

CS410

Project Progress Report

Ge Yu - gey2

Chengmin Huang - ch61

Xuehao Wang - xuehaow2

- Progress made:
 1. Understand the advantages of Latent Aspect Rating Analysis Model compared to previous two-step solution
 2. Data from TripAdvisor has been processed into SQL database
- Remaining tasks:
 1. Try to pull more datas from the website to make the model more meaningful
 2. Algorithms to calculate the K-means from different features
 3. Still try to implement algorithms described in the paper
- Challenges:
 1. Fit the model with the bag-of-word assumption
 2. How to determine the aspect segments and their weights
 3. Acceleration on the training process for the test dataset
 4. Dimensionality reduction to solve the data sparsity