EECS E6895 Adv. Bigdata Analytics - Final Project Proposal

Sun-Yi Lin (sl3833), Hao-Hsiang Chuang (hc2751)

Visualization of Yelp's Academic Dataset

Yelp is a famous recommendation service base on varies businesses, it helps people to find great local businesses including restaurant, dentists, hair stylists, mechanics, etc. It was founded in 2004 and now it has 135 million monthly visitors and 71 million reviews.

Motivation

When we first discuss what to do in the final porject, visualization quickly become our consensus. Graph is a powerful tool that could make complicated records easy to read and understand.

Yelp's Academic Dataset is the best choice to implement visualization. It provides all the data and reviews of the 250 closest businesses for 30 universities for students and academics to explore and research, and we can do so much manipulations on it.

We want to provide a intersting virtualization application for the Yelp's Academic Dataset.

Dataset, Algorithm, Tools

Dataset: [Yelp's Academic Dataset](#)

Algorithms: K-means clustering, PLA classification, LDA, etc.

Tools: Java, Python, Javascript, HTML, Eclipse, Amazon Web Services, etc.

Required algorithms and tools could be changed base on actual necessary.

Progress and Expected Contributions

Progress: Famliarize ourselves with dataset and tools.

Expected Contribution : An interested, interactive and open source application that visualize the Yelp's Academic Dataset.