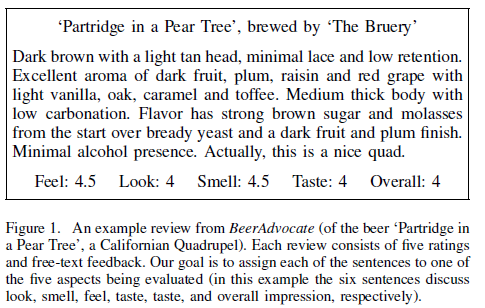
Learning Attitudes and Attributes from Multi-Aspect Reviews

I. INTRODUCTION



We consider three tasks on this type of data: First, can multi-aspect ratings be used as a form of weak supervision to learn language models capable of uncovering which sentences discuss each of the rated aspects? For example, using only multi-aspect rating data from many reviews (and no other labels), can we learn that ‘medium thick body with low carbonation’ in Figure 1 refers to ‘feel’? Moreover, can we learn that the word ‘warm’ may be negative when describing the taste of a beer, but positive when describing its color? Second, can such a model be used to summarize reviews, which for us means choosing a subset of sentences from each review that best explain a user’s rating? And third, since ratings for aspects are optional on many of the websites we consider, can missing ratings be recovered from users’ overall opinions in addition to the review content?

A. Present work

B. Contributions

The models we propose produce highly interpretable lexicons of each aspect, and their associated sentiments.

C. Further related work

II. DATASETS

III. THE PALE LAGER MODEL

PALE LAGER models aspects, and ratings on aspects, as a function of the words that appear in each sentence of a review.

We first introduce the notation used throughout the paper.

IV. LEARNING

A. Unsupervised Learning

B. Enforcing Diversity in the Predicted Output

C. Semi-Supervised Learning

D. Fully-Supervised Learning

V. LEARNING TO PREDICT RATINGS FROM TEXT

In many websites with multiple aspect ratings, ratings for aspects are optional, while only ‘overall’ ratings are mandatory. For example, our 10,989 Audible reviews represent only those where all three aspects (author, narrator, overall) were rated. In total there were 199,810 reviews in our crawl that included an overall vote but were missing an aspect rating. Predicting such missing ratings may help us to understand why users voted the way they did. We will learn models for this task from users who entered complete ratings.

VI. EXPERIMENTS

A. Review Segmentation

B. Review Summarization

In the context of our model, summarization means identifying a subset of sentences that best explain a user’s multiple-aspect rating.

1) Aspect Ranking

C. Rating Prediction

In many of the datasets we consider, only ‘overall’ ratings are compulsory while aspect ratings are optional. In this section we try to recover such missing aspect ratings.

D. Qualitative Analysis

VII. CONCLUSION

By introducing corpora of five million reviews from five sources, we have studied review systems in which users provide ratings for multiple aspects of each product. By learning which words describe each aspect and the associated sentiment, our model is able to determine which parts of a review correspond to each rated aspect, which sentences best summarize a review, and how to recover ratings that are missing from reviews. We learn highly interpretable aspect and sentiment lexicons, and our model readily scales to the real-world corpora we consider.