**Natural Language Processing and Text Mining: HW#2**

**小組成員和負責工作：**

**109590041 范遠皓 測試和報告**

**109590043 柯瑞霖 撰寫程式**

**環境**

使用的語言：Python

所需套件：

click==8.1.3

colorama==0.4.6

joblib==1.2.0

nltk==3.8.1

numpy==1.24.2

pandas==1.5.3

python-dateutil==2.8.2

pytz==2023.2

regex==2023.3.23

scikit-learn==1.2.2

scipy==1.10.1

six==1.16.0

threadpoolctl==3.1.0

tqdm==4.65.0

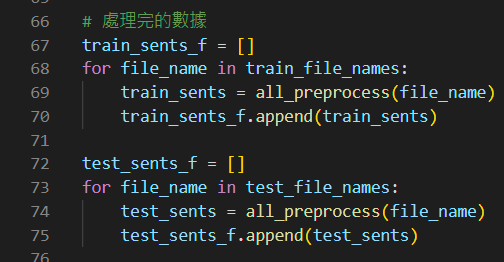
安裝辦法：

Zip檔內有requirements.txt

終端機輸入以下指令安裝

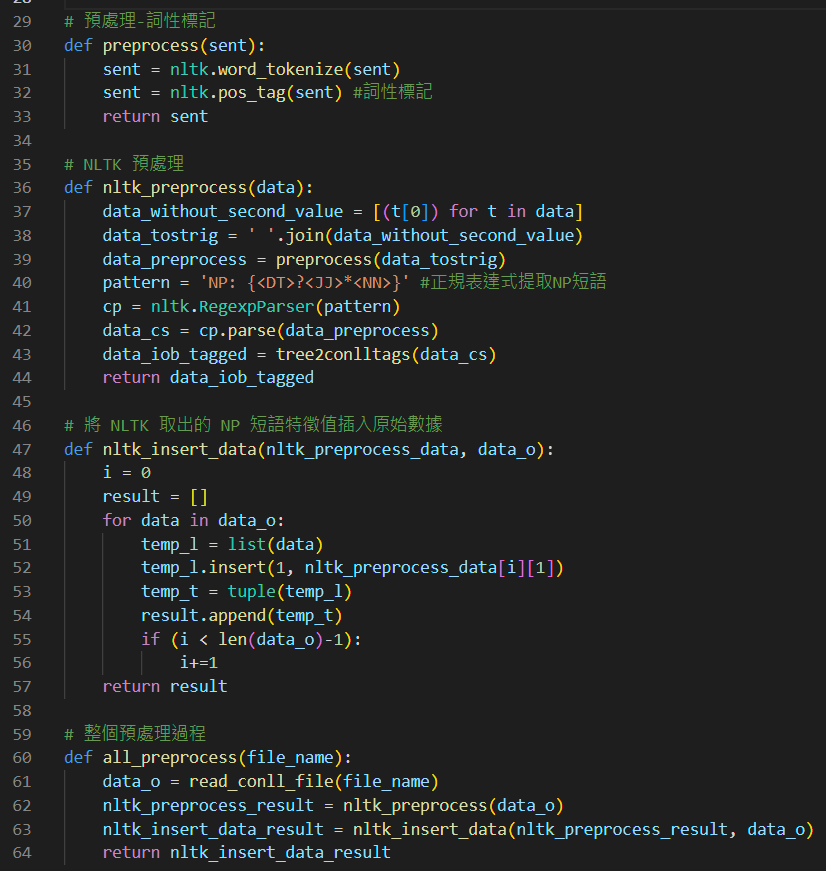
pip3 install -r requirements.txt

資料讀取並進行前置處理

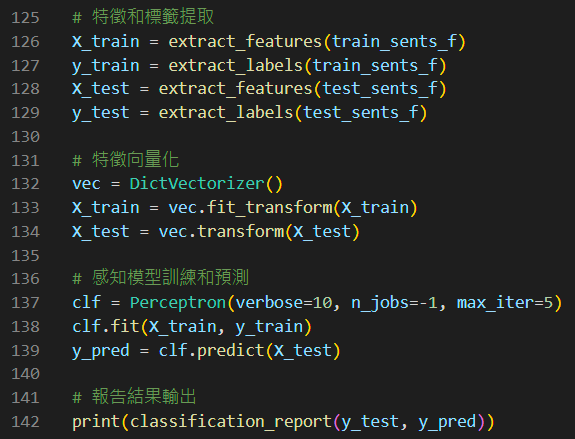


NLTK 前置處理:

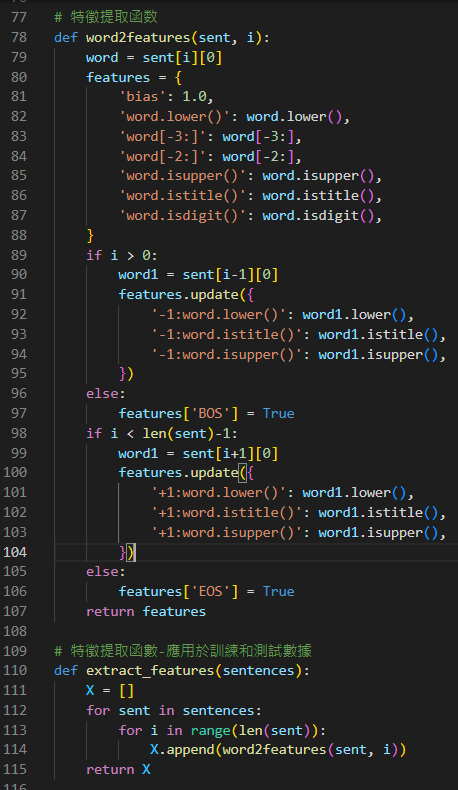
對資料進行詞性標記，然後提取其中的名詞短語 (NP)，並將其特徵值插入原始資料中。



提取特徵後放入 Perceptron 模型進行分類訓練，最後輸出結果。



特徵提取



執行結果：

