Project Progress Report CS-410 / Fall 2021
JunyangWang

Which tasks have been completed?
1. Set up Environment, made requirements.txt file and
installed compatible libraries.
2. Data Pre-Processing, removed stop words, removed Latin words.
Tokenized words and removed punctuations. Performed POS tagging and
extracted adjectives and nouns.

Which tasks are pending?
1. To execute LDA and extract topics.
2. To perform sentiment analysis on each topic at a dataset level.
3. Code cleanup and Documentation.

Are you facing any challenges?
1. Library incompatibilities with Python Version.
2. Had some trouble doing data cleaning implementation. Extracting
   meaningful words and Mapping lemmatization functions to every row in
   data frames
3. Facing some encoding error when work with datasets.