CS410 Final Project Progress Report

1. Project Team - American River

- Maximus Choi (daechoi2@illinois.edu)
- Jackson DeDobbelaere (dedobbe2@illinois.edu) – Captain
- Subramanian Shankar (ss117@illinois.edu)

2. Progress Made

For the final project, we are working on reproducing research paper Latent aspect rating analysis without aspect keyword supervision by Hongning Wang, Yue Lu, and ChengXiang Zhai. 2011. We created git repository (https://github.com/JackDeDobb/CourseProject) for common workspace. We together reviewed the paper and created strategy to divide tasks among the team members.

We collected source data sets to analyze (TripAdvisor review data set and Amazon MP3 Player review data set) from the UIUC Database and Information Systems Laboratory (http://sifaka.cs.uiuc.edu/~wang296/Data/index.html). We conducted data curation by parsing data, removing the reviews with bad data quality, converting the two data sets to a consistent format, and removing stop words and unnecessary terms and punctuations.

Using Python, we developed Latent Aspect Rating Analysis (LARA) models to ingest the data sets and produce aspect rating predictions.

3. Remaining Tasks

We will continue to debug and complete our LARA models and produce results similar to the research paper results.

Upon completing our LARA model validation, we will complete the final written report, describing our project and user guide's process and results to use our solution. We will create a video to demonstrate how to run our solution and produce results.

4. Any challenges being faced

Because the LARA model code development requires a comprehensive understanding of the algorithm and tight dependencies between functions, it has been challenging for multiple team members to modularize, distribute, and debug code development tasks effectively. To mitigate this challenge, we serialize our development and debugging process, tightly coordinate the team member’s time, do frequent knowledge transfer, and hand over tasks to the next available team member. Once we complete our LARA models debugging and validation, we can efficiently distribute the remaining tasks - creating a final report document and a video.