**Here are the steps to get the output from the provided code:**

**1. Set Up Environment:**

* Ensure you have Python installed, along with necessary libraries: requests, BeautifulSoup, pandas, nltk, textstat, and dotenv. Already mentioned in requirement.txt
* Install them via pip: pip install -r requirement.txt

**2. Prepare Files and Directories:**

* Create the required directories:
  + Input/ to store your input files (Input.xlsx, stopwords, positive, and negative word lists).
  + Output/ to store the final results.
  + extracted\_file/ to store the extracted text files from URLs.
* Place the following files in the Input/ directory:
  + **Input.xlsx**: Contains URLs in a column named URL and unique identifiers in the column URL\_ID.
  + **all\_stopwords.txt**: Contains the list of stopwords.
  + **positive-words.txt**: Contains a list of positive words.
  + **negative-words.txt**: Contains a list of negative words.

**3. Create the .env File:**

* In the root folder, create a .env file and set the paths to input/output files as follows:

INPUT\_FILE\_PATH = 'Input/Input.xlsx'

STOPWORD\_FILE\_PATH = 'Input/StopWords/all\_stopwords.txt'

POSITIVE\_FILE\_PATH = 'Input/positive-words.txt'

NEGATIVE\_FILE\_PATH = 'Input/negative-words.txt'

OUTPUT\_FILE\_PATH = 'Output/'

EXTRACTED\_FILE\_PATH = 'extracted\_file/'

**4. Run the Code:**

* **Download necessary NLTK data**: You need to download the required NLTK resources (e.g., tokenizers, stopwords). Run the following Python script once:

*import nltk*

*nltk.download('punkt')*

*nltk.download('stopwords')*

* **Run the main script**: Execute the app.py file to start the process:

*python app.py*

**5. Output:**

* The program will:
  1. **Extract data from URLs** and save them as .txt files in the extracted\_file/ directory.
  2. **Perform text analysis** on the articles with and without stopwords.
  3. **Generate two Excel files**:
     + Output.xlsx: Analysis results with stopwords.
     + Output\_NoStopwords.xlsx: Analysis results without stopwords.

Both Excel files will be saved in the Output/ directory.

**NOTE : The above approach is for new input data. If no new data is provided, simply run the app.py file to get the result by using: “python app.py” (AFTER INSATLLING THE REQUIREMENT.TXT FILES)**