Project Proposal

Course: Natural Language Processing (CS6320)

Project Title: Opinion Mining on Yelp Academic Dataset

Abstract

Opinion mining is an area of Natural Language Processing where a system is developed that can identify and classify opinion or sentiment as represented in an electronic dataset. Opinion Mining is simple words is defined as identifying the subjective sentences from an input document and successively identify the polar phrase among the sentences as positive and negative.

Project Scope

- Opinion Mining includes sequence of well-known standard mechanisms as follows:
  - Subjectivity Detection
  - Polarity Detection
  - Degree of Polarity Identification
- Project includes implementation of the Subjectivity Detection and the Polarity Detection following the algorithm specified in the paper mentioned below.
- Subjectivity Detection is the identification of the subjectivity of the sentences using subjectivity words and the dependency parsing.
- Polarity Detection includes calculating the polarity of the subjective sentences identified in the previous steps using subjective phrases and dependency relationships among them.

Dataset for the Project:

Yelp Dataset Challenge

Yelp is providing all the data and reviews of the 250 closest businesses for 30 universities for students and academics to explore and research.

The dataset is a single compressed file, composed of one ‘JSON-object’ per line. Every object contains a 'type' field, which tells you whether it is a business, a user, or a review.

Link: **https://www.yelp.com/academic_dataset**

Yelp dataset has the following Objects:

- **Business Objects:**
  - Business objects contain basic information about local businesses.
- **Review Objects:**
  - Review objects contain the review text, the star rating, and information on votes Yelp users have cast on the review.
- **User Objects:**
  - User objects contain aggregate information about a single user across all of Yelp

Reference Paper

A Review on Natural Language Processing in Opinion Mining by Debnath Bhattacharyya, Susmita Biswas and Tai-hoon Kim-International Journal of International Journal of Smart Home, Vol.4, No.2, April, 2010 Vol.4, No.2, April, 2010

Link: **http://www.sersc.org/journals/IJSH/vol4_no2_2010/3.pdf**