

## CRF 和 LSTM+CRF

### 数据集：

Penn Treebank: 使用 Pos Tagging:

变量	calloway	N	NN
释义	原词	标记首字母（为方便程序运行）	标记

ConLL2000: 使用 chunking

变量	Confidence	NN	B-NP	NP
释义	原词	标记	短语块	短语类型

ConLL2003: 使用 NER

变量	EU	NNP	B-NP	B-ORG	ORG
释义	原词	标记	短语块	实体块	实体类型

### CRF 训练结果：

当 loss<0.0001 时结束训练

数据集	F1	Accuracy	Precision	Recall	备注
Penn Treebank	93.85	99.94	96.46	92.69	使用出现超过5次的标记
ConLL2000	70.58	99.58	72.91	71.16	短语块
	85.15	99.46	89.21	82.84	短语类型
ConLL2003	78.36	98.94	80.90	76.70	实体块
	82.22	98.14	84.41	80.70	实体类型

### LSTM+CRF 训练结果：

#### 1、论文要求

batch\_size=100, embedding\_size=50, learning\_rate=0.1, hidden\_size=300, epoch=10.

数据集	F1	备注
Penn Treebank	87.04	标记
ConLL2000	87.09	短语块
ConLL2003	68.88	实体块

#### 2、自己调参

数据集	F1	备注
Penn Treebank	93.46	Lr=0.001
ConLL2000	90.06	Lr=0.001
ConLL2003	88.77	Lr=0.001

不足:

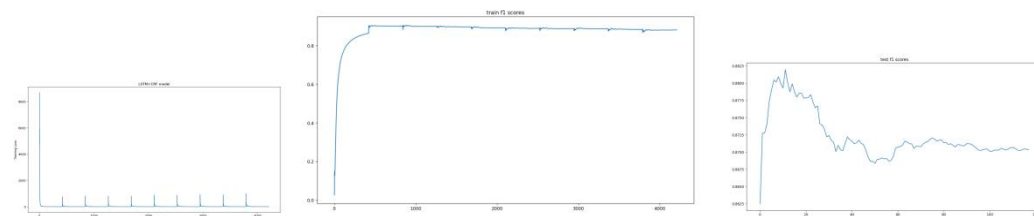
- 1、数据集没有按照论文要求进行 Random 和 Senna 词嵌入
- 2、调参只调整了学习率，没有调整其他参数

问题:

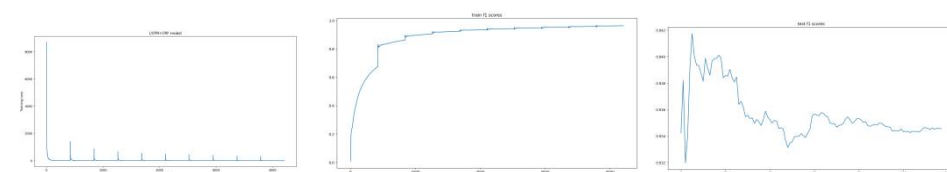
- 1、在梯度裁剪中，使用函数 `torch.nn.utils.clip_grad_norm_`，`clip_coef` 越小，则对梯度的裁剪越厉害。

已知每次循环的 `clip_coef`，怎么设置 `max_norm`，让对梯度的裁剪在合适的范围内？

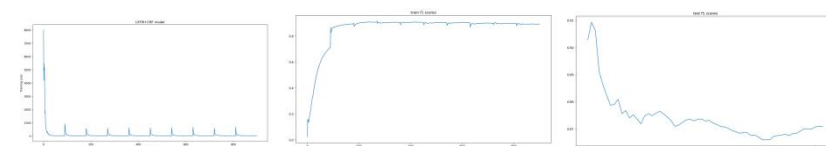
## Penn Treebank



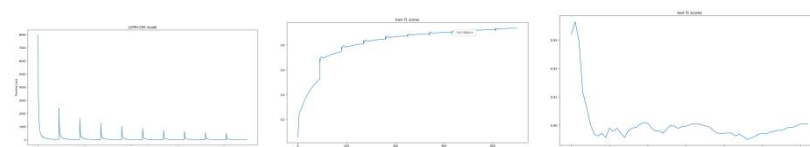
Lr = 0.001



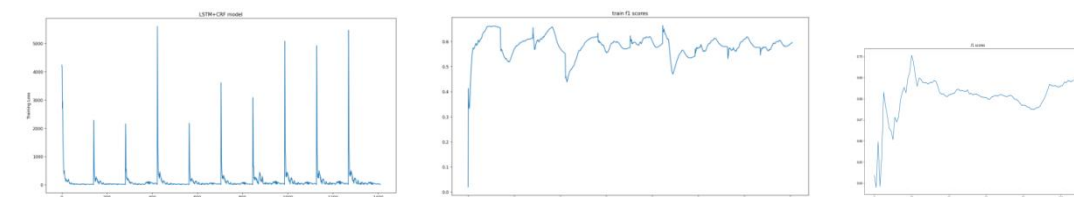
Conll2000 lr=0.1



Conll2000 lr=0.001



Conll2003 lr=0.1



Conll2003 lr=0.001

