

預訓練向量：

word2vec.py (產生預訓練向量檔案 word2vec.txt)

可調參數：

--train_data	設定使用資料
--output	設定輸出位置
--dim	設定向量維度

Example：

```
python word2vec.py --dim 200
```

NER 程式：

main.py (主程式)

model.py

eval.py

data.py

utils.py

conlleval_rev.pl

可調參數：

--train_data	設定使用的訓練資料
--test_data	設定使用的測試資料
--batch_size	設定批次訓練大小
--epoch	設定遍歷次數
--hidden_dim	設定隱藏層向量維度
--optimizer	設定最佳化方法
--lr	設定學習速率
--clip	設定梯度裁剪參數
--dropout	設定 dropout 參數
--update_embedding	訓練時是否更新 embedding
--pretrain_embedding	設定使用的預訓練向量
--embedding_dim	隨機向量的維度
--shuffle	訓練時是否 shuffle
--mode	模式選擇(train/test/demo)
--demo_model	設定要 demo 的模型

需要的檔案：

train_data	訓練資料(必要)
test_data	測試資料(必要)
word2vec.txt	預訓練向量(非必要)

Example：

```
python main.py --mode train --batch_size 16 --epoch 5 --hidden_dim 200 --optimizer Adam --lr 0.005 --clip 4 --dropout 1
```

程式流程：

1. 產生 'word2id.pkl' 檔，主要目的為將每個字給定 ID
2. 取得 'word2id.pkl' 每個字 embedding
使用向量：(1) 隨機 (2) 預訓練
3. 取得標籤(BIO)
4. 讀資料
5. 設定路徑
6. 三種模式
 - (1)train：使用 train_data 訓練，test_data 測試
 - (2)test：使用 test_data 訓練
 - (3)demo：自行輸入句子