Using Star Ratings as Time-Sensitive Performance Metrics

While Yelp star ratings are known to substantially boost business for restaurants, cafes, etc., they only indicate the overall popularity of that business since it joined Yelp. An overall rating does not explicitly provide information regarding improvement or even deterioration in critical aspects such as food quality. This is compelling research since this ultimately works towards a streamlined and most transparent approach to sharing business performance in categories that the majority of users value and that will, hopefully, be reflected in the app one day.

Research Context

Past research has focused on extracting latent subtopics from review text using Online LDA and subsequently predicting star ratings from those different subtopics. Modified LDA algorithms have also been devised to better integrate semantic analysis into generating such subtopics. This past research provides significant information of what aspects of restaurants matter most to customers and even sheds light on potential ways for restaurants to improve. However, charting progress as well as deterioration of restaurant performance in different subcategories has not thoroughly explored and is not immediately evident to a user.

Data / Design

In the business json data, there are 38059 businesses with multiple business ids and therefore multiple reviews, which we can examine to find time-sensitive performance changes (using the "stars" and "date" columns under the review json). The business id tag is the only tag that overlaps across the business and review json objects, so this would act as a link between dataframes. The amounts of data at our disposal for particular restaurants over time is a bit inconsistent—i.e. an eatery might have multiple reviews but might not have reviews across many different years/months. This prompts us to carefully group the subgroups (i.e. "coffee shops", "Asian food") that we would use LDA on to avoid analyzing businesses that are "data-rich" in different aspects. In this research project, online LDA would be implemented for more specific subcategories of Yelp data (than previously used) and we would then extrapolate from the resulting star ratings by interpreting a restaurant's progress via star rating over time.

Methods

Unlike past analyses that ran LDA across the entire Yelp dataset without discriminating based on type of restaurant, industry, etc., we aim to break the current dataset into broad subcategories: coffee shops, pizza parlors, etc. and conduct LDA on those subgroups. This would lead to more specificity in the generated word distributions and subtopics. Next, predicted star ratings of the new topics generated are produced via a simple average of reviews of that particular subtopic. (An alternate method could be using positive and negative word weights, which could be used for cross-checking). To determine restaurant performance in one of the subtopics over time, we would conduct a time series analysis to assess how the ratings have changed. For any restaurant, the star ratings from a particular subtopic would be binned (i.e. a month) to reduce noise and plotted accordingly.

Significance

Even though browsing through Yelp reviews can give a potential new customer an idea of whether or not he/she would like to try a specific eatery, it is not easy for users to tell whether Yelp star ratings truly reflect the current performance of the restaurant or whether the current rating is slightly misleading. Across subcategories for a restaurant given its subgroup, users would have a better idea of overall quality. This not only benefits consumers but also the restaurant as well, giving it a better, time-sensitive guide as to how it is doing in regards to the categories users care about the most. This certainly helps a restaurant's business strategies by exhibiting more transparency in performance heuristics.

References: 1. Personalizing Yelp Star Ratings: a Semantic Modeling Approach (J. Linshi) & 2. Improving Restaurants by Extracting Subtopics from Yelp Reviews (J. Huang, S. Rogers, E. Joo)