**CS410 Progress Report**

“TIS Chicago”

https://github.com/JosephineFalso/CourseProject

Josephine Falso

Sree Alaparthi

Narasimha Kethireddy

Within the topic of Intelligent Learning Platform, our group is building a back-end application to integrate the instructor’s lectures with the student discussions in CampusWire. Our application will highlight significant keywords in the transcription text of the Coursera videos, and the hover text will provide some additional information from the CampusWire posts related to that keyword.

Our group has decided to start with the content from week 3. There were several significant course topics discussed during that week, including precision, recall, average precision, F measure, MAP, gMAP, DCG, and nDCG. There are many postings in CampusWire related to these topics from the weekly discussions and one of the MP assignments. If time allows, our group hopes to expand the scope of the project to include additional weeks of course content.

The web scraping of the Coursera transcription text is in progress. Once a file of the transcription text has been generated, the Python library NLTK will be used to eliminate the words in the background language model and to generate the most popular keywords from the lectures. Although this task is pending, another teammate has extracted the data from CampusWire for a few of these keywords by hardcoding the terms. We have written a Python program to parse the json structures extracted from CampusWire and extract the topic and url of the posting. We then generated a csv file of the topic, url, and ranked relevance of the posting to the keyword. We have deployed the Coursera videos and transcription text to a new webpage for some of the videos in week 3. The tasks of highlighting the keywords and populating the hover text remain outstanding. For presentation of the final code delivery, we will create a Node.js application deployed on a Docker container.

Our team members have various levels of competency with Python, so the development is progressing at an inconsistent pace. Some of the concepts in this assignment are outside of the scope of the course material and require additional research. Additionally, tasks such as highlighting of the keywords in the transcription text and display of the hover text are new concepts for the team and require additional time to implement.

The reviewers gave the feedback that we may need to consider textbook data, Wikipedia data, or other datasets to populate the hover text if sufficient postings cannot be found on CampusWire. Our group hopes to expand the scope of the project to incorporate this feedback if time allows.