CS410 Project Progress Report

Team Name: ThePredictive

Team Members:

Vedant Vipul Jhaveri (vedantj2) - Captain

Sagar Dalwadi (sagardd2)

Sindhu Kopparam (sgk6)

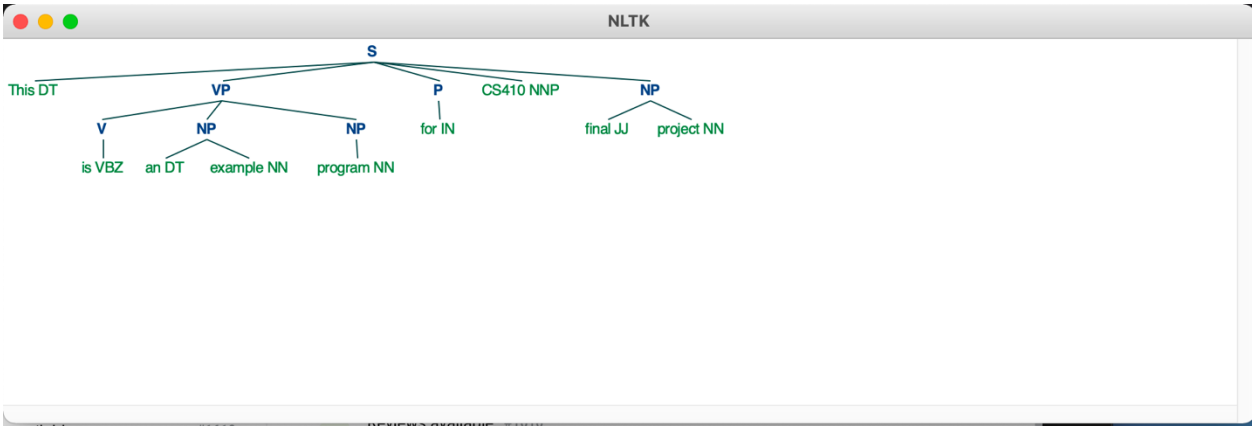
1. Which tasks have been completed?

Our project proposal was to integrate a keyword extraction function in metapy similar to the one existing in SpaCy and integrate chart parser using NLTK.

So far we have completed below tasks for respective topics:

Integrate Chart parser using NLTK:

- Completed a research with respect to implementing Parse tree using NLTK library.
- Installation of NLTK and getting familiarized with the libraries NLTK offers to implement a parse tree
- Implemented a sample python program to integrate the NLTK library that displays the parse tree for a given statement. (Here is a screenshot of an output result of the sample program that was implemented)



Integrate Key word extraction in Metapy

- Explored keyword extraction function in SpaCy
- Completed a research around how to implement a domain specific keyword extraction
- Partially implemented domain specific extraction

2. Which tasks are pending?

We will now focus on the below remaining tasks:

Integrate Chart parser using NLTK:

- Research on how a program that uses an NLTK library to display a parse tree can be integrated with MeTA
- Implementation of an integration between python program containing NLTK library and MeTA toolkit
- Preparation of data set that can be used for testing
- Testing & validation of final outcome
- Documentation & presentation

Integrate Key word extraction in Metapy

- Debug partially implemented code to further refine
- Preparation of data set that can be used for testing
- Testing & validation of final outcome
- Documentation & presentation

3. Are you facing any challenges?

- The major challenge with respect to integrating chart parser using NLTK with MeTA toolkit is the lack of documentation around enhancing MeTA toolkit and integrating with other natural language processing tools. This requires more research time than expected at the beginning.
- The challenge with the key word extraction is to gather & create data to process and test for the code
- And another challenge with respect to keyword extraction is the performance evaluation of the function in metapy when compared to spacy