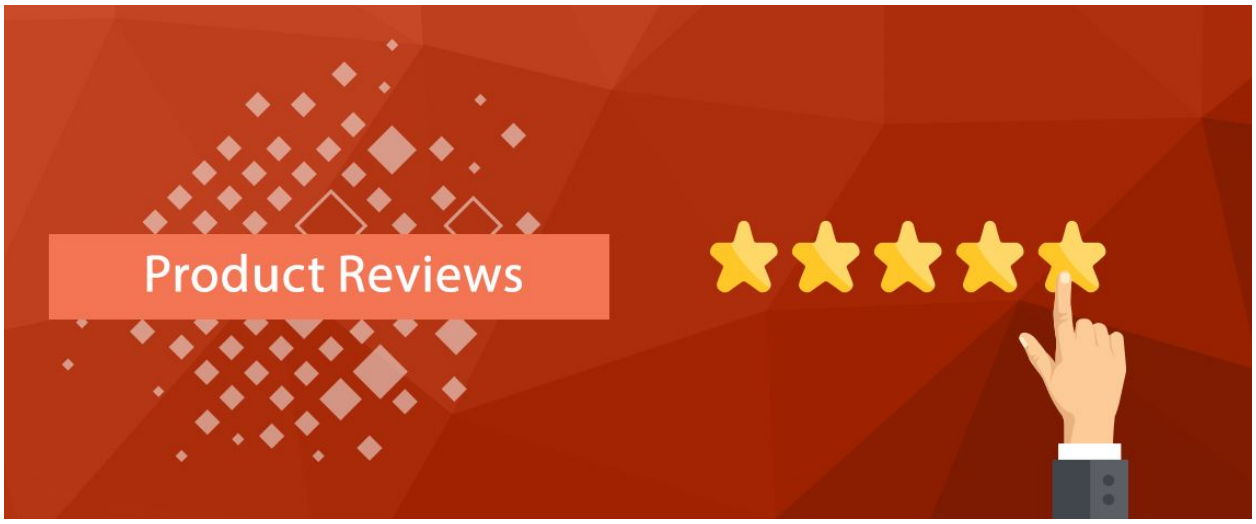

ASPECT BASED SUMMARIZATION OF REVIEWS

Team Name: Lazy and Slow

TEAM MEMBERS:

VIVEK KUMAR (2019701004)

TUSHAR CHANDRA (2019701014)

FACULTY ADVISOR:

Dr. Manish Shrivastava

IIIT -Hyderabad

Overview

Reviews posted on Web-based consumer platforms such as Amazon and Zomato form a valuable source of information prior to buying a product or hiring a service. Especially, when a multitude of reviews of a specific product are gathered, the different experiences enable the consumer to obtain a well-informed conception of the qualities and deficiencies of the respective article.

However, achieving a coherent view of a product tends to be challenging when a high volume of reviews diverge in focus and sentiment. A system that can summarize the reviews of a product would therefore be a valuable asset to these platforms. This project aims to implement and evaluate an effective approach to aspect-based sentiment analysis.

Goals

Our main goal is to create summaries that zoom in on the pros and cons of the product that are mentioned in a review.

This has wide applications enabling users at various e-commerce and service provider websites to choose better products in less time.

Specifications

FIrst of all we would like to gather sufficient dataset corresponding to products so as to be able to determine the aspects related to the product.

For example, if the dataset is of a food review app then our aspects can include food, service, price, delivery , ambience among others as the various categories in which we would try to rate the restaurant or a product.

Then our work would include tokenization , parsing and Part of Speech Tagging of the text reviews so that we are able to deal with them more effectively, and it also helps us in applying effective sentiment analysis such that we avoid ambiguity and gather the sense of the data.

We try to figure out keywords from reviews and classify them for each aspect such that it contributes either in a good or bad manner. This is proposed to be done using one of the topic modelling techniques such as Latent Dirichlet allocation method.

Later our work reduces to calculating scores and ratings for each aspect based on the sentiment of the keywords or features used.

This helps user to categorize products easily without taking the effort of going through hundreds of mixed reviews.

Milestones

I. Web scraping for data about reviews of products
II. Tokenization and Parsing of Review data
III. Aspect Identification
IV. Keyword identification using Topic modelling
V. Semantic Classification of aspect based on several reviews
VI. Development of score parameters and evaluation of aspects
VII. Report and Presentation