Extracting and Aggregating Aspect-Level Sentiment from Product Reviews

Matthew Long, Desmond C. Ong, Shane Soh
Stanford University
{mlong14, dco, shanesoh}@stanford.edu

Abstract

Previous work in *aspect-level sentiment analysis*—identifying the sentiment associated with products and their individual attributes, like battery life—have mostly been formulated as supervised learning problems, requiring known labels of both the relevant aspects and their sentiment. Here we propose a hybrid method where we first generate aspects from natural language text via unsupervised clustering of word vector representations, and secondly extract aspect-level sentiment. We then proceed to summarize these aspect-level sentiment within and across reviews and provide visualizations to aid consumers.

1 Introduction

In today's e-commercialized society, the consumer not only has access to online stores through which she might purchase anything she desires, but she also has unprecedented access to a deluge of information—most notably, product reviews written by other consumers—with which she can make her decision. Unfortunately, sifting through hundreds of reviews across tens of different websites to acquire specific information about the product and its important attributes (e.g. *battery life* for electronics) is a time-consuming chore. It is difficult to automatically summarize this information, especially because of the heterogeneity of attributes across different product categories. Because of this, there has been much recent research tackling the two separate but connected components of this problem: (1) automatic discovery of aspects, and (2) aspect-specific sentiment analysis.

Within the context of a product review, the first component, *aspect discovery*, involves identifying individual "aspects" (which could be the product itself, or features/attributes of the product). This is made more difficult by the fact that relevant aspects might differ across similar products. For example, for a given electronic item such as a mobile phone, relevant aspects could be the battery life, screen size, weight, cost, and more. Not all these aspects are relevant across other electronic items: for a computer, weight might not be an issue, and screen size might be moot for headphones. Previous methods have relied on graphical models (e.g. Latent Dirichlet Allocation [1-3], or Conditional Random Fields [4]) to model latent aspects. We propose using similar unsupervised methods to generate candidate aspects via semantic word vector representations.

The second component, *aspect-specific sentiment analysis* involves subsequently identifying the sentiment—positive, neutral, or negative judgments—associated with aspects [1-6]. In this paper, we propose extending recent successful advances in deep learning [6-7] to address this problem. In particular, we seek to improve and extend the work in [6], who used deep learning architectures like the Recursive Neural Tensor Network (RNTN) to extract aspect and sentiment in a single step. Unfortunately, supervised methods (e.g. [5-7]) that require labeled aspects-sentiment pairs for training are unscalable, and a better approach would be to automatically identify relevant aspects for each product (which, e.g. [1-3] try to do via topic modeling).

2 Problem Statement

Our proposed workflow can be divided into two main parts: (unsupervised) **Aspect Discovery** and **Aspect-Specific Sentiment Extraction**. See Fig. 1 for an illustration.

2.1 Aspect Discovery

Problem: Given an **unlabeled** set of product reviews (for a single product), identify clusters $\{C_1, \ldots, C_k\}$, where the (weighted) centroid of cluster i would represent aspect i, and member words of cluster i would represent synonyms that map onto the same aspect i.

Dataset: 6 million reviews of electronic products from Amazon.com [8].

Evaluation: **(fill in)**.

2.2 Aspect-Specific Sentiment Extraction

Input: From above, we can generate aspect labels for product reviews. We will also use an existing trained sentiment analysis model to generate sentiment labels for each token / unigram.

Problem: Given a set of product reviews (for a single product), identify (aspect, aspect-related sentiment) tuples. This is a similar problem statement as [7]: we intend to expand upon their model.

Evaluation: **(fill in)**.

Output: Given the (aspect, aspect-related sentiment) tuples, we can summarize them across reviews to obtain a "meta-review" with aggregated sentiment for each aspect. We will also visualize these results to aid consumers.

3 Technical Approach and Models

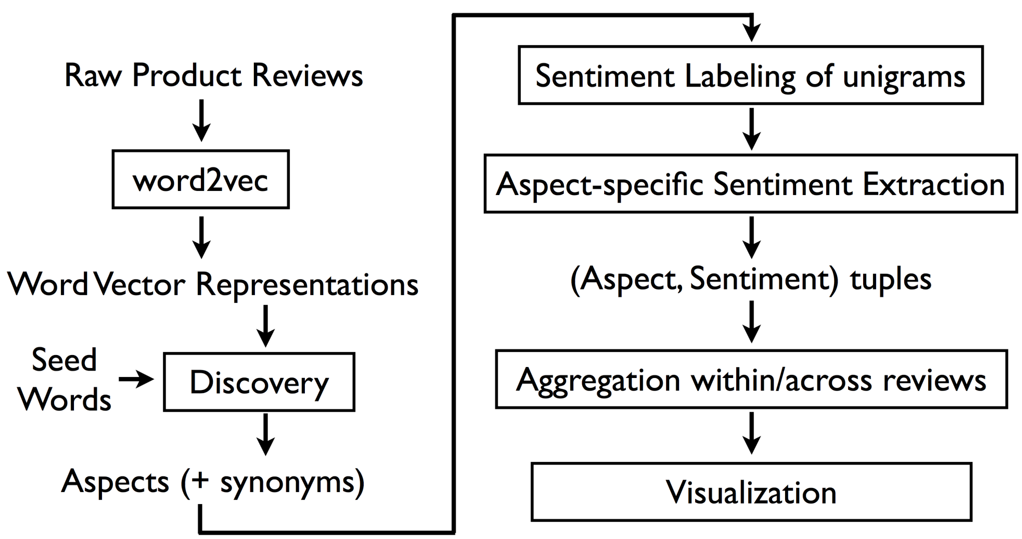

Figure 1: Proposed Workflow. Boxed items are processes, while non-boxed items are outputs. The left half of the diagram details unsupervised aspect discovery, while the right half of the diagram involves aspect-specific sentiment extraction, summarization and visualization.

3.1 Aspect Discovery

The first part of the project involves an unsupervised approach to generating aspects. First, we run word2vec to obtain a word vector representation of the data. Then, we would perform k-means clustering (or some other clustering algorithm) on the word2vec representations. This would generate clusters of similar words, such as: {build construction, build quality, durability, etc}. We can then define the weighted centroid of these clusters as **Aspects**, and the other words in the cluster as relevant synonyms. This allows us to label all the aspect-words in the dataset.

3.2 Aspect-Specific Sentiment Extraction

An additional pre-processing step would be to label all unigrams with a trained sentiment analysis model. With the aspect-labels from the previous part, we would have a dataset labeled with aspect and (unigram-) sentiment.

Next, we build off and improve on the model in [7] to use a RNTN or improved architecture to extract (aspect, aspect-specific sentiment) tuples from the data.

An application-related goal would be to have a summarization or meta-review of our aspect-related sentiment, and finally, a visualization of our output, in terms of a word cloud of important aspects (e.g., sized by word frequency and colored by sentiment), or a bar chart that shows the relevant aspects and their associated sentiment.

4 Intermediate/Preliminary Experiments & Results

State and evaluate your results up to the milestone

Acknowledgments

We would like to acknowledge funding for computing resources provided by the Deep Social Learning Lab at Stanford.

References

[1] Titov, I., & McDonald, R. T. (2008). A Joint Model of Text and Aspect Ratings for Sentiment Summarization. In *ACL* (Vol. 8, pp. 308-316).

[2] Jo, Y., & Oh, A. H. (2011). Aspect and sentiment unification model for online review analysis. In Proceedings of the fourth ACM international conference on Web search and data mining (pp. 815-824). ACM.

[3] Brody, S., & Elhadad, N. (2010). An unsupervised aspect-sentiment model for online reviews. In *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 804-812). Association for Computational Linguistics.

[4] Engonopoulos, N., Lazaridou, A., Paliouras, G., & Chandrinos, K. (2011). ELS: a word-level method for entity-level sentiment analysis. In *Proceedings of the International Conference on Web Intelligence, Mining and Semantics*

[5] Moilanen, K., & Pulman, S. (2009). Multi-entity Sentiment Scoring. In *Recent Advances in NLP* (pp. 258-263).

[6] Lakkaraju, H., Socher, R, & Manning, C. (2014). Aspect Specific Sentiment Analysis using Hierarchical Deep Learning. *NIPS Workshop on Deep Learning and Representation Learning*

[7] Socher, R., Perelygin, A., Wu, J. Y., Chuang, J., Manning, C. D., Ng, A. Y., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the conference on Empirical Methods in Natural Language Processing (EMNLP)* (Vol. 1631, p. 1642).

[8] McAuley, J., Targett, C., Shi, J., & van den Hengel, A. (2015). Image-based recommendations on styles and substitutes. *ACM Special Interest Group on Information Retrieval (SIGIR)*