CS 457 - Homework Assignment 13: NLP

Due Date: Monday, April 17 at 11:59 pm

Purpose:
Demonstrate understating of basic NLP techniques for text data

Points: 100

Deliverables: Submit ipynb code file along with your answer

- Perform NLP analysis on the following questions**.** Make sure to include interpretation of each result including text/visualizations/tables to support your answer

    1. Select any article on the website of your interest and perform Name Entity Relation (NER) analysis on it. Show all the necessary output and visualization and explain your findings. Also mention if you see any shortcoming or error in the algorithm output.
    2. Select any research paper of your interest and perform word frequency and keyword extraction analysis. Use at least 3 different techniques/libraries. Show all the necessary output and visualization and explain your findings. Compare different techniques and mention if you see any shortcoming or error in the algorithm output.
    3. Choose any keyword of your interest and pull at least 1000 (tweets) from Twitter.
        a. Clean the tweets and store them in csv file.
        b. Perform Sentiment Analysis on the tweet and show total positive and negative sentiment counts.
        c. Display one Positive and one Negative sentiment and discuss if the algorithm is accurate.
        d. Display the tweet with highest retweets and show its sentiment.
        e. Make a timeline visualization plot showing counts of positive and negative sentiments over the period of time.