
Sentiment Analysis using Natural Language Processing

Project Title

Product Feature Extraction and Sentiment Analysis in Product Reviews

Aditya Gourav (2K11/CO/007): adi.dtu@gmail.com

Aman Gupta (2K11/CO/016): aman1309@live.in

Ashish Yadav (2K11/CO/029): ashuyadav93@gmail.com

Kartik Singal (2K11/CO/052): kartiksingal1@gmail.com

Mentor: Mr. Manoj Kumar Gupta (Associate professor)



Computer Engineering, Delhi Technological University, New Delhi

Contents

- 1 Project Title
- 2 Project Abstract
- 3 Project Abstract: Introduction
- 4 Task
- 5 Evaluation
- 6 Baseline
- 7 Challenges

Project Abstract: Introduction

In this project, we plan to work on product reviews of various product classes and analyze them for finding the product features and opinion of various customers about those features. Using this analysis we aim to identify feature-wise good and bad aspects of a given product. This can be a useful practical solution to allow customers to help decide how well a product satisfies his/her needs if they are only looking for few important features in a product and don't care about other features.

We plan to use the product information available on marketing sites e.g., www.amazon.com, twitter.com to learn common features for a given product type to help reduce redundancy. Also, we aim to analyze the performance of different feature extraction techniques on different product classes. Opinion mining for specific product features would require some level of semantic understanding to separate opinions about other features mentioned in the same review. Also same word can be used to express contrasting opinions, which must be taken into account to avoid incorrect sentiment classification if only a global polarity is used for each sentiment word. Based on these challenges, we aim to achieve a robust solution for extracting features and opinion about them from the product reviews. We present the general trends and within we only focus on extracting reviews from online websites about a particular product. Additionally, we present some interesting emerging projects in the contemporary research and point out the key appliances of sentiment analysis.

Sentiment Analysis Applications

This section presents some selected and most prominent use cases for opinion mining techniques. Some have been already related in various research experiments previously mentioned while others still remain a goal to achieve in the future.

- ▶ Product benchmarking and market intelligence
- ▶ Advertisement placement
- ▶ Opinion search and retrieval
- ▶ Opinion spam detection

Task

In this project, given reviews for a product, we essentially identify its features/attributes and then extract opinion about those product features and classify them as positive or negative and grade them. We will analyze the opinion expressed for each product feature and observe the good and bad features from a customer's perspective. We will also obtain a list of commonly used features for a given product type.

Evaluation

- Evaluation of product feature extraction can be done in comparison with another system such as Hu and Liu's (2004) feature extraction system on the same dataset. Feature extraction can also be evaluated qualitatively to observe whether the most useful features of a product are obtained or not. For this we can manually create list of useful features for target products.
- The datasets have an annotated test set for sentiment analysis which can be used for evaluation of opinion mining task. We can also compare the performance with other systems (Related Papers) using the same dataset.

Baseline

We plan to use the following baselines -

- Product Feature Extraction: We can use an n-gram model to extract noun phrases/words which are usually candidate features for a product. We can also add Wordnet synset data to expand on the list of candidate features and also put an appropriate frequency threshold to discard unimportant features.
- Sentiment Analysis: We can use a list of sentiment words already marked as positive and negative and then score each sentence as positive or negative or neutral based on presence of these words in that sentence.

Challenges

- Some of the product classes do not have well-defined features like movies, books etc. For such classes, we need to identify implicit features based on what customers liked/disliked about it, which could be something like the movie plot, specific actor's performance etc.
 - Currently available datasets usually label an entire review as positive or negative opinion about a product. But even a positive-labeled review can contain negative opinion about one or more product features. This needs to be considered when mining opinion for individual features and evaluating results.
 - It would be interesting to identify the phrases within a sentence with positive or negative opinion rather than tagging the whole sentence as positive or negative.
-