Usefulness Predictions of Yelp Reviews and Review Aggregation

Christian Clark, Spencer Gray, Ryan Conley, Peter Schaldenbran

**Introduction**

Restaurant reviews posted on the website “Yelp” are given a “usefulness” score by the users, where each user can choose to assign a single point into a review’s usefulness. We look to identify and predict the patterns inherent in useful reviews, and apply this model in predicting the usefulness of reviews, perhaps most appropriately to establishments with a sparse amount of data. A possible further utilization of this model is the aggregation of reviews based on weighted usefulness scores to determine the overall characteristics of a restaurant, deemed a “meta-review”.

**Problem Definition, Solution, and Evaluation**

The initial problem of extracting the useful data from the Yelp data set is posed such that many reviews have such sparse amounts of data (reviews, ratings, etc.) that their worth in training, developing, and testing our model is null. A threshold must be determined to separate the data into reviews that have enough data for use in training, development, and testing, and those too sparse to be of use until the supervised testing phase, wherein the accuracy of the model will be subjectively analyzed by a human to determine the model’s use in predicting the usefulness of reviews pertaining to restaurants with scarce feedback.

In order to train the model, the use of several features for a review will be used and weighted, such as n-grams of popular phrases, association of phrases with a type of restaurant, length of the review, positivity/negativity, formalness, and others. These can be properly weighted using the segmented development portion of the data set. The model can finally be tested on the final segmented portion of the extracted data set.

With time constraints taken into account, there is a second, more directly useful application of our model we wish to pursue in the meta-review. Of the remaining data set as described prior, using human verification, we seek to subjectively determine how accurate the meta-review is to the restaurant’s experience and its usefulness in previewing a restaurant. It is also our hope in undertaking this endeavor that some of the more common biases in restaurant reviews may be uncovered and accounted for.

**Milestones**

* Project Proposal finalized and submitted for review (3/1)
* Research tools and techniques to be considered for use in the project (e.g. NLP libraries and language modeling techniques) and defining their application therein (3/10)
* Division of responsibility within the group and other group affairs as well as basic implementation of scoring reviews’ usefulness (3/20)
* Completed model of scoring reviews’ usefulness and predictive abilities (4/10)
* Meta-reviews (4/20)[[1]](#footnote-1)
* Fine tuning of model parameters (4/27)
* Presentation (4/28)

**Related work**

Exploring the mechanisms behind the assessment of usefulness of restaurant reviews: <http://dl.acm.org/citation.cfm?id=2768557>. This team used interviews to figure out what makes a review feel useful to a person using Yelp.

Low-Quality Product Review Detection in Opinion Summarization: <https://aclweb.org/anthology/D/D07/D07-1035.pdf>. This team defined what a high-quality review of a product was. Then, using this definition, were able to filter out poor quality reviews to enhance opinion summarization.

The Social Aspect of Voting for Useful Reviews: <http://link.springer.com/chapter/10.1007/978-3-319-05579-4_36>. This team determined what criteria people use to rate a review. They devised a regression model that predicts the usefulness rating of reviews.

Finding Thoughtful Comments from Social Media: <http://www.aclweb.org/anthology/C12-1061>. This team measured how thoughtful a comment was in social media. They could determine the quality of the comment and the opinion of the person writing it.

1. Tentative milestone, subject to omission [↑](#footnote-ref-1)