

Problem Statement:

Task: Perform text analysis using web scraping with file management.

Libraries Used:

Utilised BeautifulSoup and requests for effective web scraping.

Code Issue:

Main content extraction primarily relies on the `<div>` tag with class "td-post-content tagdiv-type," except for one case where it's under "tdb_single_content."

Solution:

Implemented conditional statements to handle both cases based on the presence of specific classes.

URLs Not Found:

Two URLs (<https://insights.blackcoffer.com/covid-19-environmental-impact-for-the-future/> and <https://insights.blackcoffer.com/how-neural-networks-can-be-applied-in-various-areas-in-the-future/>) encountered issues.

Solution:

Verified URL correctness and web page accessibility.

Ensured that HTML structure aligns with expectations.

Error Handling:

Encountered a division by zero error due to zero sentence length during text analysis.

Solution:

Implemented try-except blocks to gracefully handle division by zero errors.

Provided a mechanism to skip processing or set default values for zero-length sentences.