

1. 執行環境：jupyter notebook
2. 程式語言：python 3.12.4
3. 執行方式：學生自己是透過 anaconda 跑 jupyter notebook 的，檔案類型也是 ipynb 檔，無須預先安裝其他的外部套件，這次的作業程式碼中也沒用到。
4. 作業邏輯說明：原先是直接透過 `.split()` 函式將句子切成一個個 token，但後來發現這樣標點符號會跟字黏在一起(ex: I have an apple. 會變成 I/have/an/apple.)，所以先將標點符號去除，接著按照題目要求將所有單字變成小寫，然後自訂了簡單常見的 stopword 清單去除沒用的字。接著按照 Porter's Algo. 設計一個 stemming 的方法，把 token 還原成其字根，這裡遇到了一些問題讓輸出的結果不那麼完美，第一個便是縮寫的部分(ex: I'm 跟 don't 不知道該不該加一個判斷是去硬比對，感覺這樣很粗魯)，第二個便是去 e+ing 的動詞不知道該如何判斷其字根要加回 e，最後就是 ed 如果直接去掉會把原本結尾就是 e 的動詞破壞。最後根據 terms 的定義把每個抓出來的結果用 set 表示其唯一性，並順便根據首字母做了 index 輸出結果。