I have mainly created functions for each step to reuse it in the code parts.

First, I collected data from the URL given using Beautifulsoup

Second, I stored them in a folder

Third, I cleaned and then used stop words and stored them in other folder

Fourth, I have cleaned again with punctuation and stored them in a folder

Fifth, every file is saved with the URL name, which is used later in the Excel section.

Sixth, I have formulated all scores to give the scores

Seventh, use other library openpyxl, I have used for reading and writing data into Excel

I have used the ipynb format for easy code-making because spitting codes and watching their output is easier. I have completed the process in ipynb and converted it in py. Please use ipynb for better output

Please remove ‘’|’ in stops word because the code cannot read them for running purposes I will upload the complete set please run through my downloaded folder

In dependencies; there are my old dependencies. Please exclude them

Finally, it took much effort to debug, structure and try error methods sometimes to get the desired output.