

作业 4 – RNN+CRF 序列标注练习

命名实体（中文人名、地名、组织机构名）识别

任务：用序列标注方法将文本中的人名、地名和组织机构名识别出来。

模型参数：BILSTM+CRF

词向量维数：300

隐藏层维数：512

优化器：resprop

初始学习率：0.001

学习率衰减因子：0.95

batch_size：64

dropout：0.3

clip：5

```
30     nepochs = 100
31     dropout = 0.3
32     batch_size = 64
33     optimizer = 'rmsprop'
34     lr = 0.001
35     lr_decay = 0.95
36     clip = 5
37
38     lstm_dim = 250
39     word_dim = 300
```

测试结果：

第 64 轮结果

```
{'acc': 98.38905775075987, 'f1': 91.14850036576446,
'p': 93.4032983508246, 'r': 89.0, 'f1_per': 97.0,
'p_per': 97.48743718592965, 'r_per': 96.517412
93532339, 'f1_loc': 90.94076655052264, 'p_loc':
92.5531914893617, 'r_loc': 89.38356164383562,
'f1_org': 85.49618320610686, 'p_org':
90.3225806451612 8, 'r_org': 81.15942028985508}
```

Accuracy	F1-Measure	Precision	Recall
98.39	91.15	93.40	89.00

	F1-Measure	Precision	Recall
PER	97.00	97.49	96.52
LOC	90.94	92.55	89.38
ORG	85.50	90.32	81.16

第 82 轮结果：

```
{'acc': 98.52887537993921, 'f1': 91.42857142857142,
'p': 93.83458646616542, 'r': 89.14285714285714,
'f1_per': 98.2367758186398, 'p_per':
99.48979591836735, 'r_per': 97.01492537313433,
'f1_loc': 90.45936395759718, 'p_loc':
93.43065693430657, 'r_loc': 87.67123287671232,
'f1_org': 86.06965174129351, 'p_org':
88.71794871794872, 'r_org': 83.57487922705315}
```

Accuracy	F1-Measure	Precision	Recall
98.52	91.43	93.83	89.14

	F1-Measure	Precision	Recall
PER	98.24	99.49	97.01
LOC	90.46	93.43	87.67
ORG	86.07	88.72	83.57

每一轮总的损失值[0:81]

[5363.324743099997, 2241.2641406999996, 1557.805051,
1175.8519952000017, 931.7302836599993,
763.7542083500007, 654.99399003, 560.7510163900004,
496.6013199000005, 441.32196952999976,
392.9374539899999, 356.90063034999986,
328.8622702699997, 298.13284129, 277.85264216999985,
249.88952663999996, 234.52279871999988,
214.68443945799996, 204.15062416199993,
191.58167880199977, 175.17529298499986,
167.745009143, 159.3090481920001, 143.9641921859999,
135.003840890000008, 131.21411750399992,
124.51624749599995, 120.99739265999997,
107.196758751, 104.965666666, 99.70549397800008,
96.409844949, 91.63145081999993, 89.15878342600003,
81.90139181499997, 79.946598631, 79.78545710579999,
80.9829309145, 75.42270009499997, 70.85089436400004,
68.52049107699997, 68.57820312049998,
64.18227634670001, 65.53686197, 60.8846807282,
61.79351916220003, 61.28441696069999,
60.77538780699997, 64.12499822100006,
56.77893356670002, 56.01299447759999,
56.40532897200001, 51.48972347369996,
54.07328272950007, 49.39237267300003,
53.02695868190002, 52.94683449829997,
50.473966278300026, 49.938738740699996,
48.566684569999985, 47.20052815300005,
45.122743040000024, 48.34275493550001,
48.69931023709999, 50.244189399499994,
48.20232153289999, 45.31293331049999,
47.417113746000005, 45.1993615451, 43.57401853380004,
46.74579654700002, 45.217887618999974,
42.65489377216998, 40.99709191700003,
45.90022485416002, 44.218159106899996,
45.32957788770001, 43.0915286435, 42.32011698085997,
42.6450208128, 42.48845278719998, **41.39583528089999**]