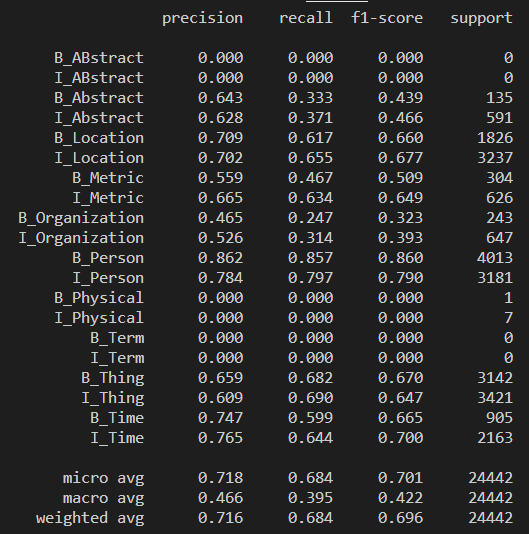
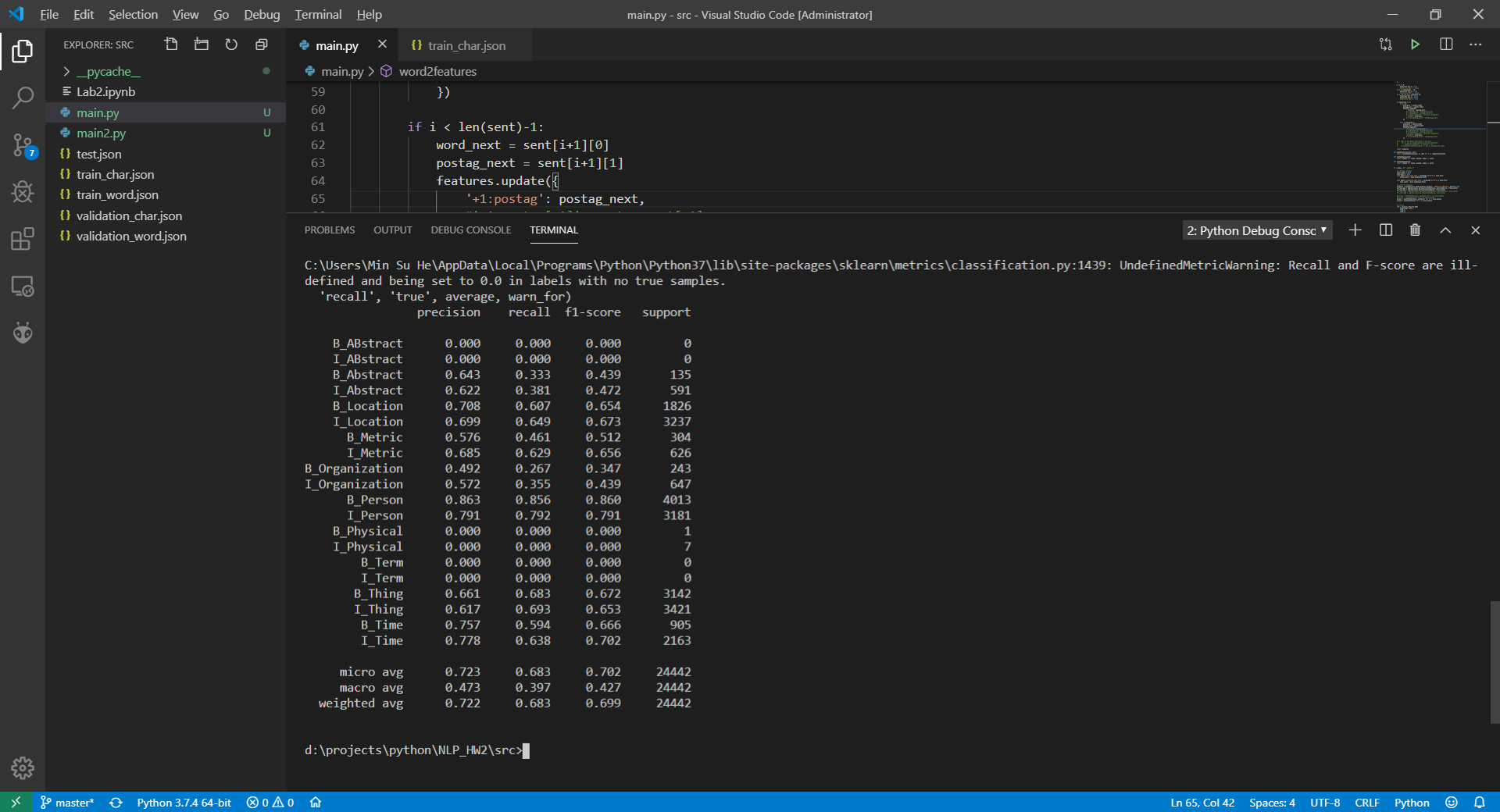
NLP HW2

105062271 何旻書

1. **問題背景**  
   由於NER（Name Entity Recognition）問題屬於一種Sequence labeling問題  
   ，即給定特定序列，必須為序列中的元素做標記、分類。我們常使用各種機率模型來求解，如Markov Model、HMM（Hidden Markov Model），在本作業中，我們使用CRF（Conditional Random Field）來解決NER問題。
2. **程式說明**  
   main.py架構主要有四個步驟:   
    1. 定義中文字的feature   
    2. 利用feature來將dataset轉為可訓練和驗證的list  
    3. 使用sklearn\_crfsuite的CRF模組來訓練  
    4. 利用validation資料來評估模型精準度  
   其中第一個步驟中，有四種方式做特徵抽取：範例使用的feature、範例使用的feature修改版，以及兩種word2vec模型
3. **運行結果**  
   1. 使用範例預設的feature（開始實作之前）  
   

2. 使用自行修改範例的feature



1. 比較與討論
2. 結論