

CIT 44400/CIT 58100

Fall 2024

Homework 6

Use the NLTK python library to do the following tasks on your document and display the output:

1. Remove stop words. (10 points)

Example output of stop word removal for the following sentence may look like this:

Sentence: "The cat is sitting on the mat."

Output: "cat sitting mat."

2. Perform Lemmatization (using NLTK's WordNetLemmatizer or other) (10 points)

Below is an example of lemmatization outputs:

Sentence: "The cats are running."

Output: "The cat be run."

3. Perform sentiment analysis at document level, paragraph level, and sentence level. (30 points)

Note: consider applying stop word removal, lemmatization, and other NLP technique to improve sentiment analysis results.

Bonus: After performing sentiment analysis, try to build a word cloud for your document. (10 points)