INSTRUCTION

This is code that performs text analysis on a list of URLs. The code performs the following tasks:

1. Imports required libraries including requests, BeautifulSoup, TextBlob, pandas, nltk, string, enchant, textstat and re.
2. Downloads necessary datasets using nltk.download.
3. Import the URLs to be analyzed from an Excel file named “Input.xlsx”.
4. Reads six text files with words and phrases to be used for positive and negative score calculation, and for filtering out stop words from the analyzed text.
5. Combines all the stop words from all six files into one list.
6. Performs a loop for each URL in the list and does the following:
7. Sends a request to the URL to retrieve its HTML content.
8. Parses the HTML content using BeautifulSoup library.
9. Calculates the positive score, negative score, average sentence length, complex words and Gunning fog index.
10. Cleans the text by removing stop words and punctuation.
11. Stores the results in a list named ‘output’.
12. Writes the results of the analysis to an output file.