Source-free Domain Adaptation for Aspect-based Sentiment Analysis

Zishuo Zhao[1], Ziyang Ma[1], Zhenzhou Lin[1], Jingyou Xie[1]
Yinghui Li[2,*], Ying Shen[1,*]
[1]School of Intelligent Systems Engineering, Sun Yat-Sen University
[2]Tsinghua Shenzhen International Graduate School, Tsinghua University
{zhaozsh,mazy23,linzhzh6,xiejy73}@mail2.sysu.edu.cn

Abstract

Unsupervised Domain Adaptation (UDA) of the Aspect-based Sentiment Analysis (ABSA) task aims to transfer knowledge learned from labeled source domain datasets to unlabeled target domains on the assumption that samples from the source domain are freely accessible during the training period. However, this assumption can easily lead to privacy invasion issues in real-world applications, especially when the source data involves privacy-preserving domains such as healthcare and finance. In this paper, we introduce the **S**ource-**F**ree Domain Adaptation Framework for **ABSA** (SF-ABSA), which only allows model parameter transfer, not data transfer, between different domains. Specifically, the proposed SF-ABSA framework consists of two parts, i.e., feature-based adaptation and pseudo-label-based adaptation. First, in feature-based adaptation, we embed the model into the feature space of the target domain through dependency relation prediction; and then transfer this feature-embedded model to the source domain. Second, we use labeled data in the source domain to obtain a well-trained source model. Finally, in pseudo-label-based adaptation, we utilize the pseudo labels predicted by the source model to obtain category anchor points, and then reassign the pseudo labels of the target domain according to the distance between the anchor points and the target data to obtain higher-quality pseudo labels for domain adaptation. Experiment results on four benchmarks show that the proposed framework performs competitively with traditional unsupervised domain adaptation methods under the premise of insufficient information, which demonstrates the superiority of our method under privacy conditions.

Keywords: Source-Free, Domain Adaptation, Aspect-based Sentiment Analysis

1. Introduction

Aspect-based sentiment analysis (ABSA) is a data mining technique that involves aspect extraction and aspect sentiment classification subtasks (Qiu et al., 2011; Poria et al., 2016; Li et al., 2018a; Wang et al., 2018; Li et al., 2019c). For instance, in the sentence "The AMD Turin Processor seems to always perform much better than Intel", we detect two aspect terms "AMD Turin Processor" and "Intel", and then classify their sentiment polarities as "Positive" and "Negative", respectively. Recently some researchers focus on End to End ABSA(E2E-ABSA), which combines the above two subtasks in an end-to-end manner and uses a unified tagging scheme to connect the two tasks into one task (Dredze et al., 2010; Xia et al., 2014; He et al., 2018; Ye et al., 2020), as illustrated in Figure 1.The unified tagging scheme turns the E2E-ABSA task into a sequence labeling task by aggregating aspect boundary tags (e.g., B, I, O denotes the beginning of, inside of, and no aspect term) and sentiment polarities (e.g. POS, NEG, NEU denotes positive, negative and neutral sentiment). Citing the same example as above, "AMD Turin Processor" and "Intel" should be tagged with B-POS, I-POS, I-POS and B-NEG, while the remaining words are tagged with O.

In previous studies, researchers have noticed that the ABSA task relies heavily on large-scale labeled data for supervised learning. Therefore, the End2End ABSA task in the field of Unsupervised Domain Adaptation (UDA) has recently attracted the attention of researchers. Domain adaptation trains a robust model in a label-rich domain and transfers the learned knowledge to a label-poor domain, which can well alleviate the label-lacking problem in the ABSA task. Many methods have been proposed, which can be broadly classified into two categories: feature-based domain adaptation (Ganin et al., 2016; Li et al., 2018b) and instance-based domain adaptation (Dredze et al., 2010; Xia et al., 2014).

Despite the effectiveness of these existing studies, domain adaptation remains a challenge in real-world applications since UDA relies heavily on source domain data. Considering that real-world data are distributed across different devices and often contain private information (e.g., data on personal phones or surveillance cameras), existing UDA methods require access to source domain data during the learning process, which may violate data privacy. To tackle such an issue, recent

*Corresponding author.

LREC-COLING 2024, pages 15076–15086
20-25 May, 2024.

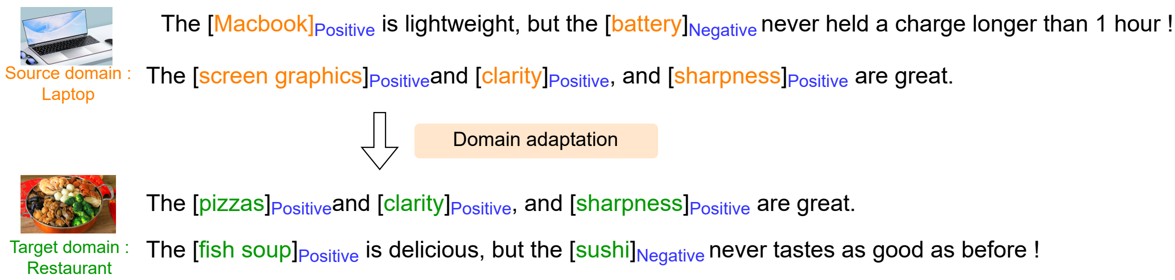

Figure 1: An overview of the Cross-Domain End-to-End ABSA task, in which the goal is to migrate source domain data to the distribution of target domain data through UDA domain adaptation method.

studies propose an interesting but challenging UDA setting called source-free UDA.

The source-free UDA has recently been explored in the field of computer vision, but its development in the field of natural language processing is still relatively under-explored. Laparra et al. (2021) designed the SemEval 2021 Task 10 where labeled source data is not accessible, only models trained on source domain data can be shared, and little or no labeled target data is available. SemEval 2021 Task 10 is the only work that explicitly tests source-free domain adaptation in NLP. A variety of techniques are applied to this task, including active learning, self-training, and data augmentation (Su et al., 2022). Active learning requires high-quality labels, with strong constraints in real-world scenarios. In the absence of source data and target labels, self-training is the most anticipated study since it can make good use of source domain models.

Despite the effectiveness of these existing methods, the source-free UDA remains some challenges: (1) Feature distribution shifts between different domains. For example, "battery" is an aspect term in the laptop domain but rarely in the service domain. The model needs some latent information to judge their potential similarity. (2) High-quality pseudo-labels. Due to the lack of labels in the target domain, we usually need to generate pseudo-labels that better represent the data distribution of the target domain, which is challenging.

In light of these challenges, we propose a Source-free domain adaption framework for ABSA (SF-ABSA). Specifically, the proposed SF-ABSA framework consists of two parts: Feature-based adaptation and Pseudo-label based adaptation. First, in feature-based adaptation, we transfer the model to the feature space of the target domain via dependency relation prediction, and then transfer the feature-embedded model to the source domain. Second, we use the labeled data in the source domain to obtain a well-trained source model. Finally, in pseudo-label based adaptation, we leverage self-supervised pseudo-label annotations to reassign the pseudo-labels predicted by the source model, and enable domain adaptation by computing the distance between anchors and target data.

The main contributions can be summarized as follows:

- To the best of our knowledge, we are the first to explore domain adaptation for the ABSA task without access to source domain data, which has great significance for protecting privacy.
- We propose the SF-ABSA framework, encompassing both feature-based adaptation and pseudo-label-based adaptation, which can transfer knowledge using only the parameters of the model.
- Experimental results show that our framework achieves competitive results compared to the state-of-the-art results on ABSA for unsupervised domain adaptation.

2. Related work

2.1. Aspect Based Sentiment Analysis

Aspect term extraction (ATE) and Aspect-level sentiment classification (ASC) are two subtasks in Aspect Based Sentiment Analysis.

Aspect term extraction. Aspect term extraction is a fundamental task of ABSA, aiming to identify specific terms or phrases within a sentence that people express opinions about. Traditional machine learning techniques have been used for ATE such as Support Vector Machines (SVM) and Conditional Random Fields (CRF). These models usually use lexical, syntactic, and contextual features (Manek et al., 2017; Xiang et al., 2018). With the advent of transformer-based models, particularly BERT (Bidirectional Encoder Representations from Transformers) and its variants such as RoBERTa and ALBERT, there has been a shift towards using these pre-trained models for ATE. Fine-tuning these models on domain-specific data has proven to be effective for aspect term extraction (Devlin et al., 2018; Liu et al., 2019). Considering the labor-intensive and costly nature of data annotation, unsupervised methods and semi-supervised methods have been explored for ATE. For instance, bootstrapping techniques and self-training are commonly applied (Qiu et al., 2011; Poria et al., 2016; Li et al., 2018a; Wang and Pan, 2018).

Aspect-level sentiment classification. Aspect-level sentiment classification aims to predict the sentiment polarity of the extracted aspect terms. We can divide its method into three parts. First is the method based on deep learning. With the advances in neural networks, models like Transformer, Bert and RoBERTa have been proposed and brought large performance improvements (Tang et al., 2016; Wang et al., 2018; Li et al., 2019c). Secondly, some studies have specifically explored the modeling of a sentence's syntactic structure for prediction, given that the relationship between aspects and their associated opinions, as determined by their structural relations, frequently suggests emotional orientation. Various methods exist for incorporating dependency and syntactic information into ASC tasks, as exemplified by (Kiritchenko et al., 2014; Brun et al., 2014). Tiredly, some works have explored enhancing ASC by integrating external knowledge bases (Chen and Huang, 2019; Zhou et al., 2020; Yu et al., 2022).

2.2. Conventional Domain Adaptation

Since fine-grained sentiment polarity annotation of sentences is costly, it is impossible to have sufficient labeled data in each new domain. In the domain adaptation field, researchers usually unify two subtasks into one end-to-end training task, which is generally called End-to-End Aspect-Based Sentiment Analysis (E2E-ABSA). Existing research on E2E-ABSA for domain adaptation mainly focuses on the coarse-grained domain adaptation problem to learn domain-invariant representations, including semi-supervised methods (He et al., 2018; Ye et al., 2020), domain adversarial networks (Ganin et al., 2016; Li et al., 2018b), and pivot-based methods (Wang et al., 2016). Another line of work focuses on re-weighting source instances (Dredze et al., 2010; Xia et al., 2014).

2.3. Source-free Domain Adaptation.

In the field of computer vision, the source-free setting has attracted the attention of researchers, and many source-free UDA methods have emerged. Liang et al. (2020) utilizes information maximization via a pseudo-labeling strategy to adapt the trained source model to target features. Xia et al. (2021) introduces a new target classifier to align two domains (Source domain similar domains and dissimilar domains) via adversarial training manner. Ding et al. (2022) utilizes the Gaussian distribution assumption, uses the feature of the target domain data to approximate the distribution of the source domain data, and solves the problem of inaccessible source domain data.

The application of source-free domain adaption in natural language processing is still relatively limited. Though there is partially related research on continual learning (de Masson D'Autume et al., 2019; Sun et al., 2019) and generalization of pretrained models (Hendrycks et al., 2020). Laparra et al. (2021) designed the SemEval 2021 Task 10 dataset on two tasks, i.e., negation detection and time expression recognition. Su et al. (2022) extended self-training, active learning and data augmentation baseline.

3. Task Definition

The E2E-ABSA task includes two subtasks, i.e., Aspect extraction and Aspect sentiment classification. We connect the two tasks by viewing them as a sequence tagging task with a set of unified labels. Assuming that we have a sentence $x = \{w_1, w_2, w_3, ..., w_T\}$, we get its embedding $e = \{e_1, e_2, e_3, ..., e_T\}$ by extracting the features. The goal of the task is to predict the label $y = \{y_1, y_2, y_3, ..., y_T\}$ with y_i corresponding to each word and satisfying that $y_i \in \Psi^u = \{$B-POS, I-POS, B-NEG, I-NEG, B-NEU, I-NEU, O$\}$.

In this paper, we focus on Source-free Unsupervised Domain Adaptation, where source domain data cannot be obtained, only model parameters trained on the source domain can be obtained, and meanwhile labeled data are not available in the target domain. Given a set of unlabeled instances from a target domain $D_u = \{w_i^u\}_{i=1}^{N_U}$, and model M_i that can be transmitted between different domains, our goal is to predict token labels for target test instances: $y_i^t = M_{final}(w_i^t)$.

4. Methods

4.1. Overview

Our task is an End2End ABSA task that connects two ABSA subtasks (Aspect extraction and Aspect sentiment classification) in an end-to-end manner. Our model framework is based on Aspect Based Sentiment Analysis datasets of different distributions. We let D_s denote source domain data and D_t denote target domain data ($D_s \neq D_t$). In order to complete the task under the premise of protecting data privacy, we only allow the transfer of model parameters between the source domain and the target domain, not allowing the transfer of data. On this basis, we eliminate the difference between the target domain and the source domain from two aspects:

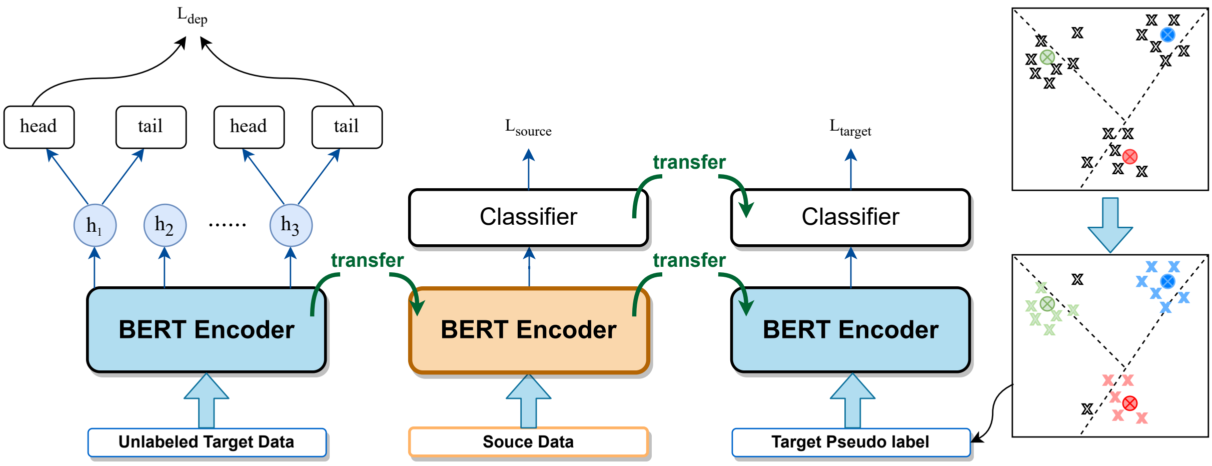

Figure 2: An overview of the proposed SF-ABSA framework. SF-ABSA framework consists of two parts: feature-based adaptation and pseudo-label-based adaptation. First, in feature-based adaptation, we embed the model into the feature space of the target domain through dependency relation prediction, and then transfer this feature-embedded model to the source domain. Second, we use labeled data in the source domain to obtain a well-trained source model. Finally, in pseudo-label based adaptation, we utilize the pseudo labels predicted by the source model to obtain category anchor points, and then reassign the pseudo labels of the target domain according to the distance between the anchor points and the target data to obtain higher-quality pseudo labels for domain adaptation.

(1) **Feature-Based Domain Adaptation.** With domain-shared knowledge and using Mask Language Model (MLM) task, we make the model embed the target domain data into feature space in advance. Then we transport the feature-embedded model to the source domain to train on source domain data.

(2) **Domain Adaptation with Self-supervised Pseudo-labeling.** After we obtain the trained source domain model, we utilize the source domain model to predict the labels and get the feature vectors from the hidden layer of the model. We use the labels and corresponding features as anchor points to reassign the labels of the target domain data according to the distance between the anchor points and the data instances.

We train the above two parts in order and finally get our SF-ABSA framework, as illustrated in Figure 2.

4.2. Feature-Based Domain Adaptation

Feature extraction. We first convert the word sequence $\{w_1, w_2, w_3, \ldots, w_T\}$ into a continuous embedding $\{e_1, e_2, e_3, \ldots, e_T\}$. The embedding of each word is composed of four different types of embeddings $e_i = \{t_i, s_i, p_i, tag_i\}$, where t_i is the word embedding of the word w_i, s_i is segment embedding, p_i is positional embedding and tag_i is POS tag embedding. The first three embeddings are the input of Bidirectional Encoder Representations from Transformers (BERT) defined by (Devlin et al., 2018). The fourth one, POS tag embedding, is a randomly initialized matrix trained by unlabeled data in the target domain. The embedding $e_i = \{e_1, e_2, e_3, \ldots, e_T\}$ is put into BERT to extract context-aware features $H_i = \{H_1, H_2, H_3, ..., H_T\}$.

$$H = BERT(e). \tag{1}$$

Dependency Relation Prediction. The syntactic dependency is very helpful for the ABSA task, which is capable to indicate the dependency between words and words in sentences, thus helping us to identify the aspect term in the sentence. In the ABSA task in the field of domain transfer, although the aspect terms in different domains are different and the opinion words are also different, they all follow the same grammatical rules. Therefore, We can regard syntactic dependencies as domain-independent macroscopic features that are important for passive unsupervised domain transfer tasks. So we embed the dependency features of the target domain data into the feature space of the model. We put the above context-aware feature H_i into two different nonlinear transformation functions, with $H_{head} = h_1^{head}, h_2^{head}, h_3^{head}, ..., h_T^{head}$ and $H_{tail} = h_1^{tail}, h_2^{tail}, h_3^{tail}, ..., h_T^{tail}$ being the two outputs respectively.

$$h_i^{head} = tanh(W_1 h_i + b_1), \tag{2}$$

$$h_i^{tail} = tanh(W_2 h_i + b_2), \tag{3}$$

where W_1 and W_2 are learnable parameters, h_i^{head} and h_i^{tail} represent the head token and child token in the dependency tree. Assuming that the i-th word and j-th word in a sentence are connected in the dependency tree, and they are the head token and child token in the dependency relationship respectively, we use m_{ij} to predict the type of their dependency relationship:

$$m_{ij} = [h_i^{head}; h_j^{tail}; h_i^{head} - h_j^{tail}; h_i^{head} \cdot h_j^{tail}], \quad (4)$$

where the semicolon indicates concatenation operation, the minus sign indicates element-wise subtraction and the dot indicates element-wise dot multiplication. Then m_{ij} is converted to p_{ij}^{dep}:

$$p_{ij}^{dep} = softmax(W_{dep} m_{ij} + b_{dep}), \quad (5)$$

where W_{dep} is the weight matrix for relation classification, and p_{ij}^{dep} represents the probabilities of all the forty-seven classes of dependency relationships. In the dependency tree, if two words are connected, we will predict their dependency relationship, and we will not predict the words that are not connected. So we use $\mathbb{I}(ij)$ to express whether the i-th word is connected to the j-th word. It can be considered that $\mathbb{I}(ij)$ is an indicator function, if it is connected in the dependency tree, it is 1, otherwise, it is 0. The optimization function is as follows:

$$L_{dep} = \sum_{X^t} \sum_{i}^{T} \sum_{j}^{T} \mathbb{I}(ij) CrossEntropy(p_{ij}^{dep}, y_{ij}^{dep}), \quad (6)$$

where X_t is unlabeled data in the target domain, T is the maximum length of sentences and y_{ij}^{dep} is the actual dependency relationship. Based on the above optimization model, we embed the model into the feature space of the target domain data and obtain the model M_0.

4.3. Source Model Generation

After the feature embedding in Section 4.2, we obtain the model M_0 embedded in the feature space of the target domain data. We transfer the model from the target domain to the location of the source domain and use this model as the basic model for source domain training. Then we train the model on the source domain data by minimizing the cross-entropy where the source domain is located:

$$L_{source} = CrossEntropy(M_0(x_s), y_{label}), \quad (7)$$

where x_s is the source domain data and y^{label} is the ABSA label corresponding to the source domain data.

After training at the location of the source domain, we obtain the well-trained source model M_s. We transfer the source model M_s to the target domain location for the domain transfer process of the target domain.

4.4. Domain Adaptation with Self-supervised Pseudo-labeling

In many recent works, pseudo-labeling is an important technique to obtain category information of unlabeled samples, usually by exploiting high-confidence outputs derived from models trained on the source domain. However, due to the domain shift, the pseudo-labels obtained by predicting the target domain data through the source model usually have a lot of noise. For example, in the E2E-ABSA field, when the model is trained in the source domain and predicts samples in the target domain, due to different domain keywords (e.g., "battery" in the laptop field and "E-trade" in the service field), the model The output is usually unable to predict the answer very accurately, such as [0.36,0.32,0.01,0.02], this kind of input is usually simply classified as [1,0,0,0], but the original unconfident prediction contains more syntactic level information that cannot be captured. So how to obtain higher quality pseudo-labels is a critical issue. We use the idea of clustering to propose the idea of using the class center anchor to reassign labels to solve this problem.

Through the training in Section 4.3, we obtained the source domain model M_s that was transferred from the location of the source domain. The source model mainly includes two parts, feature extractor (BERT) F_s and classifier (Classifier) C_s. The feature extractor F_s is used to extract the corresponding feature distribution of the sentence and the classifier C_s is used to predict and classify each word in the sentence according to The unified tagging scheme. First, we use the trained source model to predict the unlabeled data of the target domain, and obtain the preliminary sample category information:

$$y^t = argmax(\delta_k(M_s(x_t))), \quad (8)$$

where y^t is the pseudo-label predicted by the source model M_s, $\delta_k(a) = \frac{exp(a_k)}{\Sigma_i exp(a_i)}$ is softmax function and x_t represents target domain data. However, this pseudo-label usually contains much noise caused by domain shift. Therefore, based on the predicted category, we calculate the category center point of each category and use it as the anchor point feature of the category:

$$A_k^{(0)} = \frac{\Sigma_{x_t \in X_t} \delta_k(M_s(x_t)) * F_s(x_t)}{\Sigma_{x_t \in X_t} \delta_k(M_s(x_t))}, \quad (9)$$

where $M_s(x_t) = C_s(F_s(x_t))$, and $F_s(x_t)$ outputs the feature extracted from target domain data. We weight the output confidence to the sample features and then average the features of the same category, and use it as the category center feature of this category. This can characterize the feature distribution of different categories of the target domain more stably and reliably. We then reassign

Dataset	Domain	Sentences	Training	Testing
L	Laptop	3,845	3,045	800
R	Restaurant	6,035	3,877	2,158
D	Device	3,836	2,557	1,279
S	Service	2,239	1,492	747

Table 1: Statistics of adopted datasets.

the labels of the target domain samples by computing the distance of each feature to the class center feature:

$$\hat{y}^t = arg \min_k Dist(A_k^{(0)}, F_s(x_t)), \tag{10}$$

where $Dist(a, b) = \frac{1}{2}(1 - \frac{a^T b}{|a| \cdot |b|})$ is cosine distance, and $\hat{y}^t$ is the reassigned label. Then we use the new pseudo-label to re-find the category center feature of the target domain, and iterate as such:

$$A_k^{(m)} = \frac{\Sigma_{x_t \in X_t} \mathbb{I}(\hat{y}^t = k) * F_s(x_t)}{\Sigma_{x_t \in X_t} \mathbb{I}(\hat{y}^t = k)}, \tag{11}$$

$$\hat{y}^t = arg \min_k Dist(A_k^{(m)}, F_s(x_t)), \tag{12}$$

where m represents the current iteration times and $\mathbb{I}(\cdot)$ is an indicator function. The iteration will stop when the anchor point of the target domain feature converges. When the iteration stops, we utilize a confidence threshold $\tau \in (0, 1)$ to filter out lower-quality answers to improve the quality of the pseudo-labels.

$$D_t' = \{(x_i^t, \hat{y}^t) | Dist(F_s(x_i^t), A_{\hat{y}_i^t}) < \tau\}_{i=1}^{n_t}. \tag{13}$$

We take the final (x^t, y^t) as the final high-quality labels, and use these labels as the data of the target domain to perform the domain adaptation process on the source model M_s. In the above formula (11), we update the category features of the target domain. In our experiment, the performance of the model is the best when the number of iteratively updating category feature anchors is 1.

5. Experiment

5.1. Datasets

We conduct experiments on four benchmark datasets: Laptop (L), Restaurant (R), Device (D) and Service (S). Restaurant (R) is the union set of the restaurant datasets from SemEval ABSA challenge 2014, 2015 and 2016 (Pontiki et al., 2014, 2015, 2016). Laptop (L) containing user reviews from the laptop domain, it is from SemEval-2014 (Pontiki et al., 2014) ABSA challenge. Device (D) is a combination of device reviews from 5 different digital products (Toprak et al., 2010). While similar in nature to the Laptop dataset, this dataset would encompass a broader range of electronic or mechanical devices. Service (S) contains reviews from web services, which is introduced by (Hu and Liu, 2004). Detailed statistics are shown in Table 1.

5.2. Experimental settings

we perform our experiments on 10 source→target pairs based on the four datasets mentioned above. Each of these four datasets is used as the source domain and remains are used as the target domain for knowledge transfer. For each source→target transfer experiment, our source domain data is containing the sentences and their corresponding sentiment labels, and the target domain data is not containing the corresponding sentiment labels. We followed the setting of (Li et al., 2019b) and removed D→L and L→D because the datasets of L and D are very similar. Due to our setting for data privacy protection, which only allows model parameter transfer, not data transfer, between different domains. In other words, the source and target domain data cannot appear in the training period of the model at the same time. The evaluation of the model is based on the test datasets of the target domain data. The dependency analysis of the target domain data is performed using Spacy, with a total of 47 dependencies between words and words.

In previous unsupervised domain adaptation for ABSA tasks, the selection of the base model is pre-trained model BERT$_B$ and extended version BERT$_E$. For BERT$_B$, it refers to the uncased BERT$_{Base}$ model pre-trained by (Devlin et al., 2018). For BERT$_E$, it refers to an extended version of BERT$_{Base}$ model from (Li et al., 2019a), which fine-tuning the pre-trained BERT$_B$ model on product reviews from a combination of Yelp Challenge datasets and the Electronics datasets from Amazon. Since BERT$_E$ incorporates domain knowledge in advance, it is often better to use BERT$_E$ as the base model. However, in order to compare more fairly with some other methods, and make our SF-ABSA framework can be extended to more NLP tasks, we do not use BERT$_E$ as our base model.

The GTX 3090 was used for this experiment. The batch size is set to 32, and the learning rate is set to $5 \cdot 10^{-5}$. Due to the size of the datasets, the epoch was set to 1.

The evaluation metric is Micro-F1. We also used precision and recall as references in the experiment to analyze the effect of our framework, but we did not put all of them into the result table. We continue to use the evaluation criteria of the previous task of unsupervised domain adaptation, only exact match could be counted as correct.

Methods	R→S	L→S	D→S	S→R	L→R	D→R	S→L	R→D	AVG
Source-Only									
BERT-Base (Devlin et al., 2018)	19.48	25.78	30.31	42.2	40.38	30.06	29.20	29.47	30.36
Source-Required									
Hier-Joint (Ding et al., 2017)	15.56	13.90	19.04	31.10	33.54	32.87	22.65	24.53	23.71
RNSCN (Wang et al., 2018)	20.04	16.59	20.03	33.21	35.65	34.60	18.87	33.26	26.09
AD-SAL (Li et al., 2019a)	28.01	27.20	26.62	41.03	43.04	41.01	27.04	35.44	33.71
BERT-DANN (Gong et al., 2020)	21.60	25.10	18.62	45.84	41.73	34.68	30.47	34.41	30.83
BERT-UDA (Gong et al., 2020)	33.12	27.89	28.03	47.09	45.46	42.68	**34.77**	34.93	35.98
CDRG (Indep) (Yu et al., 2021)	34.10	33.97	31.08	44.46	44.96	39.42	26.81	25.25	34.27
CDRG (Merge) (Yu et al., 2021)	35.14	**38.14**	37.22	47.92	**49.79**	**47.64**	33.69	27.46	**38.98**
Source-Free									
SF-ABSA (Feature-based only)	26.44	26.05	31.83	**48.78**	40.72	41.16	34.33	**36.64**	35.87
SF-ABSA (All)	**35.67**	29.62	**45.93**	44.62	44.23	35.43	34.01	28.56	37.14

Table 2: Comparison results of different methods for Cross-Domain End-to-End ABSA based on Micro-F1.The table consists of three parts: (1) Source-Only; (2) Source-Required; (3) Source-Free.

Methods	R→S	L→S	D→S
Source-Required			
BERT-DANN (Gong et al., 2020)	21.60	25.10	18.62
BERT-UDA (Gong et al., 2020)	33.12	27.89	28.03
CDRG (Yu et al., 2021)	35.14	**38.14**	37.22
GCDDA (Li et al., 2022)	32.07	27.22	28.52
Source-Free			
SF-ABSA(our)	**35.67**	29.62	**45.93**

Table 3: Comparison results of different methods for Cross-Domain End-to-End ABSA, which utilizing the service dataset as the transfer target.

5.3. Baselines

We compare with several highly competitive DA methods as follows:

- Hier-Joint (Ding et al., 2017): A recurrent neural network with syntactic rule-based auxiliary tasks for cross-domain AE.
- RNSCN (Wang et al., 2018): A recursive neural structural correspondence network based on syntactic structures and auto-encoders.
- AD-SAL (Li et al., 2019a): A Selective Adversarial Learning method to achieve local semantic alignments for fine-grained domain adaptation.
- BERT-base (Devlin et al., 2018): indicates directly fine-tuning BERT-base-uncased model on the source training data.
- BERT-DANNN: We respectively use BERT-base as the base model, and simultaneously perform adversarial training on each word, which can be viewed as the BERT version of the widely used DANN approach proposed by (Ganin et al., 2016).
- BERTB-UDA (Gong et al., 2020): our recent unified feature and instance-based domain adaptation method based on BERT-base, respectively.
- CDRG (Yu et al., 2021): they generate the target domain reviews with independent and merge training strategies.
- GCDDA (Li et al., 2022): a generative framework for cross-domain data augmentation, utilizing a pre-trained BART sequence-to-sequence model, designed to synthesize target-domain data accompanied by detailed annotations..

We are the first study to consider data privacy issues in the cross-domain E2E-ABSA task. In this paper, we select conventional UDA ABSA methods as baselines, although none of them consider the source domain data privacy issue. In addition, our proposed frameworks are all based on the premise of privacy non-disclosure, so our method is more information-deficient compared with conventional unsupervised domain adaptation methods. Therefore, we can take the conventional unsupervised domain adaptation ABSA task as the upper bound of our experiments (sufficient source domain information) and the adaptation task directly fine-tuning base-model on the source training data as the lower bound of our experiments (lack of source domain information). If our method can achieve comparable performance to conventional unsupervised domain adaptation, it demonstrates the robust performance of our method in the absence of information.

5.4. Results

The results are shown in Table 2.The table consists of three parts: **(1) Source-Only:** target performance without domain adaptation; **(2) Source-Required:** target performance with source-required domain adaptation; **(3) Source-Free:** tar-

Methods	BERT-Base	SHOT	SF-ABSA
R→S	42.2	26.52	**35.67**
L→S	20.99	17.00	**29.62**
D→S	13.64	10.45	**45.93**
S→R	42.2	37.95	**44.62**
L→R	39.14	33.57	**44.45**
D→R	30.06	26.94	**35.43**

Table 4: Under the same source-free setting, the performance of each model.

get performance with source-free domain adaptation.

For Source-Only, the BERT-Base model has been trained on the source domain dataset and then evaluated on each target dataset without any domain adaptation method. For the Source-Required, we adopt state-of-the-art methods. These methods are mainly based on adversarial training approaches and generation-based methods, which learn domain-invariant representations by minimizing the feature differences between the source domain and target domain or generating high-quality target domain reviews. However, these methods heavily rely on the availability of source and target domain data, which may lead to privacy leakage issues. In contrast, the proposed SF-ABSA framework does not require access to source domain data and can preserve data privacy.

First, the proposed SF-ABSA framework performs better compared to BERT-base, which indicates that SF-ABSA has better domain adaptability. Furthermore, our SF-ABSA framework achieves comparable performance to source-required methods on the premise of only transferring model parameters without accessing source-domain data, which demonstrates the significance of our approach in the privacy-preserving domain.

In addition, we find that both feature-based and pseudo-label-based methods have their own advantages. When the target domain is the **Service** domain dataset, the pseudo-label based approach has a significant effect on performance improvement, as shown in the Table 3, which we believe is due to the overall smaller size of the Service dataset and the relatively small total number of words contained in the sentences, which results in less noise in the computation of the category feature anchors and higher accuracy of pseudo-label in the reassignment. For example, in the D→S task, our method outperforms the state-of-the-art method by 8.5 percentage points. In the case of longer sentences, the dependency information in the feature-based method can improve model performance, but the pseudo-label method weakens model performance.

R→S	likely, proud, impressive, work, contentious, bearing, hated,beauty, canned, mistake, madden, nicely,catalogue,difficulty
L→S	kind, fork, appears, weary, desk, projects, enjoy, kindly, well-served, general, comprehensive well, quality, nice
D→S	desk, providing, interaction, recommend, satisfaction, robust, experience, sophisticated, comprehensive, nicely, good, appearing

Table 5: Words with higher instance weights in our UDA approach.

Domain	m=1	m=2	m=3
R→S	**35.67**	31.61	30.72
L→S	**29.62**	25.12	25.68
D→S	**45.93**	41.98	40.56
S→R	**44.62**	41.79	41.32
L→R	**44.45**	42.54	40.19
D→R	**35.43**	32.15	31.08

Table 6: Effect of the number of iterations to reassign pseudo-labels on model performance.

5.5. Analysis

In the above experiments, the conditions on which the various models are based are not consistent, and it is unfair to directly compare with each other, so we choose the classic baseline SHOT in the source-free field for comparison. SHOT is a method in the field of computer vision, but it has a strong generalization ability and can be generalized to our ABSA task for comparison. Tabel 4 shows the comparison of our proposed method with other methods under the same source-free setting. We compared our method with SHOT under the same source-free setting, and we found that our method has obvious advantages. SHOT's method, although highly generalizable, still does not achieve great performance, illustrating the progress of our method in the field of privacy.

Tabel 5 shows the words with higher instance weights in our UDA approach. When calculating the category center point, the feature vectors of these words will have relatively large weights. We can think that these words can express both the information of the source domain and the information of the target domain.

Table 6 shows the effect of the hyperparameter m on the performance of the model in the pseudo-label-based method. m represents the number of times to iteratively calculate the category center point and then reassign the pseudo-label. We found that the pseudo-label obtained after one iteration is the best.

6. Conclusion

In this paper, we explore the ABSA task of source-free unsupervised domain adaptation. We propose a joint framework of feature-based methods and pseudo-labeling methods. Our framework achieves comparable performance to conventional unsupervised domain adaptation methods under the premise of insufficient information and without access to source domain data. This demonstrates the superiority of our method under the source-free setting. In the future, we will further explore ways to obtain high-quality pseudo-labels across modalities in the source-free setting.

7. References

Caroline Brun, Diana Nicoleta Popa, and Claude Roux. 2014. Xrce: Hybrid classification for aspect-based sentiment analysis. In *SemEval@ COLING*, pages 838–842. Citeseer.

Fang Chen and Yongfeng Huang. 2019. Knowledge-enhanced neural networks for sentiment analysis of chinese reviews. *Neurocomputing*, 368:51–58.

Cyprien de Masson D'Autume, Sebastian Ruder, Lingpeng Kong, and Dani Yogatama. 2019. Episodic memory in lifelong language learning. *Advances in Neural Information Processing Systems*, 32.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Ning Ding, Yixing Xu, Yehui Tang, Chao Xu, Yunhe Wang, and Dacheng Tao. 2022. Source-free domain adaptation via distribution estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7212–7222.

Ying Ding, Jianfei Yu, and Jing Jiang. 2017. Recurrent neural networks with auxiliary labels for cross-domain opinion target extraction. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31.

Mark Dredze, Alex Kulesza, and Koby Crammer. 2010. Multi-domain learning by confidence-weighted parameter combination. *Machine Learning*, 79(1-2):123–149.

Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. 2016. Domain-adversarial training of neural networks. *The journal of machine learning research*, 17(1):2096–2030.

Chenggong Gong, Jianfei Yu, and Rui Xia. 2020. Unified feature and instance based domain adaptation for aspect-based sentiment analysis. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7035–7045.

Ruidan He, Wee Sun Lee, Hwee Tou Ng, and Daniel Dahlmeier. 2018. Adaptive semi-supervised learning for cross-domain sentiment classification. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3467–3476.

Dan Hendrycks, Xiaoyuan Liu, Eric Wallace, Adam Dziedzic, Rishabh Krishnan, and Dawn Song. 2020. Pretrained transformers improve out-of-distribution robustness. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 2744–2751.

Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177.

Svetlana Kiritchenko, Xiaodan Zhu, Colin Cherry, and Saif Mohammad. 2014. Nrc-canada-2014: Detecting aspects and sentiment in customer reviews. In *Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014)*, pages 437–442.

Egoitz Laparra, Xin Su, Yiyun Zhao, Özlem Uzuner, Timothy A Miller, and Steven Bethard. 2021. Semeval-2021 task 10: source-free domain adaptation for semantic processing. Association for Computational Linguistics (ACL).

Junjie Li, Jianfei Yu, and Rui Xia. 2022. Generative cross-domain data augmentation for aspect and opinion co-extraction. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 4219–4229.

Xin Li, Lidong Bing, Piji Li, and Wai Lam. 2019a. A unified model for opinion target extraction and target sentiment prediction. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 6714–6721.

Xin Li, Lidong Bing, Piji Li, Wai Lam, and Zhimou Yang. 2018a. Aspect term extraction with history attention and selective transformation. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence*, pages 4194–4200.

Zheng Li, Xin Li, Ying Wei, Lidong Bing, Yu Zhang, and Qiang Yang. 2019b. Transferable end-to-end aspect-based sentiment analysis with selective adversarial learning. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 4590–4600.

Zheng Li, Ying Wei, Yu Zhang, and Qiang Yang. 2018b. Hierarchical attention transfer network for cross-domain sentiment classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32.

Zheng Li, Ying Wei, Yu Zhang, Xiang Zhang, and Xin Li. 2019c. Exploiting coarse-to-fine task transfer for aspect-level sentiment classification. In *Proceedings of the AAAI conference on artificial intelligence*, volume 33, pages 4253–4260.

Jian Liang, Dapeng Hu, and Jiashi Feng. 2020. Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation. In *International Conference on Machine Learning*, pages 6028–6039. PMLR.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Asha S Manek, P Deepa Shenoy, and M Chandra Mohan. 2017. Aspect term extraction for sentiment analysis in large movie reviews using gini index feature selection method and svm classifier. *World wide web*, 20:135–154.

Maria Pontiki, Dimitrios Galanis, Harris Papageorgiou, Suresh Manandhar, and Ion Androutsopoulos. 2015. Semeval-2015 task 12: Aspect based sentiment analysis. In *Proceedings of the 9th international workshop on semantic evaluation (SemEval 2015)*, pages 486–495.

Maria Pontiki, Dimitris Galanis, Haris Papageorgiou, Ion Androutsopoulos, Suresh Manandhar, Mohammed AL-Smadi, Mahmoud Al-Ayyoub, Yanyan Zhao, Bing Qin, Orphée De Clercq, et al. 2016. Semeval-2016 task 5: Aspect based sentiment analysis. In *ProWorkshop on Semantic Evaluation (SemEval-2016)*, pages 19–30. Association for Computational Linguistics.

Maria Pontiki, Haris Papageorgiou, Dimitrios Galanis, Ion Androutsopoulos, John Pavlopoulos, and Suresh Manandhar. 2014. Semeval-2014 task 4: Aspect based sentiment analysis. *SemEval 2014*, page 27.

Soujanya Poria, Erik Cambria, and Alexander Gelbukh. 2016. Aspect extraction for opinion mining with a deep convolutional neural network. *Knowledge-Based Systems*, 108:42–49.

Guang Qiu, Bing Liu, Jiajun Bu, and Chun Chen. 2011. Opinion word expansion and target extraction through double propagation. *Computational linguistics*, 37(1):9–27.

Xin Su, Yiyun Zhao, and Steven Bethard. 2022. A comparison of strategies for source-free domain adaptation. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8352–8367.

Fan-Keng Sun, Cheng-Hao Ho, and Hung-Yi Lee. 2019. Lamol: Language modeling for lifelong language learning. In *International Conference on Learning Representations*.

Duyu Tang, Bing Qin, and Ting Liu. 2016. Aspect level sentiment classification with deep memory network. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 214–224.

Cigdem Toprak, Niklas Jakob, and Iryna Gurevych. 2010. Sentence and expression level annotation of opinions in user-generated discourse. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pages 575–584.

Shuai Wang, Sahisnu Mazumder, Bing Liu, Mianwei Zhou, and Yi Chang. 2018. Target-sensitive memory networks for aspect sentiment classification. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*.

Wenya Wang and Sinno Jialin Pan. 2018. Recursive neural structural correspondence network for cross-domain aspect and opinion co-extraction. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2171–2181.

Wenya Wang, Sinno Jialin Pan, Daniel Dahlmeier, and Xiaokui Xiao. 2016. Recursive neural conditional random fields for aspect-based sentiment analysis. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 616–626.

Haifeng Xia, Handong Zhao, and Zhengming Ding. 2021. Adaptive adversarial network for source-free domain adaptation. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9010–9019.

Rui Xia, Jianfei Yu, Feng Xu, and Shumei Wang. 2014. Instance-based domain adaptation in nlp via in-target-domain logistic approximation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 28.

Yanmin Xiang, Hongye He, and Jin Zheng. 2018. Aspect term extraction based on mfe-crf. *Information*, 9(8):198.

Hai Ye, Qingyu Tan, Ruidan He, Juntao Li, Hwee Tou Ng, and Lidong Bing. 2020. Feature adaptation of pre-trained language models across languages and domains with robust self-training. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 7386–7399.

Haibo Yu, Guojun Lu, Qianhua Cai, and Yun Xue. 2022. A kge based knowledge enhancing method for aspect-level sentiment classification. *Mathematics*, 10(20):3908.

Jianfei Yu, Chenggong Gong, and Rui Xia. 2021. Cross-domain review generation for aspect-based sentiment analysis. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 4767–4777.

Jie Zhou, Jimmy Xiangji Huang, Qinmin Vivian Hu, and Liang He. 2020. Sk-gcn: Modeling syntax and knowledge via graph convolutional network for aspect-level sentiment classification. *Knowledge-Based Systems*, 205:106292.