Christian Clark

Spencer Gray

Ryan Conley

Peter Schaldenbrand

Project Proposal: Human Language Technologies

There are two main parts to the project we are proposing. The first, is to be able to determine the usefulness of each sentence within a review. Then, we will use this qualification to pull the most useful data from reviews for a business in order to create a more concise, useful review.

Evaluation:

We will use part of the data from the Yelp data set to train and develop our programs. We will use the rest of the data to test our programs. We will evaluate how accurate our rating of usefulness for a review is by comparing it to how other people rated the review. The evaluation of the meta-review will have to be subjective and based off of our own opinion. We will create examples of this meta-review, and ask people to determine if it actually compiled the most useful data from reviews of the business.

Tentative Timeline:

3/1 Project Proposal finished and submitted for review.

3/4 Research and analyze the related work, try to figure out what implementation tools we will need, and brain storm a way to implement quantifying the usefulness of a review over spring break on an individual basis.

3/15 Create a plan for programming the quantification of the usefulness of a review based on our individual research. Create a skeleton idea for the program, figure out who is doing what, and begin the implementation.

3/25 Have a functioning program that is able to quantify the usefulness of a sentence within a review. Begin using this program to create another program that will extract useful details from reviews.

4/14 Have a working program to extract useful details from reviews for a single business. Then compile those details into one meta-review. Then begin fine tuning everything and evaluating the results.

4/21 Have all programming and evaluation of results finished. Spend the next week working on the presentation.

4/28 Presentation.

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.