Programming Assignment 1

- Write a program to extract terms from a document.
  - You have to do:
    - Tokenization.
    - Lowercasing everything.
    - Stemming using Porter's algorithm.
    - Stopword removal.
    - Save the result as a txt file.

  - Text collection:
    - One English news document
      (https://cool.ntu.edu.tw/files/8065606/download?download_frd=1).

  - **Please zip and submit [1.]the result of document 1, [2.]the source code, and [3.]a report to TA.**
    - 2 weeks to complete, that is **2025/9/22**.