An Entity-Aware Adversarial Domain Adaptation Network for Cross-Domain Named Entity Recognition (Student Abstract)

Qi Peng, [1,2] **Changmeng Zheng,** [1,2] **Yi Cai,** [1,2*] **Tao Wang,** [3] **Haoran Xie,** [4] **Qing Li** [5]
[1] School of Software Engineering, South China University of Technology, Guangzhou, China
[2] Key Laboratory of Big Data and Intelligent Robot (South China University of Technology), Ministry of Education
[3] Department of Biostatistics and Health Informatics, King's College London
[4] Department of Computing and Decision Sciences, Lingnan University, Hong Kong, China
[5] Department of Computing, Hong Kong Polytechnic University, Hong Kong, China
{se_pengqi, sethecharm}@mail.scut.edu.cn, ycai@scut.edu.cn, {wtgmme, hrxie2}@gmail.com, csqli@comp.polyu.edu.hk

Abstract

Existing methods for named entity recognition are critically relied on labeled data. To handle the situation that the data is fully-unlabeled, we propose an entity-aware adversarial domain adaptation network, which utilizes the labeled source data and then adapts to unlabeled target domain. We first apply adversarial training to reduce the distribution gap between different domains. Furthermore, we introduce an entity-aware attention to guide adversarial process to achieve the alignment of entity features. The experiment shows that our model outperforms the state-of-the-art approaches.

Introduction

Named entity recognition (NER) is a task that seeks to locate and classify named entity mentions in unstructured text into pre-defined categories. Existing NER models achieve good performance by using a large amount of labeled data. However, In some domain the data is fully-unlabeled. To deal with the problem, existing domain adaptation method utilizes language model to transfer NER knowledge from labeled source to unlabeled target domain (Jia, Liang, and Zhang 2019). However, the data distributions between different domains are disparate, which are not taken into account by language model. As a result, knowledge from source domain cannot be fully transferred to the target domain, which leads to poor performance in target domain. We consider it is a good choice to utilize adversarial training, which learns transferable features to match data distributions across domains and minimizes domain discrepancy distance.

Although adversarial training can align distributions between different domains, it aligns the whole source and target distributions without considering the fine-grained alignment on entity features, as shown in Figure 1. The false alignment of entity features will cause such consequence that the entity information in source cannot be well transferred to target domain, which will lead to incorrect entities prediction in target domain. To deal with the problem, we need to give higher attention to entity features in adversarial training to achieve entity-level alignment.

*Corresponding author: Yi Cai, ycai@scut.edu.cn

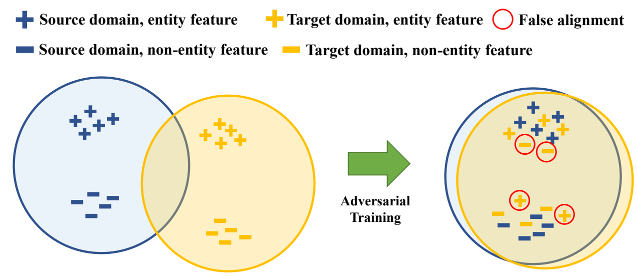

Figure 1: The false alignment of entity features across different domains. Adversarial training aligns the whole domain distributions while ignores entity-level alignment.

In this paper, we propose an entity-aware adversarial domain adaptation network for cross-domain NER. We apply adversarial training to reduce domain distribution shift. In addition, we introduce an entity-aware attention mechanism, which uses the entity information to guide the adversarial training to align entity features across different domains. We evaluate the proposed method through conducting experiment on different datasets. The experimental result shows that the proposed model outperforms other existing models.

Our Model

Input Layer For a given sentence $\boldsymbol{x}$ in source or target domain, we first map each word to word-level embedding using pretrained Glove. Then we propose two feature extractors, the NER-specific LSTM and the Adversarial-specific (Adv-specific) LSTM, and the hidden states of different feature extractor can be computed as follows.

$$\boldsymbol{h}^N = BiLSTM(\boldsymbol{x}; \theta_N) \quad (1)$$

$$\boldsymbol{h}^A = BiLSTM(\boldsymbol{x}; \theta_A) \quad (2)$$

where θ_N and θ_A represent the parameters of NER-specific LSTM and Adversarial-specific LSTM, respectively.

Entity-Aware Attention Layer To obtain domain-invariant entity features, we propose entity-aware attention to exploit entity information to guide the adversarial training. Entity-aware attention mechanism will utilize implicit entity information to give entity feature higher attention score, which influences adversarial training process and gets a better entity-level alignment. We choose multi-head attention as our attention mechanism:

$$\boldsymbol{h}^{EA} = MHA(\boldsymbol{h}^N, \boldsymbol{h}^A) \tag{3}$$

where $MHA(\cdot)$ denotes multi-head attention; $\boldsymbol{h}^N$ is the query sequence. $\boldsymbol{h}^N$ is NER-specific feature extractor output, thus its hidden states contain entity information implicitly, which will guide $\boldsymbol{h}^A$ focus on the entity feature. Finally the entity-aware adversarial feature $\boldsymbol{h}^{EA}$ is generated.

Adversarial Training To reduce the distribution gap of entity-aware adversarial features $\boldsymbol{h}^{EA}$ between source and target domain, we exploit adversarial training. Specially, we set a domain discriminator, which takes the feature $\boldsymbol{h}^{EA}$ as input and attempts to discriminate the domain of the feature. The overall domain adversarial training loss:

$$L_{Adv} = \min_{\theta_d}(\max_{\theta_f}(\frac{1}{n_s + n_t}\sum_{i=1}^{n_s+n_t} D_i)) \tag{4}$$

n_s and n_t indicate the number of training samples in source and target domain, respectively; θ_d and θ_f denotes the parameters of domain discriminator and feature extractor, respectively; D_i is the loss of ith sample in the discriminator.

Label Decoder Given $\boldsymbol{h}^c = \boldsymbol{h}^N \oplus \boldsymbol{h}^{EA}$ as input feature, we use standard conditional random field as our NER label decoder. The label decoder is trained by source data in the training stage because only source domain has labeled data.

Experiment

Datasets

CoNLL. A collection of newswire articles from Reuters Corpus (Tjong Kim Sang and De Meulder 2003). It belongs to newswire domain and the language style is formal.

Twitter. The dataset is a collection of tweets posted by Twitter users (Zhang et al. 2018), which belongs to social media domain. The language style is more casual and the semantic information is sparse in most cases.

Baseline Methods

We compare our models with several unsupervised cross-domain NER models, including *Transfer* model (We train Bi-LSTM+CRF on source data and test on target data. *Trans.* for short.), *DANN* (Ganin et al. 2016), *Cross-Domain+LM* (*CDLM* for short.) (Jia, Liang, and Zhang 2019).

Experimental Results

We use precision, recall and F1 score as our evaluation metrics. The results are shown in table 1. $C \rightarrow T$ means we use CoNLL as source data and Twitter as target data; $T \rightarrow C$ means we use Twitter as source data and CoNLL as target data. We can find that when transferring from CoNLL to Twitter or from Twitter to CoNLL, the SOTA method *CDLM* has a bad performance. It may be due to the context patterns between news and twitter differ greatly, thus the language model can not accomplish knowledge transfer well.

Model	$C \rightarrow T$			$T \rightarrow C$		
Metric	P(%)	R(%)	F(%)	P(%)	R(%)	F(%)
Trans.	44.84	57.86	50.52	74.23	63.54	68.47
DANN	45.59	58.70	51.33	74.56	64.43	69.13
CDLM	59.41	44.53	50.91	61.69	52.00	56.55
Our	51.73	58.01	**54.69**	76.84	65.60	**70.77**

Table 1: Experiments on different transferring situations.

Model	**P(%)**	**R(%)**	**F1(%)**
Base	45.12	56.68	50.24
Base+AT	47.15	58.61	52.26
Base+AT+EA-att	51.73	58.01	54.69

Table 2: Results of ablation study.

While our model take the distribution of entity features into account without the influence of context patterns, thus our model can get a better performance. To evaluate the contribution of adversarial training and entity-aware attention, we implement the ablation experiment, which is shown in table 2. We choose $C \rightarrow T$ as transferring setting. *AT* and *EA-att* represent the added adversarial training and entity-aware attention module, respectively. With *AT* and *EA-att* module, the transferring result has a significant improvement.

Conclusion

In this paper, we propose an entity-aware adversarial domain adaptation network. Our model bridges the distribution gap between different domains by applying adversarial training. Furthermore, we introduce the entity-aware attention to achieve entity-level alignment during adversarial training. Experiments on different datasets show that our model outperforms other existing models.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities, SCUT(No.D2182480), the National Key Research and Development Program of China, the Science and Technology Programs of Guangzhou (No.201802010027, 201902010046), National Natural Science Foundation of China (62076100).

References

Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; Marchand, M.; and Lempitsky, V. 2016. Domain-adversarial training of neural networks. *The Journal of Machine Learning Research* 17(1): 2096–2030.

Jia, C.; Liang, X.; and Zhang, Y. 2019. Cross-Domain NER using Cross-Domain Language Modeling. In *ACL*, 2464–2474.

Tjong Kim Sang, E. F.; and De Meulder, F. 2003. Introduction to the CoNLL-2003 shared task: language-independent named entity recognition. In *NAACL*, 142–147.

Zhang, Q.; Fu, J.; Liu, X.; and Huang, X. 2018. Adaptive Co-attention Network for Named Entity Recognition in Tweets. In *AAAI*, 5674–5681.