ICANN19
28th International Conference on Artificial Neural Networks

An Empirical Study of Multi-domain and Multi-task Learning in Chinese Named Entity Recognition

Yun Hu [1,2] , Mingxue Liao [2], Pin Lv [2] and Changwen Zheng [2]
[1]University of Chinese Academy of Sciences, Beijing, Chian
[2]Institute of Software, Chinese Academy of Sciences, Beijing, China
Corresponding author : Yun Hu (huyun2016@iscas.ac.cn)

Abstract

Named entity recognition (NER) often suffers from lack of annotation data. Multi-domain and multi-task learning solve this problem in some degree. However, previous multi-domain and multi-task learning are often studied in English. In the other part, multi-domain and multi-task learning are often researched independently. In this paper, we first summarize the previous works of multi-domain and multi-task learning in NER. Then, we introduce the multi-domain and multi-task learning in Chinese NER. Finally, we explore the universal models between multi-domain and multi-task learning. Experiments show that the universal models can be used in Chinese NER and outperform the baseline model.

Introduction

In this paper, we explore the performance of the multi-domain and multi-task methods in Chinese NER task. The source domain is Chinese news domain, and the target domain is Chinese weibo domain. The source task is CWS, and the target task is NER. The first block is the Chinese-English translation pair for understanding. The second block is from weibo NER. The third block is from weibo CWS. The fourth block is from news NER.

嘿嘿,坐标洛阳,想看岳云鹏说的洛阳机场天价面
Heihei, in Luoyang, want to see the high-price noodles which Yue Yunpeng ate before in Luoyang Airport

副省长张大伟一行来到洛阳,视察洛阳机场航站楼扩建工程
Vice Governor Zhang Dawei arrived in Luoyang and checked up on Terminal Extension Project of Luoyang Airport

嘿 嘿 ， 坐 标 洛 阳 ， 想 看 岳 云 鹏 说 的
O O O O O B-GPE I-GPE O O B-PER I-PER I-PER O O
洛 阳 机 场 天 价 面
B-LOC I-LOC I-LOC I-LOC O O O

嘿嘿 ， 坐标 洛阳 ， 想看 岳云鹏 说的 洛阳机场 天价面
B E S B E B E S B E B I E B E B I I E B I E

副 省 长 张 大 伟 一 行 来 到 洛 阳 ，
O O O B-PER I-PER I-PER O O O O B-LOC B-LOC O
视 察 洛 阳 机 场 航 站 楼 扩 建 工 程
O O B-LOC I-LOC I-LOC I-LOC O O O O O O O

Previous works

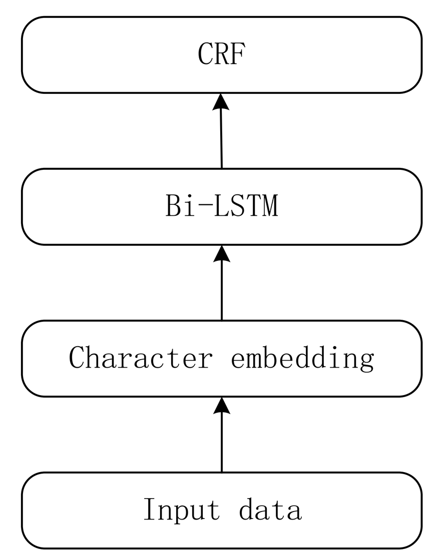
A summary of the multi-domain and multi-task learning. Details can be seen in paper.

	Multi-domain	Multi-task	Model
Yang et al. [26]	English	English	SHA
Peng and Dredze [18]	Chinese	Chinese	SHA
Lee et al. [8]	English	-	SHA
Collobert et al. [4]	-	English	SHA
Peng and Dredze [17]	-	Chinese	FEAT
Cao et al. [2]	-	Chinese	ADV
Peng and Dredze [16]	-	Chinese	BV
Changpinyo et al. [3]	-	English	BV
He and Sun [7]	Chinese	-	BV
Wang et al. [22]	Chinese	-	BV
Lin et al. [12]	English	-	BV

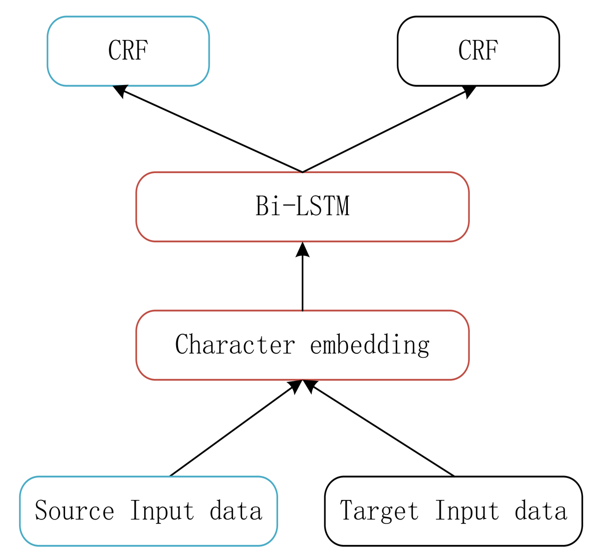
Model

All the models are composed by some basic modules. We discuss the basic modules first. Four types of modules are considered: Character embedding, Bi-LSTM, CRF and Classifier.

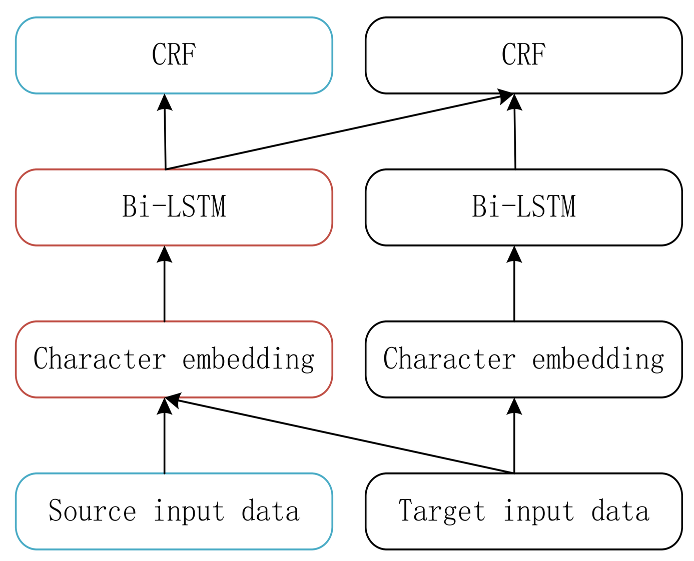
The multi-domain and multi-task models are divided into four types: SHA, FEAT, ADV and BV (variant of base model). The base model and three types of architectures in multi-domain and multi-task learning are shown below. The blue block is source part, the black block is target part, and the red block is the share part.



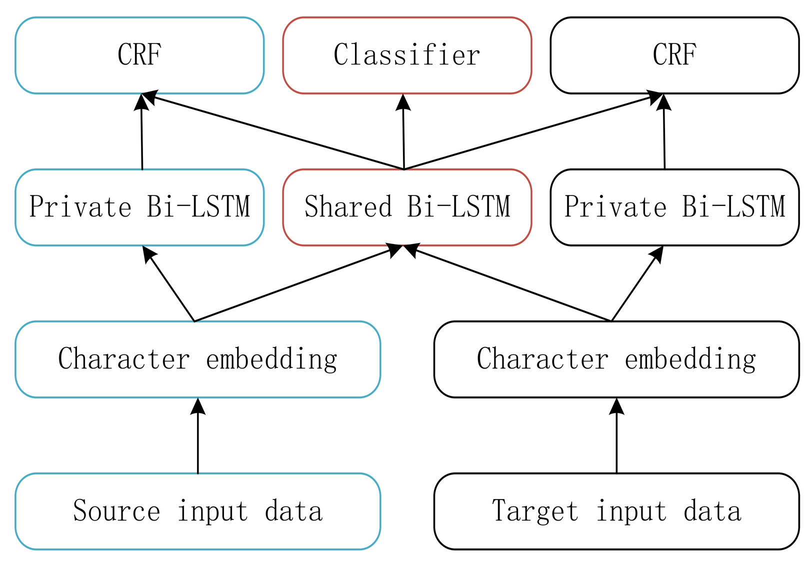
(a) Base model



(b) SHA model



(c) FEAT model



(d) ADV model

Results

The overview results of multi-domain and multi-task learning in Chinese weibo NER. P represents precision, R represents recall, and F represents F1 score.

	Multi-domain			Multi-task		
	P	R	F	P	R	F
Base	56.07	44.50	49.62	56.07	44.50	49.62
SHA-INIT	58.06	41.28	48.25	62.00	42.66	50.54
SHA-CRF	65.52	17.43	27.53	69.49	18.81	29.61
SHA-MUL	61.64	46.39	52.94	60.65	48.45	53.87
FEAT-INIT	53.55	49.11	51.23	62.93	44.13	51.88
FEAT-CRF	59.64	45.41	51.56	62.24	46.77	53.41
FEAT-MUL	59.86	45.36	51.61	61.39	50.00	55.11
ADV	57.06	52.06	54.45	60.92	47.42	53.33

The overview results of multi-domain and multi-task learning in Ontonote dataset.

	Multi-domain			Multi-task		
	P	R	F	P	R	F
Base	48.23	44.60	46.34	48.23	44.60	46.34
SHA-INIT	48.39	43.45	45.79	47.84	44.26	45.98
SHA-CRF	36.34	25.17	29.74	30.31	19.43	23.68
SHA-MUL	59.78	55.24	57.42	62.93	44.85	52.37
FEAT-INIT	52.79	49.49	51.09	51.56	47.70	49.55
FEAT-CRF	51.90	51.15	51.52	52.45	47.75	49.99
FEAT-MUL	53.32	58.37	55.73	57.24	45.52	50.71
ADV	55.40	55.60	55.50	50.56	50.47	50.51

Conclusion

In this paper, we focus on utilizing Chinese news domain information and Chinese word segmentation information to improve the performance of Chinese weibo named entity recognition by multi-domain and multi-task learning. Three types of universal model architectures are explored. Experiments show that the universal models outperform the baseline model.

ManuscriptID:299