41] proposed a transferable adversarial training method to bridge the gap between the source and target domains. Despite their advantages, traditional methods do not handle the relationship between the target features and the classification decision boundary. Thus, Saito et al. [16] proposed producing task-specific classifiers to align the sample distributions between both domains. Tang et al. [42] explored a mutually inhibitory strategy to generate a relationship between category and prediction results. However, these approaches do not consider the intra-domain structure and only achieve coarse pairwise matching. Fortunately, Li et al. [43] proposed an optimal transport strategy for fine-grained domain adaptation. Based on existing studies, we propose a novel adaptive method that not only focuses on the structure in the intra-domain, but also finely designs the discrepancy between domains in a flexible adversarial training manner.

3. Methodology

3.1. Overall Idea

In the task of CDSC, we are given review data from two domains, one is the labeled source domain data $D^s = \{X^s, Y^s\}$, and the other is the unlabeled target domain data $D^t = \{X^t\}$. Where, $\{X^s, Y^s\} = \{x_i^s, y_i^s\}_{i=1}^{N_s}$ presents N_s labeled samples and $X^t = \{x_i^t\}_{i=1}^{N_t}$ indicates N_t unlabeled samples.

As described by Saito, two task-specific classifiers and a feature generator can handle the relationship between the target samples and the decision boundary. However, their method still does not consider the errors caused by sample dispersion and the domain discrepancy. To overcome these limitations, we propose a DA-SDS approach to perform the sentiment transfer across domains. In our model, we propose a shrinking category subspace strategy and a fine-tuning domain discrepancy strategy to reduce misclassification and nagetive tranfer.

As shown in Fig. 2(a), the sample distributions in the two domains were scattered and unaligned before domain adaptation. Fig. 2(b)-Fig. 2(e) show the process of domain adaptation for the CDSC task. As shown in Fig. 2(b), we first designed two sentiment classifiers C_1 and C_2 that have different characteristics to classify the sentiment polarity of the source domain, and shrink their category subspace to improve the misclassification caused by sample discretization. To eliminate the domain shift, we propose two types of discrepancy evaluation methods: classification discrepancy and domain discrepancy. For the classification discrepancy, we use the maximum and minimum methods [16] to eliminate the misclassification caused by target samples outside the decision boundary. The classification discrepancy region is denoted by black lines in Fig. 2(b), which indicates the target samples outside the support of the source. The classification discrepancy is first maximized to detect these target samples by training two classifiers, as shown in Fig. 2(c). Then, it is minimized to move the target samples inside the support of the source by training the generator, as shown in Fig. 2(d). At the same time, the category subspace of the target domain is shrunk to promote the classification in Fig. 2(d). For the domain discrepancy, we design a weighted optimal transport distance to achieve precise pairwise matching between the source and target domains. And, the domain discrepancy is reduced to align the feature distribution between domains, as shown in Fig. 2(c). Finally, the samples of target domain are well classified by proposed domain adaptation strategy, as shown in Fig. 2(e).

3.2. Model structure

In this section, we introduce the proposed DA-SDS model in detail. As shown in Fig. 3, the model includes a feature generator F and two sentiment classifiers C_1 and C_2, and the parameters of all components are shared in both domains.

Because bidirectional encoder representations from transformers (BERT) can construct the contextual representation using a transformer encoder by discriminating surrounding sentences, we apply the BERT and MLP encoder as the main components of the feature generator. For samples of the source and target domains, their features can be presented as $G_f(X^s)$ and $G_f(X^t)$. These features are then fed into two classifiers for sentiment prediction. The final predictive output is the sum of the two classifiers.

As previously described, our model needs to shrink the classification subspace in the intra-domain, reduce the domain discrepancy, adjust the classification discrepancy, and perform

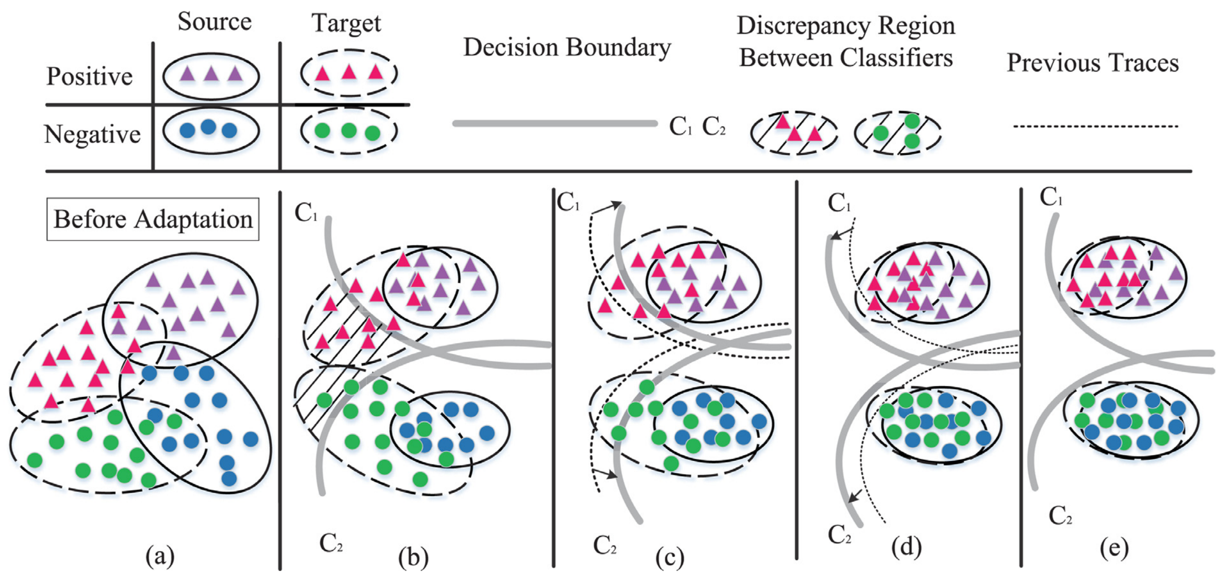


Fig. 2. Domain adaptation process of the proposed DA-SDS method. (a) The distribution of the two domains before adaptation. (b) The distribution after shrinking the subspace of the source data and training the two classifiers with labeled source data. (c) The distribution variation caused by maximizing the discrepancy between two classifiers in the target domain and dynamically reduce the discrepancy between both domains. (d) The distribution variation caused by minimizing the discrepancy between two classifiers in the target domain and shrinking its subspace. (e) The final obtained distribution of our model. (Best viewed in color.).

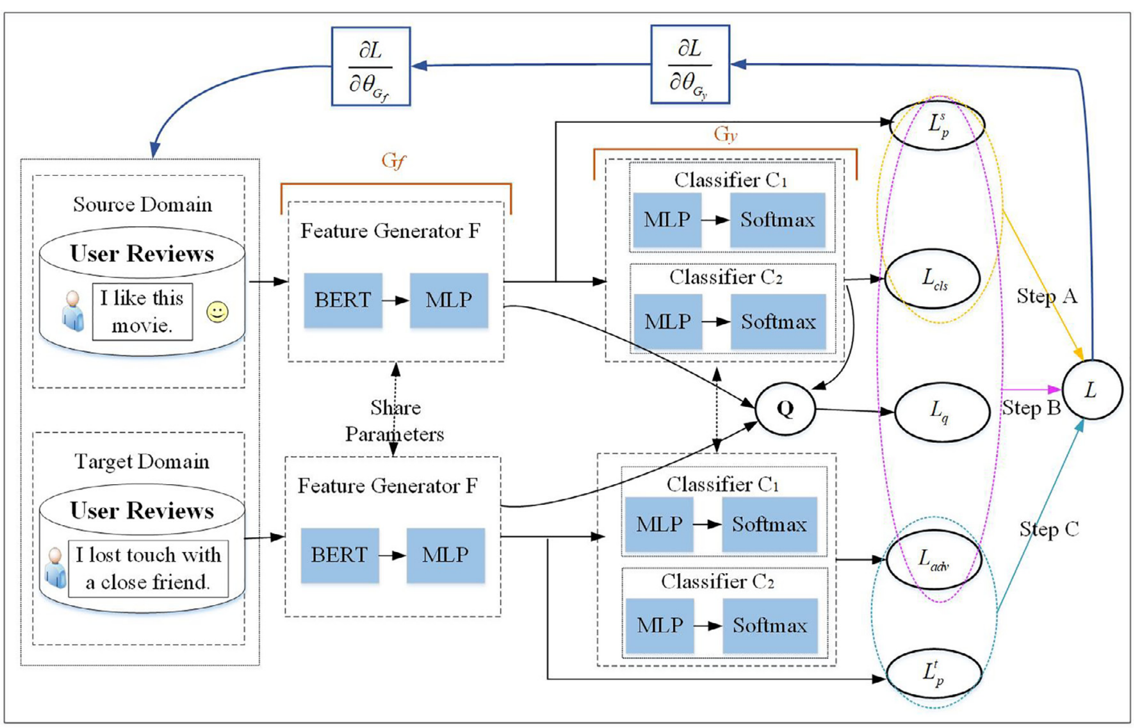


Fig. 3. The architectures of Domain Adaptation with a Shrinkable Discrepancy Strategy (DA-SDS), where G_f is the parameter set of feature generator, G_y is the parameter set of sentiment classifiers; L_{cls} is the standard cross-entropy loss; L_p^s and L_p^t are the proposed shrinking losses, L_q is the weighted domain loss, L_{adv} is the classification loss.

sentiment classification to realize domain adaptation for the CDSC task. When we train the sentiment classifiers, the cross-entropy loss function is used, and it can be presented as follows:

$$L_{cls} = -\frac{1}{N_s}\sum_{i=1}^{N_s}\sum_{j=1}^{K} y_i^s(j) log \hat{y}_i^s(j), \tag{1}$$

where $\hat{y}_i^s \in \{0,1\}$ is the ground truth label in the source domain, and K denotes the number of different polarities.

3.2.1. Shrinking Subspace

The samples in a domain are generally discrete, and those far away from their corresponding category centers are easily misclassified. Therefore, we use the method of shrinking the category subspace to improve it.

To make the samples of the same category as close as possible in the feature space, we design a discriminative centroid loss based on the center loss. However, we have different designs for different domains according to their data characteristics. For the source domain, the center loss is expressed as:

$$L_p^s = \sum_{i=1}^{N_s} \left\| G_f(x_i^s) - c_{y_i^s}^s \right\|_2^2, \tag{2}$$

where $c_{y_i^s}^s$ is the y_i^s-th class center in the source domain, and it can be approximately computed by averaging features of several batch-size samples as:

$$c_{y_i^s}^s = \frac{1}{S} \sum_{i=1}^{N_b} G_f(x_i^s) \phi(y_i^s, k), \tag{3}$$

where $\phi(y_i^s, k) = 1$ if $y_i^s = k$, otherwise $\phi(y_i^s, k) = 0$. $S = \sum_{i=1}^{N_b} \phi(y_i^s, k)$ and $k \in \{1, 2\}$ is sentiment polarity. Here, we apply N_b samples instead of all samples to reduce the complexity of the procedure and we set $N_b = 4 \times n_{batch}$.

For the target domain, the center loss is presented as follows.

$$L_p^t = \sum_{i=1}^{N_t} \left\| G_f(x_i^t) - c_{y_i^t}^t \right\|_2^2, \tag{4}$$

where,

$$c_{y_i^t}^t = \frac{1}{T} \sum_{i=1}^{N_b} G_f(x_i^t) \phi(y_i^t, k). \tag{5}$$

Since the target samples are unlabeled, y_i^t is not given. Based on the fact that the two classifiers can almost accurately classify labeled source samples, we design a pseudo label strategies to determine the value of y_i^t. Thus, we use the trained sentiment classifiers to predict the pseudo labels y_i^t of target domain in this algorithm.

3.2.2. Weighted domain discrepancy

In cross-domain scenarios, the measurement of the domain discrepancy often determines the extent to which the domain shifts. Because the target domain is unlabeled, the general domain discrepancy only calculates the distribution difference between the two domains as a whole, which lacks the emotional polarity information. In fact, eliminating the difference between samples of the same category in the two domains can promote the final sentiment prediction.

To eliminate the domain discrepancy in samples of similar categories between the two domains, we assign the target samples a "pseudo" label by learning from labeled source data. However, the pseudo label is only the result of the prediction, and has uncertainty to the true label. Inspired by previous work [43,44], we evaluate the uncertainty of a target sample i belonging to class k with spatial prototypical information, and the uncertainty value is defined as:

$$Q(i,k) = \frac{d_{A(k)} E(i,k) + (2 - d_{A(k)}) Z(i,k)}{\sum_{m=1}^{K} (d_{A(m)} E(i,m) + (2 - d_{A(m)}) Z(i,m))}, \tag{6}$$

where, E presents the spatial prototypical information and it is defined as:

$$E(i,k) = \frac{e^{-d(G_f(x_i^t), c_k^s)}}{\sum_{m=1}^{K} e^{-d(G_f(x_i^t), c_m^s)}}, \tag{7}$$

where $d(G_f(x_i^t), c_k^s)$ is the distance between the target feature $G_f(x_i^t)$ and k-th source class center c_k^s, and it can be defined as:

$$d(G_f(x_i^t), c_k^s) = H(c_k^s, c_k^s) - 2H(G_f(x_i^t), c_k^s) + H(G_f(x_i^t), G_f(x_i^t)). \tag{8}$$

Here, we select Gaussian kernel function H to compute the distance of two samples.

Since the prediction results of the target domain have certain uncertainties, the following functions are defined in order to express the likelihood of the distribution of the target domain. The sharpen probability function is presented as:

$$Z(i,k) = P(y = k | softmax(\frac{G_y(G_f(x_i^t))}{\tau})). \tag{9}$$

From the above formula, we can see that $E(i,k)$ and $Z(i,k)$ both represent the likelihood that the i-th target domain sample belongs to category k, where $E(i,k)$ is measured by the distance from the target sample feature to the source category center c_k^s, and $Z(i,k)$ is measured by the classifier G_y. Actually, in the early training stage, $E(i,k)$ measured by the distance is more reliable than $Z(i,k)$ estimated by classifier G_y, but it reverses in the later stage. Thus, we use A-distance as the weight to adjust the inportance between $E(i,k)$ and $Z(i,k)$, and it can be defined as follows.

$$d_{A(k)}(E_k^s, E_k^t) = 2(1 - 2\epsilon(h_k)), \tag{10}$$

where $\epsilon(h_k)$ represents the error of a linear SVM classifier h_k to classify two domains.

In the early training stage of the domain adaptation, the feature distributions of two domains are expected to be quite different in the feature space, then the SVM should obtain perfect classification performance, which means error $\epsilon(h_k) \to 0$ and $d_{A(k)} \to 2$. It leads that the second term of Eq. 6 vanishes, and the first term $E(i,k)$ weight the uncertainty of domain discrepancy. With domain adaptive training, the distributions of two domains coincide with each other, and the SVM cannot differentiate features between two domains. Thus, $\epsilon(h_k) = 0.5$ and $d_{A(k)} = 0$. It leads that the first term of Eq. 6 vanishes, and the second term $Z(i,k)$ weight the uncertainty of domain discrepancy. Based on these well-founded reasons, we define the weighted domain discrepancy as follows.

$$L_q = \sum_{k=1}^{K} \sum_{i=1}^{n_t} Q(i,k) \left\| G_f(x_i^t) - c_k^s \right\|_2^2, \tag{11}$$

where c_k^s is the k-th class center in the source domain.

3.2.3. Classification discrepancy

The classification disrepancy aims to move target samples outside the support of the source to the inside of the support of the source. The maximizing and minimizing classification disrepancy method is applied in this study. we maximize the discrepancy to effectively detect target samples outside the support of the source by training classifiers (C_1 and C_2). Without this step, the two classifiers can be very similar ones and cannot detect target samples outside the support of the source. We then minimize the discrepancy by training the generator to fool the classifiers.

We use the absolute value of the output difference between two sentiment classifiers as the classification disrepancy of target domain. The loss function is defined as:

$$L_{adv} = \frac{1}{K} \sum_{k=1}^{K} \sum_{i=1}^{n_t} |p_{1k}(x_i^t) - p_{2k}(x_i^t)|, \tag{12}$$

where the p_{1k} and p_{1k} denote probability outputs of classifier C_1 and classifier C_2 for class k respectively.

3.2.4. Adaptation process

In order to realize the ideas in Section 3.1, we specifically present the domain adaptation process of the proposed model. The generator and classifiers first should classify the samples of source domain correctly, and then the weighted domain discrepancy and the classification discrepancy are utilized to dynamically align the features between both domains. The adaptation process is performed in three steps.

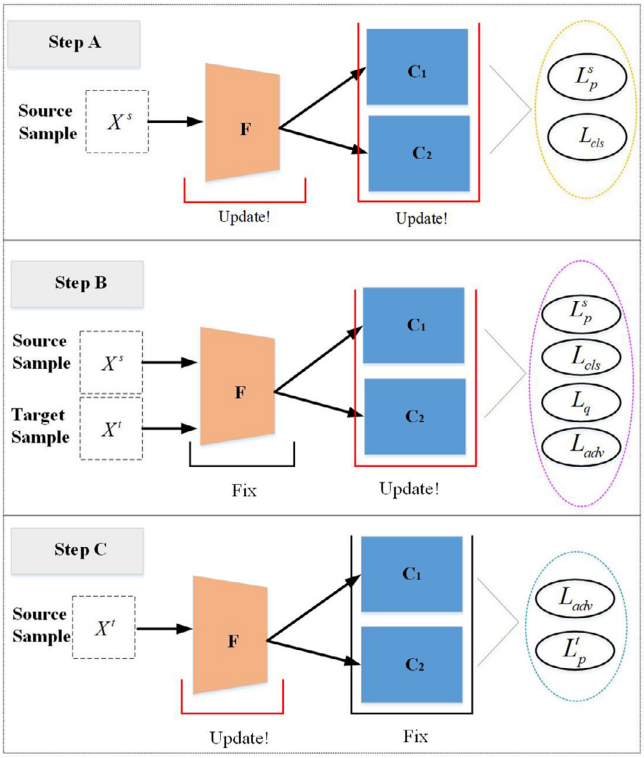


Fig. 4. Adversarial training steps of our method.

Step A The generator and classifiers are trained to indicate the sentiment poalrity of source domain by labeled source data. In this step, we also shrink the category subspace of the source domain to enhance the classification effect. As Fig. 4 shown, the parameters of both generator and classifier should be trained in step A. The objective is presented as follows.

$$\min_{F,C_1,C_2} L_{cls} + \alpha_1 L_p^s. \tag{13}$$

Step B We mainly apply the classification discrepancy to identify the target samples excluded by the support of the source and domain discrepancy to align features between both domains. To ensure the basic conditions for the source domain samples to be correctly classified, we add the loss function of the step A into step B. Thus, the objective of step B is presented as:

$$\min_{C_1,C_2} L_{cls} + \alpha_1 L_p^s + \beta L_q - \gamma L_{adv}, \tag{14}$$

where β and γ are hyperparameters, they and α_1 are used to balance different loss functions. Note that in this step we fixed the feature generator to train only the parameters of the classifiers.

Step C In step C, the classification discrepancy is reduced so that the target samples outside the decision boundary can be supported by the source data and target samples is shrunk closer to the source sample. And the feature generator is trained to minimize the classification discrepancy for fixed classifiers. As Fig. 4 shown, the objective is presented as:

$$\min_{F} L_{adv} + \alpha_2 L_p^t. \tag{15}$$

In our approach, these three steps need to be repeated and their order is not important.

4. Experiment

4.1. Datasets

Our experiments apply the Amazon Review dataset, which is widely used in CDSC task. It includes reviews from four domains: Books (B), DVDs (D), Electronics (E) and Kitchen appliances (K). Each domain includes 2000 labeled data, of which 1000 are positive and 1000 are negative. The relevant statistical information is shown in Table 1.

The CDSC task is to realize the sentiment transfer from the labeled domain to the unlabeled domain. Therefore, for these four domains, we performed 12 cross-domain tasks to verify the effectiveness of the algorithm. For each task, we apply a 5-fold cross-validation to get the best results.

4.2. Parameter Settings

In the experiments, we apply BERT$_{base}$ as the basic component of context feature generator. The maximum sequence length is 256 in BERT. We apply Adam as the adaptive optimizer and the initial learning rate is 10-3e. The number of batch size is 64 and the dropout rate is 0.5. The weights of these loss functions can vary across different domains and related study on its effect on the prediction performance will be shown later through a sensitivity analysis.

4.3. Comparison methods

We present some comparison methods according to previous studies and describe their basic design for subsequent comparison.

(1) **Non-adaptation**. The non-adaptation method does not utilize any adaptation strategy in our model. Specially, this method does not consider domain discrepancy and classification discrepancy.
(2) **DANN** [12]. DANN uses domain-adversarial neural networks to provide transferable shared information for cross-domain tasks.
(3) **HATN** [2]. HATN utilizes hierarchical attention network to achieve sentiment transfer, which mainly extract pivot and non-pivot features to assist classification task.
(4) **HATN (BERT)** [9]. It replaces the main feature extractor of HATN with BERT.
(5) **IATN** [13]. It combines sentence and aspect attention learning mechanisms with an interactive attention transfer mechanism to obtain domian-shared features across domains.
(6) **TAT (mSDA)** [41]. It applies the adversarial training method to eliminate domain shift and generates transferable examples from source domain to target domain. And it used mSDA [45] representations as input of words.
(7) **CapsuleDAR** [3]. It integrates semantic rules into Capsule network to represent text and uses CORAL loss to obtain domain-shared features.

Table 1
Statistics of the Amazon dataset. "Voc." represents the number of vocabulary. "Avg-len of doc." represents the average number of words in a review.

Domain	classes	Positive	Negative	Voc.	Avg-len of doc.
Books	2	1000	1000	26278	158.6
DVD	2	1000	1000	26940	171.5
Electronics	2	1000	1000	13256	102.9
Kitchen	2	1000	1000	11187	87.3

(8) **BERT-DAAT** [9]. It explores post-training methods that includes the masked task of target domain and a reconstructed domain-distinguish task to complete domain adaptation with BERT model.
(9) **WTN** [7]. It uses BERT to obtain the contextual embedding and minimizes the Wasserstein distance to eliminate differences between domains.
(10) **PERL** [29]. It extends contextualized embedding models such as BERT with pivot-based fine-tuning to solve the problem of domain adaptation for CDSC task.
(11) **PTASM (BERT)** [14]. It combines parameter transferring method and attention sharing mechanism to transfer the sentiment across domain, and uses BERT to encode context features.
(12) **PTASM (Glove)** [14]. It applies Glove word embedding to replace pre-train BERT embedding to encode text for PTASM model.
(13) **DA-SDS (Glove)**. The proposed DA-SDS method uses Glove word embeddings and the feature extractor of HATN to replace BERT extractor.
(14) **DA-SDS (BERT)**. The proposed DA-SDS method uses pre-train BERT to generate text features.

4.4. Evaluation Criteria

According to previous studies, we apply accuracy rate as the evaluation standard for a comparison between the proposed method and the benchmark method.

4.5. Experimental results and analyses

In this subsection, we present experimental results and analyses. First, the results were compared with other benchmark experiments. Then, we show the visualization of features to observe the sample distributions of the two domains in the process of the proposed domain adaptation. Subsequently, we conducted ablation studies to verify the importance of the different components. Finally, the sensitive analysis on the model was provided.

4.5.1. Comparison with Existing Approaches

Because the current model can be divided into two types of benchmark models with and without BERT, we conducted comparative experiments with the two benchmark models. The comparison results between baselines without BERT and the DA-SDS model are displayed in Table 2, and those between baselines with BERT and the DA-SDS model are displayed in Table 3.

In Table 2, we first perform experiments with the non-adaptation strategy, which means that the proposed DA-SDS model does not use shrinking subspace, domain discrepancy, or classification discrepancy. Compared with the non-adaptation method, the other adaptation methods obtained better results. This proves that appropriate domain adaptation strategies have an absolutely positive effect on the emotional transfer of CDSC tasks. For example, compared with the non-adaptation method, the domain adaptation methods (DANN, HATN, IATN, TAT (mSDA), PTASM (Glove), DA-SDS (Glove) and DA-SDS (BERT)) improved the average accuracy of the CDSC by 0.1054, 0.1158, 0.1086, 0.1131, 0.1271, 0.1301, and 0.1649, respectively.

To compare with the baseline model without BERT, we apply Glove word embeddings and the feature extractor of HATN to replace BERT, and the results are given in the DA-SDS (Glove) method in Table 2. Compared with other baseline methods (DANN, HATN, IATN, TAT (mSDA), PTASM (Glove)), our DA-SDS (Glove) improved the average accuracy of CDSC by 0.0247, 0.0143, 0.0215, 0.0170 and 0.0030, respectively. When using the BERT structure, our DA-SDS (BERT) improved the average accuracy by 0.0595, 0.0491, 0.0563, 0.0518 and 0.0378, respectively. These results prove that our adaptive strategy has more advantages for emotional transfer than the baseline methods. The effect of our proposed DA-SDS (BERT) model is much better than that of DA-SDS (Glove), which also proves that the BERT structure could significantly improve classification performance in the CDSC task.

In Table 3, we compare the results of the CDSC tasks between DA-SDS and the transfer methods with BERT. The HATN (BERT), BERT-DAAT, WTN, PERL, PTASM (BERT) and DA-SDS (BERT) methods combine BERT with the downstream task model, which has the advantages of large-scale pre-training models. As shown in Table 3, these methods perform well with results of 0.8869, 0.9012, 0.9040, 0.8750, 0.9110, and 0.9148 on the Amazon dataset, respectively. Compared with other baseline methods (HATN (BERT), BERT-DAAT, WTN, PERL, PTASM (BERT)), our DA-SDS (BERT) improved the average accuracy of CDSC by 0.0279, 0.0136, 0.0108, 0.0398 and 0.0038, respectively. In these methods, our DA-SDS (BERT) method obtains the best average accuracy, which proves that the proposed method can reduce negative transfer and effectively achieve cross-domain sentiment transfer.

4.5.2. Visualization of Features

To intuitively evaluate the impact of shrinking subspace and weighted domain discrepancy on the proposed DA-SDS model, we present a visualization of features of the variants of DA-SDS for sample distribution of two domains. Fig. 5 shows the sample distribution of four models for the B → D CDSC task. In Fig. 5, we apply t-SNE to map features of source and target data points into a two-dimensional plane.

Table 2
Accuracies results in CDSC tasks between DA-SDS and transferring methods without BERT. (The literature subscript indicates the source of comparison results, and the best results are in bold.)

Task	Non-adaptation	DANN [2]	HATN [2]	IATN [13]	TAT mSDA [41]	PTASM (Glove) [14]	DA-SDS (Glove)	DA-SDS (BERT)
B → D	0.7510	0.8342	0.8707	0.8680	0.8680	0.8820	0.8831	**0.9128**
B → E	0.7340	0.7627	0.8575	0.8650	0.8590	0.8640	0.8648	**0.9022**
B → K	0.7310	0.7790	0.8703	0.8590	0.8860	0.8710	0.8722	**0.9066**
D → B	0.7510	0.8077	0.8778	0.8700	0.8640	0.8830	0.8835	**0.9118**
D → E	0.7550	0.7635	0.8632	0.8690	0.8640	0.8790	0.8795	**0.9208**
D → K	0.7630	0.7815	0.8747	0.8580	0.8940	0.8780	0.8791	**0.9115**
E → B	0.7210	0.7353	0.8403	0.8180	0.8370	0.8620	0.8638	**0.9145**
E → D	0.7430	0.7627	0.8432	0.8410	0.8350	0.8750	0.8760	**0.9088**
E → K	0.7880	0.8453	0.9008	0.8870	0.9040	0.8830	0.8971	**0.9328**
K → B	0.7210	0.7417	0.8488	0.8470	0.8140	0.8710	0.8813	**0.9223**
K → D	0.7430	0.7532	0.8472	0.8440	0.8470	0.8790	0.8822	**0.9140**
K → E	0.7980	0.8553	0.8933	0.8760	0.8920	0.8960	0.8976	**0.9198**
Average	0.7499	0.8553	0.8657	0.8585	0.8630	0.8770	0.8800	**0.9148**

Table 3
Accuracies results in CDSC tasks between DA-SDS and transferring methods with BERT. (The literature subscript indicates the source of comparison results, and the best results are in bold.)

Task	HATN (BERT) [9]	BERT-DAAT [9]	WTN [7]	PERL [29]	PTASM(BERT) [14]	DA-SDS (BERT)
B → D	0.8936	0.8970	0.9090	0.8780	0.9120	**0.9128**
B → E	0.8721	0.8957	0.8840	0.8720	0.9010	**0.9022**
B → K	0.8941	0.9075	0.8960	0.9020	0.9060	**0.9066**
D → B	0.8981	0.9086	0.9080	0.8560	0.8990	**0.9118**
D → E	0.8699	0.8930	0.9150	0.8930	0.9110	**0.9208**
D → K	0.8759	0.9050	0.8910	0.9040	0.9080	**0.9115**
E → B	0.8710	0.8891	0.9010	0.8390	0.9140	**0.9145**
E → D	0.8881	0.9013	0.8920	0.8480	0.9070	**0.9088**
E → K	0.9201	0.9318	0.9320	0.9120	0.9170	**0.9328**
K → B	0.8788	0.8798	0.9160	0.8300	0.9210	**0.9223**
K → D	0.8789	0.8881	0.8890	0.8560	**0.9140**	**0.9140**
K → E	0.9031	0.9172	0.9190	0.9120	0.9190	**0.9198**
Average	0.8869	0.9012	0.9040	0.8750	0.9110	**0.9148**

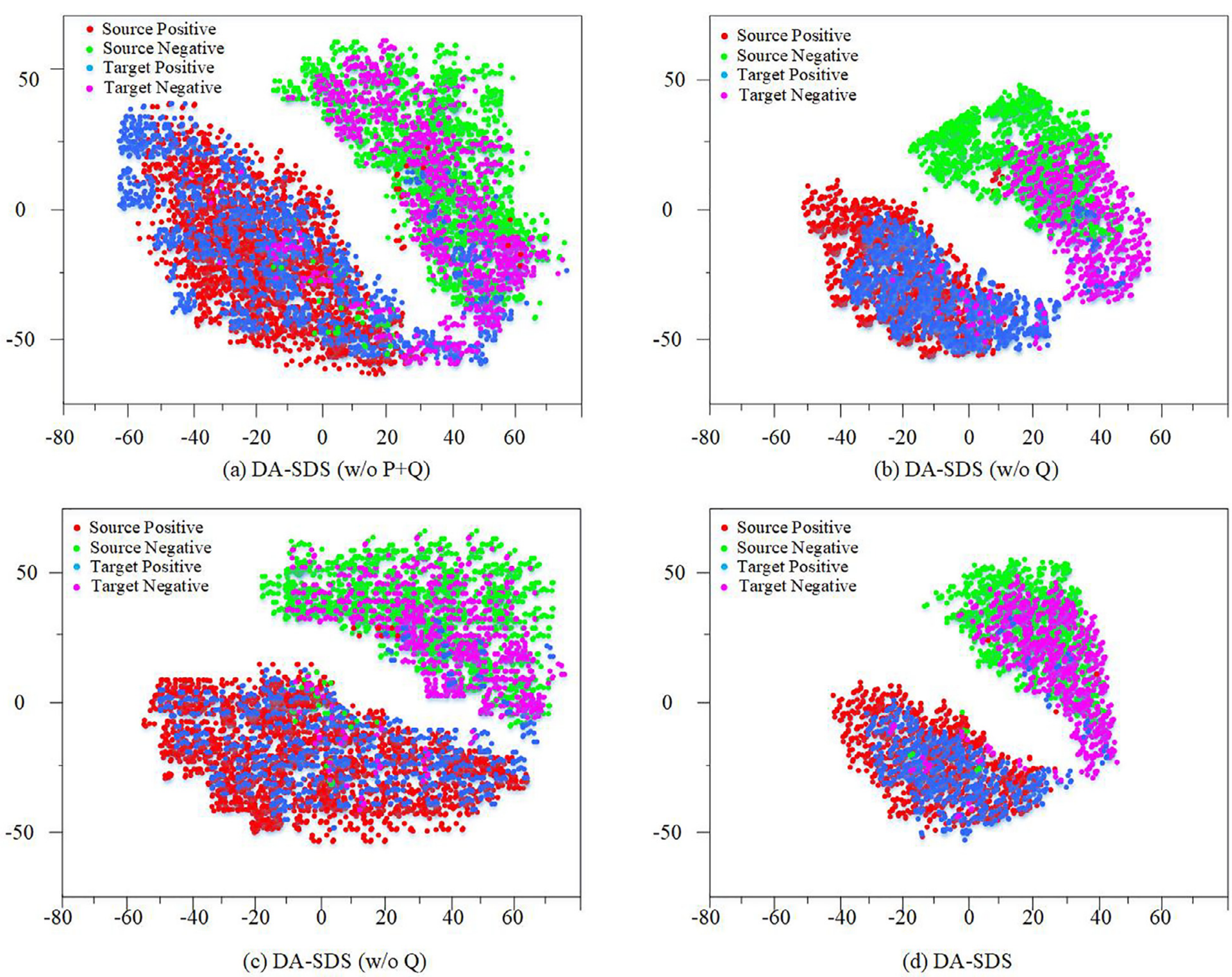


Fig. 5. The effect of shrinking subspace and weighted domain discrepancy on the distribution of the extracted features. The figure shows t-SNE visualization of the DA-SDS's features for the B → D task.

We compare the feature distributions of different models after domain adaptation in Fig. 5, where (DA-SDS w/o P + Q) indicates the variant of the proposed model in which DA-SDS does not use shrinking subspace and weighted domain discrepancy; (DA-SDS w/o P) indicates that DA-SDS does not use shrinking subspace; (DA-SDS w/o Q) indicates that DA-SDS does not use weighted domain discrepancy. From Fig. 5 (a), we observe that samples of different polarities in the source domain are well classified, while for the target domain, some samples of different polarities are misclassified and decision boundaries are not clear to the samples of the target domain. Compared with Fig. 5 (a), the shrinking subspace in Fig. 5 (b) increases the sample compactness and improves the sample dispersion, and makes the decision boundaries clearer; the weighted domain discrepancy in Fig. 5 (c) causes better alignment of the features between both domains, and reduces the number of misclassified samples in the target domain. In Fig. 5 (d), the proposed DA-SDS model not only obtains clear classification boundaries, but also the sample features of different domains

Table 4
Classification accuracy of DA-SDS without some components on the Amazon reviews dataset.

Task	DA-SDS	DA-SDS (w/o P + Q)	DA-SDS (w/o P)	DA-SDS (w/o Q)
B → D	0.9128	0.8893 (-0.0235)	0.9011 (-0.0017)	0.8986 (-0.0142)
B → E	0.9022	0.8796 (-0.0226)	0.8903 (-0.0119)	0.8865 (-0.0157)
B → K	0.9066	0.8804 (-0.0262)	0.8933 (-0.0133)	0.8901 (-0.0165)
D → B	0.9118	0.8901 (-0.0217)	0.9001 (-0.0117)	0.8982 (-0.0136)
D → E	0.9208	0.8906 (-0.0302)	0.9108 (-0.0100)	0.9068 (-0.0140)
D → K	0.9115	0.8895 (-0.0220)	0.9043 (-0.0072)	0.8994 (-0.0121)
E → B	0.9145	0.8787 (-0.0358)	0.9021 (-0.0124)	0.8946 (-0.0199)
E → D	0.9088	0.8802 (-0.0286)	0.8993 (-0.0095)	0.8867 (-0.0221)
E → K	0.9328	0.9033 (-0.0295)	0.9208 (-0.0120)	0.9134 (-0.0194)
K → B	0.9223	0.8966 (-0.0257)	0.9116 (-0.0107)	0.9038 (-0.0185)
K → D	0.9140	0.8993 (-0.0147)	0.9028 (-0.0112)	0.9001 (-0.0139)
K → E	0.9198	0.8890 (-0.0308)	0.8905 (-0.0293)	0.9076(-0.0122)
Average	0.9148	0.8889 (-0.0259)	0.9022 (-0.0125)	0.8988 (-0.0160)

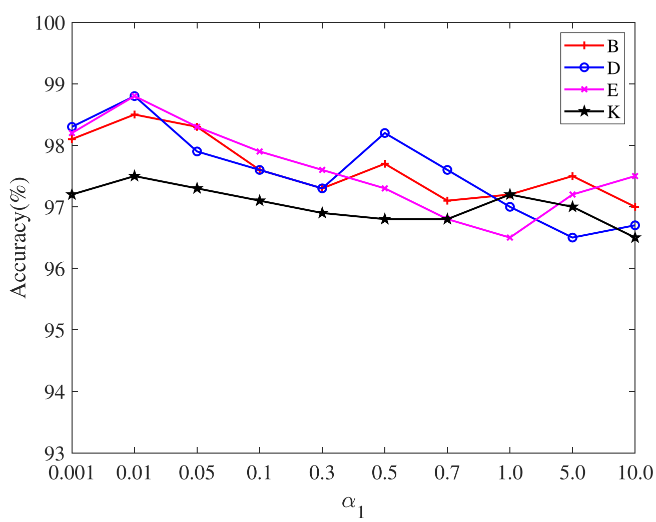


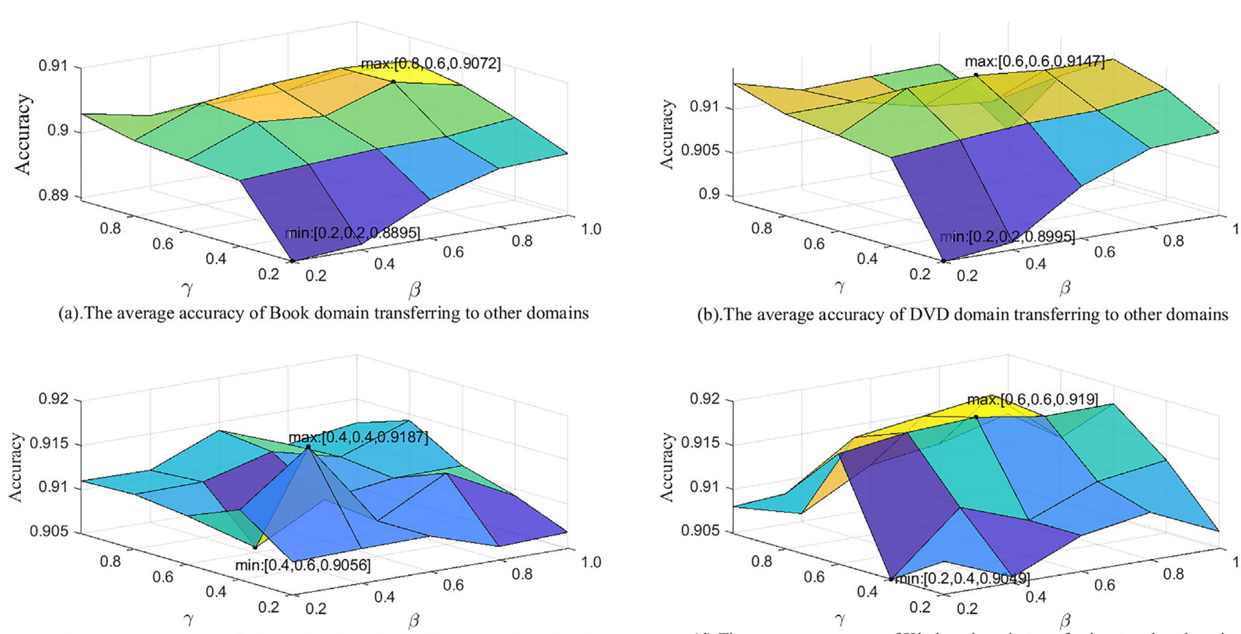
Fig. 6. Effects of transfer loss parameters α_1 on classification performance of a domain in step A.

become compact and aligned. These graphs show that shrinking subspace and weighted domain discrepancy can reduce negative transfer and increase the transferability of features across domains.

4.5.3. Ablation Studies

To explore the effect of different components, including shrinking subspace tasks and weighted domain discrepancy steps, we performed some ablation experiments, and the experimental results are given in Table 4. As defined in the previous section, w/o stands for without, P indicates shrinking subspace tasks, and Q indicates weighted domain discrepancy steps.

In Table 4, we present the classification accuracy of DA-SDS and its three variations, where the blue font indicates the extent to the accuracy was reduced compared to the original DA-SDS model. From Table 4, we observe that when one or both of the components were deleted from the proposed DA-SDS model, the effect of CDSC tasks is correspondingly reduced. When both components were removed, the average accuracy of the DA-SDS model was reduced by 2.59%, which is the most serious decline in all ablation experiments. When the weighted domain discrepancy steps were



(a).The average accuracy of Book domain transferring to other domains

(b).The average accuracy of DVD domain transferring to other domains

(c).The average accuracy of Electronics domain transferring to other domains

(d).The average accuracy of Kitchen domain transferring to other domains

Fig. 7. Effects of transfer loss parameters γ and β in DA-SDS model.

removed, the average accuracy was reduced by 1.60%. When the shrinking subspace tasks were deleted, the average accuracy wss reduced by 1.25%. This shows DA-SDS (w/o Q) has a larger decrease than DA-SDS (w/o Q).

These results demonstrate that shrinking subspace tasks and weighted domain discrepancy steps are both beneficial to CDSC tasks, and weighted domain discrepancy is more efficient than shrinking subspace to align cross-domain features and achieve sentiment transfer. In addition, the structure of DA-SDS (w/o P + Q) is consistent with the Saito's domain adaptation model [16]. The comparison results of DA-SDS and DA-SDS (w/o P + Q) show that the proposed two components can effectively improve the accuracy of cross domain sentiment classification. Thus, the shrinking subspace tasks and weighted domain discrepancy steps proposed in this study are sufficiently effective for CDSC tasks.

4.5.4. Sensitive Analysis

In this work, we define four loss weights α_1, γ, β and α_2 to balance different loss functions. We first analyze the influence of parameter α_1 on the classification performance of a labeled domain from the step A. As shown in Fig. 6, the classification accuracies of four domains are variable when α_1 = [0.001, 0.01, 0.05, 0.1, 0.3, 0.5, 0.7, 1.0, 5.0, 10.0]. From Fig. 6, we observe that the results of four domains are consistent, and they all get the highest accuracy when the parameter α_1 is set to 0.01. Based on this experience, we also set the parameter α_1 = 0.01 in the step B. In our experiment, because α_2 is also a parameter of shrinking loss, we set $\alpha_2 = \alpha_1$.

To analyze the influence of parameters γ, β on CDSC tasks, we present the average results of cross-domain tasks with four histograms to find the optimal combination of parameters in Fig. 7. For each domain, We construct three domain pairs and calculate their average accuracy as the result to find the optimal values of the parameters γ and β. We set $\gamma = [0.2, 0.4, 0.6, 0.8, 1.0]$ and $\beta = [0.2, 0.4, 0.6, 0.8, 1.0]$ to test the accuracy for various combinations of γ and β. From Fig. 7, we find that when the source domains are Book, DVD, Electronics, and Kitchen, the optimal parameter settings are (γ = 0.8, β = 0.6), (γ = 0.6, β = 0.6), (γ = 0.4, β = 0.4), and (γ = 0.6, β = 0.6) respectively.

5. Conclusions and future works

A CDSC approach based on DA-SDS was proposed in this study. It adjusts the intra-domain and inter-domain structures to reduce negative transfer with shrinking subspace tasks and weighted domain discrepancy steps. The shrinking loss is used to gather samples around the boundary of the classification decision to the center of the category, and the weighted domain loss is presented to assess the differences between domains. This method can realize effective sentiment transfer for cross-domain reviews. Comprehensive experiments prove that our DA-SDS model outperforms state-of-the-art results on Amazon datasets, and ablation studies prove that both of our proposed strategies can improve the performance of CDSC tasks.

Further, we plan to apply our proposed model to multi-source cross-domain sentiment analysis and explore domain adaptation methods for computer vision.

CRediT authorship contribution statement

Yanping Fu: Conceptualization, Methodology, Software, Visualization, Writing - original draft, Writing - review & editing. **Yun Liu:** Data curation, Investigation, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research was funded by the Fundamental Research Funds for the Central Universities (Grant No. 2020YJS012).

References

[1] B. Liu, Sentiment analysis and opinion mining, Synthesis lectures on human language technologies 5 (2012) 1–167.
[2] Z. Li, Y. Wei, Y. Zhang, Q. Yang, Hierarchical attention transfer network for cross-domain sentiment classification, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 32, 2018.
[3] B. Zhang, X. Xu, M. Yang, X. Chen, Y. Ye, Cross-domain sentiment classification by capsule network with semantic rules, IEEE Access 6 (2018) 58284–58294.
[4] R. Li, Q. Jiao, W. Cao, H.-S. Wong, S. Wu, Model adaptation: Unsupervised domain adaptation without source data, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 9641–9650.
[5] J. Liang, R. He, Z. Sun, T. Tan, Exploring uncertainty in pseudo-label guided unsupervised domain adaptation, Pattern Recognition 96 (2019) 106996.
[6] J. Liang, R. He, Z. Sun, T. Tan, Exploring uncertainty in pseudo-label guided unsupervised domain adaptation, Pattern Recognition 96 (2019) 106996.
[7] Y. Du, M. He, L. Wang, H. Zhang, Wasserstein based transfer network for cross-domain sentiment classification, Knowledge-Based Systems 204 (2020) 106162.
[8] G. Kang, L. Jiang, Y. Yang, A.G. Hauptmann, Contrastive adaptation network for unsupervised domain adaptation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 4893–4902.
[9] C. Du, H. Sun, J. Wang, Q. Qi, J. Liao, Adversarial and domain-aware bert for cross-domain sentiment analysis, in: Proceedings of the 58th annual meeting of the Association for Computational Linguistics, 2020, pp. 4019–4028.
[10] K. Bousmalis, N. Silberman, D. Dohan, D. Erhan, D. Krishnan, Unsupervised pixel-level domain adaptation with generative adversarial networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 3722–3731.
[11] M. Long, Z. Cao, J. Wang, M.I. Jordan, Conditional adversarial domain adaptation, arXiv preprint arXiv:1705.10667 (2017).
[12] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky, Domain-adversarial training of neural networks, The Journal of Machine Learning Research 17 (2016), 2096–2030.
[13] K. Zhang, H. Zhang, Q. Liu, H. Zhao, H. Zhu, E. Chen, Interactive attention transfer network for cross-domain sentiment classification, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 33, 2019, pp. 5773–5780.
[14] C. Zhao, S. Wang, D. Li, X. Liu, X. Yang, J. Liu, Cross-domain sentiment classification via parameter transferring and attention sharing mechanism, Information Sciences 578 (2021) 281–296.
[15] F. Yao, Y. Wang, Domain-specific sentiment analysis for tweets during hurricanes (dssa-h): A domain-adversarial neural-network-based approach, Computers, Environment and Urban Systems 83 (2020) 101522.
[16] K. Saito, K. Watanabe, Y. Ushiku, T. Harada, Maximum classifier discrepancy for unsupervised domain adaptation, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 3723–3732.
[17] Y. Pan, T. Yao, Y. Li, Y. Wang, C.-W. Ngo, T. Mei, Transferrable prototypical networks for unsupervised domain adaptation, in: in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 2239–2247.
[18] D. Bollegala, T. Mu, J.Y. Goulermas, Cross-domain sentiment classification using sentiment sensitive embeddings, IEEE Transactions on Knowledge and Data Engineering 28 (2015) 398–410.
[19] Y. Fu, Y. Liu, Cross-domain sentiment classification based on key pivot and non-pivot extraction, Knowledge-Based Systems 228 (2021) 107280.
[20] Z. Li, Y. Zhang, Y. Wei, Y. Wu, Q. Yang, End-to-end adversarial memory network for cross-domain sentiment classification, in: IJCAI, 2017, pp. 2237–2243.
[21] D. Bollegala, D. Weir, J.A. Carroll, Learning to predict distributions of words across domains, in: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2014, pp. 613–623.
[22] Y. Hao, T. Mu, R. Hong, M. Wang, X. Liu, J.Y. Goulermas, Cross-domain sentiment encoding through stochastic word embedding, IEEE Transactions on Knowledge and Data Engineering 32 (2019) 1909–1922.
[23] J. Liu, S. Zheng, G. Xu, M. Lin, Cross-domain sentiment aware word embeddings for review sentiment analysis, International Journal of Machine Learning and Cybernetics 12 (2021) 343–354.
[24] J. Blitzer, R. McDonald, F. Pereira, Domain adaptation with structural correspondence learning, in: Proc. 2006 conference on empirical methods in natural language processing, 2006, pp. 120–128.

[25] J. Blitzer, M. Dredze, F. Pereira, Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification, in: Proc. 45th annual meeting of the association of computational linguistics, 2007, pp. 440–447.
[26] Y. Ganin, V. Lempitsky, Unsupervised domain adaptation by backpropagation, in: Proc. International conference on machine learning, 2015, pp. 1180–1189.
[27] Z. Li, Y. Wei, Y. Zhang, Q. Yang, Hierarchical attention transfer network for cross-domain sentiment classification, in: Proc. AAAI Conference on Artificial Intelligence, volume 32, 2018.
[28] L. Li, W. Ye, M. Long, Y. Tang, J. Xu, J. Wang, Simultaneous learning of pivots and representations for cross-domain sentiment classification., in: Proc. AAAI, 2020, pp. 8220–8227.
[29] E. Ben-David, C. Rabinovitz, R. Reichart, Perl: Pivot-based domain adaptation for pre-trained deep contextualized embedding models, Transactions of the Association for, Computational Linguistics 8 (2020) 504–521.
[30] A. Gretton, D. Sejdinovic, H. Strathmann, S. Balakrishnan, M. Pontil, K. Fukumizu, B.K. Sriperumbudur, Optimal kernel choice for large-scale two-sample tests, in: Advances in neural information processing systems, Citeseer, 2012, pp. 1205–1213.
[31] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, V. Lempitsky, Domain-adversarial training of neural networks, The Journal of Machine Learning Research 17 (2016), 2096–2030.
[32] S. Abney, Semisupervised learning for computational linguistics, CRC Press, 2007.
[33] A. Søgaard, Simple semi-supervised training of part-of-speech taggers, in: Proceedings of the ACL 2010 Conference Short Papers, 2010, pp. 205–208.
[34] K. Saito, Y. Ushiku, T. Harada, Asymmetric tri-training for unsupervised domain adaptation, in: International Conference on Machine Learning, PMLR, 2017, pp. 2988–2997.
[35] G. Rotman, R. Reichart, Deep contextualized self-training for low resource dependency parsing, Transactions of the Association for, Computational Linguistics 7 (2019) 695–713.
[36] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, V. Stoyanov, Unsupervised cross-lingual representation learning at scale, arXiv preprint arXiv:1911.02116 (2019).
[37] J. Shen, Y. Qu, W. Zhang, Y. Yu, Wasserstein distance guided representation learning for domain adaptation, in: Thirty-Second AAAI Conference on Artificial Intelligence, 2018.
[38] Z. Sun, Y. Hu, L. Yang, S. Lu, J. Mei, Y. Han, X. Li, Stc-nas: Fast neural architecture search with source-target consistency, Neurocomputing (2021).
[39] Z. Meng, J. Li, Y. Gong, B.-H. Juang, Adversarial teacher-student learning for unsupervised domain adaptation, in: 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2018, pp. 5949–5953.
[40] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, Advances in Neural Information Processing Systems 27 (2014).
[41] H. Liu, M. Long, J. Wang, M. Jordan, Transferable adversarial training: A general approach to adapting deep classifiers, in: International Conference on Machine Learning PMLR, 2019, pp. 4013–4022.
[42] H. Tang, K. Jia, Discriminative adversarial domain adaptation, in: Proceedings of the AAAI Conference on Artificial Intelligence, volume 34, 2020, pp. 5940–5947.
[43] R. Xu, P. Liu, L. Wang, C. Chen, J. Wang, Reliable weighted optimal transport for unsupervised domain adaptation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 4394–4403.
[44] Y. Li, Z. Yang, Y. Wang, C. Xu, Adapting neural architectures between domains, Advances in Neural Information Processing Systems 33 (2020).
[45] M. Chen, Z. Xu, K. Weinberger, F. Sha, Marginalized denoising autoencoders for domain adaptation, arXiv preprint arXiv:1206.4683 (2012).

Yanping Fu received the Master's degree in July 2015 from the School of Information Science & Engineering, Lanzhou University, Lanzhou, China. Since 2017, she has been a Ph.D. candidate at the School of Electronic and Information Engineering, Beijing JiaoTong University, Beijing, China. Her current research interests include natural language processing, data mining, cross-domain sentiment analysis and deep learning.

Yun Liu is a Professor of Communication and Information Systems, Beijing Jiaotong University; Dean of Communication Engineering Department, Beijing Jiaotong University; Director of key Laboratory of Communication and Information Systems, Beijing Municipal Commission of Education; Director of Institute of Network Consensus Security, Beijing Jiaotong University; Vice-chair of Teachers and Staff Representative Committee, Beijing Jiaotong University. She is currently a Fellow of IET, UK and specialist enjoying special government allowance. In addition, she is an evaluation expert of State Scientific and Technological reward, State Natural Sciences Fund in communication, National High Technology Research and Development Program (HTRDP), an advanced counselor of China Tietong Company, and an advanced counselor of Beijing Municipal Office of Internet Propaganda and Management.