Cross-domain Sentiment Classification based on Adaptive Center Contrastive Learning

Stanford CS224N Custom Project

Xinqi Fan
Stanford University
xinqifan@stanford.edu

Abstract

Sentiment analysis refers to automatically recognizing affects from natural language. Since people use different words in each context, it leads to the domain shift problem, which can cause a significant decrease in training and testing different domains. In this work, we proposed an adaptive center contrastive learning (ACCL) method to tackle the problem. First, to learn semantically meaningful features and minimize inter-class variance, we introduced representation centers learned by center contrastive learning (CCL). Second, to reduce the representation discrepancy between the source and target domain, we proposed an adaptive contrastive learning (ACL) strategy, which used entropy-based pseudo-labels generation for high confidence target domain samples and trained them with the CCL, which can learn a shared representation between source and target domain. We conducted experiments on a widely-used cross-domain sentiment analysis dataset - the Amazon review dataset. The ablation study demonstrated the effectiveness of the proposed method. Compared with other methods, our methods also showed satisfactory performance on many source and target domain pairs.

1 Key Information to include

- Mentor: Anna Goldie
- External Collaborators (if you have any): No
- Sharing project: No

2 Introduction

Sentiment analysis refers to automatically recognizing affects from natural language [1]. It has a wide range of applications in business, education, healthcare, manufacturing, and many other fields for understanding reviews and recommendations. In recent years, deep learning based methods have dramatically improved sentiment analysis performance. However, some methods perform poorly in real-world applications, because their training data comes from a different domain than the deployment scenario. This problem is called domain shift [2]. Such domain shift problem makes a significant decrease of performance for the model trained on one domain (source domain), but applied in another domain (target domain) [3]. Therefore, there is a great need to solve this problem by making the model trained on the source domain perform well on the target domain, which is also called the domain adaptation (DA) task.

There are a few approaches designed to tackle the DA task. The main idea of most approaches is to learn domain invariant features, i.e., making the distribution of features from the source and target domain similar [3]. To achieve this aim, [4] proposed a method to reconstruct target domain inputs. More recently, self-supervised contrastive learning (CL) has been proposed, which does not require the model to reconstruct the input but explicitly minimizes the discrepancy of the representations.

However, the self-supervised learning approaches do not utilize the semantically meaningful labels, which are available, during the representation learning. Another way to solve the DA task is to improve the model's generalization. A powerful pre-trained language model, bidirectional encoder representations from transformers (BERT), has been proposed in [5]. It forms language modeling tasks using masks and lets the BERT model learn general knowledge of the languages. However, this knowledge may not help the DA task because the learned knowledge is domain and task agnostic.

Therefore, we proposed an adaptive center contrastive learning (ACCL) method to tackle the problem. First, to learn semantically meaningful features and minimize inter-class variance, we introduced representation centers learned by center contrastive learning (CCL). Second, to reduce the representation discrepancy between the source and target domain, we proposed an adaptive contrastive learning (ACL) strategy, which used entropy-based pseudo-labels generation for high confidence target domain samples and trained them with the CCL, which can learn a shared representation between source and target domain. We conducted experiments on a widely-used cross-domain sentiment analysis dataset - the Amazon review dataset. The ablation study demonstrated the effectiveness of the proposed method. Compared with other methods, our methods also showed satisfactory performance on many source and target domain pairs.

The rest of the report is organized as follows. Related work is discussed in Section 2. The proposed approach is introduced in Section 3. We talked about experiments in Section 4, and qualitative analysis in Section 5. Finally, we conclude the report and point out future work in Section 6.

3 Related Work

3.1 Sentiment Analysis

The sentiment can be measured in various ways, while the sentiment classification problem considers the sentiment as discrete classes, such as positive, negative, and neutral [6]. There are fine-grained sentiment classification studies for identifying happiness, sadness, anger, disgust, surprise, and fear, but these tasks are usually called emotion classifications [6]. Regarding different formats, there are sentence-level sentiment classification and document-level sentiment classification. For sentence-level sentiment classification, Kim [7] first introduced the convolutional neural network (CNN) to the field, while Wang et al. [8] applied a disconnected recurrent neural network (RNN). For the document-level sentiment classification, [9] proposed a gated recurrent neural network (GRNN) using CNN or RNN at the low level and hierarchically combining the representations to obtain the document representation. However, because of the differences between domains, a sentiment classifier trained on one domain may not perform well on new domains, necessitating the research to bridge this domain gap.

3.2 Domain Adaptation

Since people use different words in each context, it leads to the domain shift problem [2]. Such domain shift problem makes a significant decrease of performance for the model trained on one domain (source domain), but applied in another domain (target domain) [3]. Therefore, the DA problem is proposed to alleviate the domain shift problem and maintain a relatively satisfactory performance on the target domain. The DA task can be classified into unsupervised and semi-supervised domain adaptation. In particular, we focus on the unsupervised domain adaptation (UDA) problem, where the model can access the labeled source domain data, and unlabeled target domain data. In terms of discrepancy-based approaches, deep domain confusion (DDC) [10] used maximum mean discrepancy (MMD) to pull together the source and target domain features, and then deep adaptation network (DAN) [11] improved it by considering multiple kernels. Domain adversarial neural network (DANN) introduced the adversarial idea training idea to fool the domain classifier so that it cannot distinguish the inputs from the source and target domain [12]. [4] proposed a method to reconstruct target domain inputs. [13] designed an easy-to-hard strategy to select reliable pseudo-labels progressively. Our method is inspired by the idea of generating pseudo-labels.

3.3 Contrastive Learning

CL has been a very popular representation learning method in recent years. [14] proposed the instance discrimination task and the memory bank strategy. Contrastive predictive coding

(CPC) [15] regarded the future input features as positive samples, while contrastive multiview coding (CMC) [16] treated the different views of the same input as positive samples. In addition, CPC suggested using the Info Noise-Contrastive Estimation (InfoNCE) loss, which is called the contrastive loss in many studies. Momentum contrast (MoCo) [17] used a momentum encoder to update the encoder and a queue to decouple the batch size. The simple contrastive representation learning (SimCLR) framework [18] unified the CL framework by using data augmentation for creating multiple views, and proving the usefulness of adding projection layers. [2] applied the SimCLR to learn useful representations from both source and target domain, so that the domain discrepancy can be minimized automatically. However, SimCLR does not use labels, which may push the representations of the same class away from each other. Therefore, supervised contrastive learning (SCL) was proposed to pull together features for the same class, and push apart features for different classes [19]. Our work will use the idea of SCL to learn semantically meaningful representations.

4 Approach

To solve the domain shift problem for the sentiment classification, we proposed an ACCL method in Fig. 1. First, to learn semantically meaningful features and minimize inter-class variance, we introduced representation centers learned by CCL. Second, to reduce the representation discrepancy between the source and target domain, we proposed an ACL strategy, which used entropy-based pseudo-labels generation for high confidence target domain samples and trained them with the CCL, which can learn a shared representation between source and target domain.

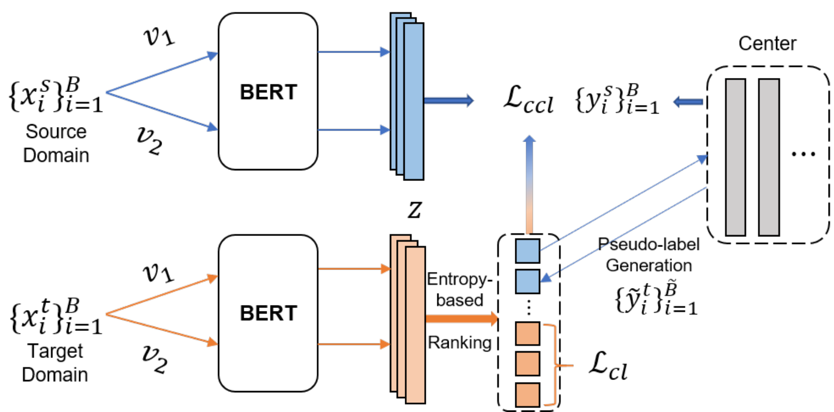

Figure 1: The pipeline of our framework. The blue lines indicate the labeled data flow, while the orange lines indicate the unlabeled data flow. B stands for the batch size, and $\tilde{B}$ is the portion with confidence predictions.

4.1 Problem Formulation

For the cross-domain sentiment classification task, a source domain dataset $D_s = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$ and a target domain dataset $D_t = \{(x_i^t)\}_{i=1}^{N_t}$ are provided. The two datasets are sampled from two different distributions: $p_s(X,Y)$ and $p_t(X,Y)$. Each sample from the source dataset x_i^s has a corresponding label y_i^s, while the samples from the target dataset x_i^t do not have labels. A learned model is expected to perform well on the target dataset.

4.2 Baseline

The feature extractor is a pretrained BERT [5] with a few projection layers f. Two different views x_i and $x_{j(i)}$ of the same input sample $i \in I$ are generated by data augmentation in advance, and fed into the model. The output representation is denoted as $z \in \mathbb{R}^d$. In particular, the representations are $z_i = f(x_(i))$ and $z_{j(i)} = f(x_(j(i)))$. The model computes contrastive loss for unlabeled data as

$$\mathcal{L}_{\text{cl}} = -\sum_{i \in I} \log \frac{\exp\left(z_i \cdot z_{j(i)}/\tau_{cl}\right)}{\sum_{a \in A(i)} \exp\left(z_i \cdot z_a/\tau_{cl}\right)}, \tag{1}$$

where τ_{cl} is a temperature parameter, and $A(i) \equiv I \backslash \{i\}$.

4.3 Center Contrastive Learning

Since the standard self-supervised CL does not use available labels to learn semantic representations, we leverage the SCL framework with labels to learn semantic representation. Furthermore, to minimize intra-class variance, we introduced representation centers. In detail, we created K (the number of classes in the dataset) learnable representation centers $c \in \mathbb{R}^d$, and applied supervised contrastive learning for pulling samples of the same class to the corresponding centers as

$$\mathcal{L}_{\text{ccl}} = \sum_{i \in I} \frac{-1}{|P(c_i)|} \sum_{p \in P(c_i)} \log \frac{\exp\left(c_i \cdot z_p/\tau_{ccl}\right)}{\sum_{b \in B} \exp\left(c_i \cdot z_b/\tau_{ccl}\right)}, \tag{2}$$

where $P(c_i)$ is the set for instances with c_i labels, and B is a batch. Here, an instance means a sample, while a center is an embedding where all instances from the same class should be close to. Note that there are two differences between the original CL and our CCL: one is that our CCL has learnable center embeddings. The second is that we used labels.

4.4 Adaptive Contrastive Learning

To reduce the representation discrepancy between the source and target domain, we proposed an ACL strategy. First, we obtained entropy for all target domain samples. Then, we selected the high confidence target domain samples, and generated pseudo-labels for them. Finally, we train the samples with labels from the source domain, and the samples with pseudo-labels from the target domain jointly using CCL.

To achieve this, it is based on the assumption of [20] that the model can make confident predictions by showing low entropy, while not confident predictions by showing high entropy. Therefore, we calculate the entropy as

$$E_i = \sum_c -z_i^t log(z_i^t), \tag{3}$$

where $z_i^t = f(x_i^t)$ is the representation computed from the target set samples.

Then, we use a threshold $T_E = \frac{|x^{t_c}|}{|x^t|}$, where $|x^t|$ is the cardinality of images in the target domain, $|x^{t_c}|$ is the cardinality of images with confident predictions in the target domain. Note that this threshold is not a hard number, but is based on the number of images in the target domain. The advantage of it is that it can be generalized to other domain case. Since we used the minibatch-based optimization, we denote the $|x^{t_c}|$ as $\tilde{B}$ for the high confidence predictions portion in a minibatch B. For the high confidence potion, we denote their representations as z_i^{tc}. These representations are used to compute the similarity with the class centers, and obtain the pseudo-labels from the higher scores as

$$\tilde{y}_i^t = \text{argmax}_k \{c_1 \cdot z_p, \ldots, c_k \cdot z_p, \ldots, c_K \cdot z_p\}. \tag{4}$$

After obtaining these pseudo-labels, we can use them to compute the CCL. Note that the source domain and target domain share the same center embeddings for the same class, so the domain gap is minimized.

4.5 Data Augmentation

Besides the back translation used by [2], we used (1) random augmenter by substituting characters randomly by WordNet synonyms, (2) contextual augmenter for word substitution using RoBERTa as used in [21]. The probabilities of both augmentations are set as 0.3. We provide a comparison for using these data augmentation techniques in Section 5.4.1.

5 Experiments

5.1 Data

The Amazon review dataset [22] contains reviews of 4 products, books (B), DVD (D), electronics (E), and kitchens (K), as four domains. Each domain has 1000 positive and 1000 negative reviews, and a few unlabeled reviews. We followed the training and testing settings as [2] by having 1400 samples for training and 600 samples for testing. Furthermore, we consider all 12 possible domain pairs as D $\rightarrow$ B, E $\rightarrow$ B, K $\rightarrow$ B, B $\rightarrow$ D, E $\rightarrow$ D, K $\rightarrow$ D, B $\rightarrow$ E, D $\rightarrow$ E, K $\rightarrow$ E, B $\rightarrow$ K, D $\rightarrow$ K, E $\rightarrow$ K, ranther than picking 6 pairs only as [2].

5.2 Evaluation method

For evaluation, we calculated the overall top-1 accuracy on the target domain test set as

$$Accuracy = \frac{N_c}{N}, \tag{5}$$

where N_c is the number of correct predictions, N is the total number of samples on the target domain test set. It is worth noting that we are not allowed to use the target domain labels during training due to the DA task setting.

5.3 Experimental details

Our model was built using the PyTorch deep learning framework. We employed an adaptive momentum decoupled weight decay (AdamW) optimizer with an initial learning rate of 2×10^{-5} and $1/10$ total steps for the learning rate warm up. The model was trained with batch size 6 for 20 epochs on a computer with 2 NVIDIA 2080Ti 11G GPU cards and 8 CPU logical cores. We set the entropy threshold T_E as 0.3, the temperature τ_{cl} of $\mathcal{L}_{cl}$ as 0.05, and the temperature τ_{ccl} of $\mathcal{L}_{ccl}$ as 0.07.

5.4 Results

In order to validate the effectiveness of the proposed method, we conducted ablation studies, and compared our performance with other methods. In addition, we also investigated the performance using different data augmentations.

5.4.1 Quantitative results for data augmentations

We show the accuracies on the target domain test set for different data augmentations discussed in Section 4.5 in Table 1. It shows that the back translation has the best performance. One possible reason is that the back-translation needs to translate the sentence to another language and then translate it back, so it creates more variants for the sentences, while keeping the same semantic meaning. The contextual augmentation method achieved a slightly lower but similar accuracy as the back-translation method, but the random augmentation method has the worst result.

Table 1: Quantitative results (K $\rightarrow$ E) for different data augmentations.

Data augmentation	**Accuracy** (%)
Back trans.	92.41
Random aug.	91.74
Contextual aug.	92.35

5.4.2 Quantitative ablation study

We show the quantitative results on the target domain test set for adding new modules in Table 2. It shows that we have slightly improved the accuracy by adding the CCL loss to the baseline. This improvement may come from using labels to extract more semantically meaningful representation during representation learning. Then, we can observe that it has a relatively larger improvement by adding the ACL strategy, since we can reduce the source and target domains' discrepancy by using the generated pseudo-labels in the target domain as guidance.

Table 2: Quantitative results (K $\rightarrow$ E) for the ablation study.

CCL	**ACL**	**Accuracy** (%)
✗	✗	92.27
✓	✗	92.30
✓	✓	92.41

5.4.3 Qualitative comparison with other methods

We compare our method with other widely-used methods on the cross-domain sentiment analysis in Table 3. The methods include the domain adversarial neural network (DANN) [23], the pivot based language model (PBLM) [24], the hierarchical attention transfer network (HATN) [25], the adversarial category alignment network (ACAN) [26], the interactive attention transfer network [27], fine-tuned BERT [5], domain-aware and adversarial BERT (DAAT) [28], and the CLIM implemented by us.

For the averaged accuracy of 12 domain pairs, we can see that our ACCL method achieved the best performance, exceeding the DAAT and CLIM by about 1% and 0.1%, respectively. In general, we find the results are higher when the target domains are electronics or kitchens rather than books or DVDs. We think this phenomenon is caused by the fact that these reviews in these two domains have very general words and rarely use specific terminologies. However, reviewers mentioned many terms for the books or DVDs domains that may only be used in those domains, which makes them hard to be understood for the models trained in another domain.

Table 3: Comparison with other methods on the Amazon review dataset.

S → T	**Accuracy** (%)								
	DANN	PBLM	HATN	ACAN	IATN	BERT	DAAT	CLIM	ACCL
D → B	81.70	82.50	86.30	82.35	87.00	89.40	90.86	91.25	**91.53**
E → B	78.55	71.40	81.00	79.75	81.80	86.50	88.91	89.08	**89.24**
K → B	79.25	74.20	83.30	80.80	84.70	87.55	87.98	**88.52**	88.41
B → D	82.30	84.20	86.10	83.45	86.80	88.96	89.70	**90.21**	89.97
E → D	79.70	75.00	84.00	81.75	84.10	87.95	90.13	**91.53**	91.42
K → D	80.45	79.80	84.50	82.10	84.10	87.30	88.81	**89.90**	89.37
B → E	77.60	77.60	85.70	81.20	86.50	86.15	89.57	91.92	**92.74**
D → E	79.70	79.60	85.60	82.80	86.90	86.55	89.30	90.15	**91.03**
K → E	86.65	87.10	87.00	86.60	87.60	90.45	91.72	92.27	**92.41**
B → K	76.10	82.50	85.20	83.05	85.90	89.05	90.75	**93.20**	92.67
D → K	77.35	83.20	86.20	78.06	85.80	87.53	90.50	91.89	**92.38**
E → K	83.95	87.80	87.90	83.35	88.70	91.60	93.18	93.21	**93.50**
Average	80.29	80.40	85.10	82.15	85.90	88.25	90.12	91.09	**91.22**

6 Analysis

Qualitative evaluation of our method is provided in Table 4, where we have shown some sample review's results obtained with (w/) our ACCL or without (w/o) our ACCL method. Most of these reviews are selected from the electronics domain with the model trained in the kitchen domain.

From Table 4, we can see that when the reviews have explicit sentiment-related words, like excellent, it is easier to obtain the correct prediction (as shown in the first row). However, it becomes harder when there are no explicit words like in the second row. Our method with ACCL can correctly predict it as a positive result. When the review becomes longer, like in the third and sixth rows, it becomes tough to predict the correct sentiment. Since the sentiment may change from one sentence to another sentence across the whole review, our model may get confused from part of the review.

7 Conclusion

In this work, we proposed an ACCL method to tackle the DA task in sentiment analysis. First, this study has shown that the CCL is useful for learning semantically meaningful representations. The second major finding is that the pseudo-labels can minimize the domain discrepancy through learning shared centers between source and target domains. The ablation study and comparison with other methods proved the effectiveness of ACCL by showing satisfactory performance. In the future, we will further improve the method by using better initial states for the centers rather than randomly

Table 4: Qualitative analysis of reviews. + means a positive review, while - means a negative review.

Review	w/o ACCL	w/ ACCL	Ground truth
It certainly did. Excellent color qualit	+	+	+
I purchased this for my wife who wanted a case for her	-	+	+
I bought these speakers some time back. I was a bit skeptical because the sub woofer looked a little wierd. but once I plugged it in they sounded really good	+	-	+
Only thing I did not care for was the design on the top of the disk. Other than that, these are quality DVD	-	+	+
I like the features. I'm satisfied with the sound. The user interface could be more intuitive.	+	+	-
I bought this machine based on the good reviews I found here. I have had it for one month and it does nothing but jam. Now the two week return policy at the store has expired and I am left with a jamming piece o' junk.	+	+	-

initializing them. In addition, we should run multiple experiments to reduce the stochasticity of the results when there is more time available.

8 Acknowledgement

I would like to thank Dr. Anna Goldie for her valuable suggestions.

References

[1] Marouane Birjali, Mohammed Kasri, and Abderrahim Beni-Hssane. A comprehensive survey on sentiment analysis: Approaches, challenges and trends. *Knowledge-Based Systems*, 226:107134, 2021.

[2] Tian Li, Xiang Chen, Shanghang Zhang, Zhen Dong, and Kurt Keutzer. Cross-domain sentiment classification with contrastive learning and mutual information maximization. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 8203–8207. IEEE, 2021.

[3] Mei Wang and Weihong Deng. Deep visual domain adaptation: A survey. *Neurocomputing*, 312:135–153, 2018.

[4] Muhammad Ghifary, W Bastiaan Kleijn, Mengjie Zhang, David Balduzzi, and Wen Li. Deep reconstruction-classification networks for unsupervised domain adaptation. In *European Conference on Computer Vision*, pages 597–613. Springer, 2016.

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

[6] JianHua Yuan, Yang Wu, Xin Lu, YanYan Zhao, Bing Qin, and Ting Liu. Recent advances in deep learning based sentiment analysis. *Science China Technological Sciences*, 63(10):1947–1970, 2020.

[7] Yoon Kim. Convolutional neural networks for sentence classification. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pages 1746–1751, Doha, Qatar, October 2014. Association for Computational Linguistics.

[8] Baoxin Wang. Disconnected recurrent neural networks for text categorization. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pages 2311–2320, 2018.

[9] Duyu Tang, Bing Qin, and Ting Liu. Document modeling with gated recurrent neural network for sentiment classification. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1422–1432, 2015.

[10] Eric Tzeng, Judy Hoffman, Ning Zhang, Kate Saenko, and Trevor Darrell. Deep domain confusion: Maximizing for domain invariance. *arXiv preprint arXiv:1412.3474*, 2014.

[11] Mingsheng Long, Yue Cao, Zhangjie Cao, Jianmin Wang, and Michael I Jordan. Transferable representation learning with deep adaptation networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(12):3071–3085, 2018.

[12] Yaroslav Ganin and Victor Lempitsky. Unsupervised domain adaptation by backpropagation. In *International Conference on Machine Learning*, pages 1180–1189. PMLR, 2015.

[13] Chaoqi Chen, Weiping Xie, Wenbing Huang, Yu Rong, Xinghao Ding, Yue Huang, Tingyang Xu, and Junzhou Huang. Progressive feature alignment for unsupervised domain adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 627–636, 2019.

[14] Zhirong Wu, Yuanjun Xiong, Stella X Yu, and Dahua Lin. Unsupervised feature learning via non-parametric instance discrimination. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3733–3742, 2018.

[15] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. Representation learning with contrastive predictive coding. *arXiv preprint arXiv:1807.03748*, 2018.

[16] Yonglong Tian, Dilip Krishnan, and Phillip Isola. Contrastive multiview coding. In *European Conference on Computer Vision*, pages 776–794. Springer, 2020.

[17] Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9729–9738, 2020.

[18] Ting Chen, Simon Kornblith, Mohammad Norouzi, and Geoffrey Hinton. A simple framework for contrastive learning of visual representations. In *International Conference on Machine Learning*, pages 1597–1607. PMLR, 2020.

[19] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning. *Advances in Neural Information Processing Systems*, 33:18661–18673, 2020.

[20] Tuan-Hung Vu, Himalaya Jain, Maxime Bucher, Matthieu Cord, and Patrick Pérez. Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2517–2526, 2019.

[21] Dejiao Zhang, Feng Nan, Xiaokai Wei, Shang-Wen Li, Henghui Zhu, Kathleen Mckeown, Ramesh Nallapati, Andrew O Arnold, and Bing Xiang. Supporting clustering with contrastive learning. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5419–5430, 2021.

[22] John Blitzer, Mark Dredze, and Fernando Pereira. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 440–447, 2007.

[23] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *The Journal of Machine Learning Research*, 17(1):2096–2030, 2016.

[24] Yftah Ziser and Roi Reichart. Pivot based language modeling for improved neural domain adaptation. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1241–1251, 2018.

[25] Zheng Li, Ying Wei, Yu Zhang, and Qiang Yang. Hierarchical attention transfer network for cross-domain sentiment classification. In *Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence*, AAAI'18/IAAI'18/EAAI'18. AAAI Press, 2018.

[26] Xiaoye Qu, Zhikang Zou, Yu Cheng, Yang Yang, and Pan Zhou. Adversarial category alignment network for cross-domain sentiment classification. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 2496–2508, 2019.

[27] Kai Zhang, Hefu Zhang, Qi Liu, Hongke Zhao, Hengshu Zhu, and Enhong Chen. Interactive attention transfer network for cross-domain sentiment classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 5773–5780, 2019.

[28] Chunning Du, Haifeng Sun, Jingyu Wang, Qi Qi, and Jianxin Liao. Adversarial and domain-aware BERT for cross-domain sentiment analysis. In *Proceedings of the 58th Annual Meeting of the Association of Computational Linguistics*, pages 4019–4028, 2020.