

College Reviewing System

Submitted in partial fulfillment of the requirements
of the degree of

Bachelor of Engineering

in

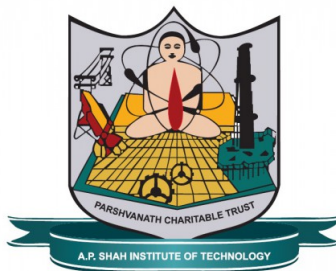
Computer Engineering

by

Kalpesh Gohil (16202004)
Yash M Jain (16202006)
Sanat Yadav (15102038)
Karan Vishwakarma (15102062)

Guide

Prof. Pravin Adivarekar



Department of Computer Engineering

A.P. Shah Institute of Technology
G.B.Road,Kasarvadavli, Thane(W), Mumbai-400615
UNIVERSITY OF MUMBAI
2017-2018

CERTIFICATE

This is to certify that the project Synopsis entitled “*College Reviewing System using Sentimental Analysis*” is a bonafide work of “*Kalpesh Gohil (16202004), Yash M Jain (16202006), Sanat Yadav(15102038), Karan Vishwakarma(1512062)*” submitted to the University of Mumbai in partial fulfillment of the requirement for the award of the degree of *Bachelor of Engineering in Computer Engineering*.

(Name and Sign)

External Examiner

Prof.Sachin H Malave
Head Of Department

Dr. Uttam. D. Kolekar
Principal

Abstract

The use of reviews has created many opportunities for people to publicly voice their opinions. Reviews in the internet could be in millions for services which make it difficult to track and understand customer opinions. Sentiment analysis is an emerging area of research to extract the subjective information to track and understand customer opinions. The reviews provide accessible and plentiful data for relatively easy analysis for a range of applications. This system seeks to apply and extend the current work in the field sentiment analysis on college reviews data retrieved from another websites using web mining technique. Naive Bayes and decision list classifiers are used to tag a given review as positive or negative. The features, such as bag-of- words and bigrams, are compared to one another in their effectiveness in correctly tagging reviews. Recent studies analyzed this reviews and found that it includes information useful for college, such as user requirements, ideas for improvements, user sentiments about specific features, and descriptions of experiences with these features.

Introduction

Reviews means the text which is given by user related with our system service. This review is useful for get detail about our system means what is better in our system and which changes is require for make our system better. Using these reviews we can take opinion of the user of our system. A user review refers to a review written by a user for a product or a service based on her experience as a user of the reviewed service or product. Most of college website gets reviews from users so the college system see view services is better and which services is require changes. Using this reviews college can make changes in services. Popular sources for consumer reviews are e-commerce sites like Amazon or Zappos, and social media sites like Trip Advisor and Yelp. E-commerce sites often have consumer reviews for products and sellers separately. Usually, consumer reviews are in the form of several lines of texts accompanied by a numerical rating. This text is meant to aid in shopping decision of a prospective buyer. A consumer review of a product usually comments on how well the product measures up to expectations based on the specifications provided by the manufacturer or seller. It talks about performance, reliability, quality defects, if any, and value for money. Sentiment analysis can be defined as a process that automates mining of attitudes, opinions, views and emotions from text, speech, tweets and database sources through Natural Language Processing (NLP). Sentiment analysis involves classifying opinions in text into categories like "positive"; or "negative"; or "neutral";. It's also referred as subjectivity analysis, opinion mining, and appraisal extraction. The words opinion, sentiment, view and belief are used interchangeably but there are differences between them.

Objectives

- 1)The system is designed to gain reviews from the students and these reviews are divided into positive, negative and neutral.
- 2)The main objective of this system is to seek reviews for the respective college from the trusted alumni students.

Literature Review

The bag-of- words model is one of the most widely used feature model for almost all text classification tasks due to its simplicity coupled with good performance. The model represents the text to be classified as a bag or collection of individual words with no link or dependence of one word with the other, i.e. it completely disregards grammar and order of words within the text. This model is also very popular in sentiment analysis and has been used by various researchers. The simplest way to incorporate this model in our classifier is by using unigrams as features. Generally speaking n-grams is a contiguous sequence of n words in our text, which is completely independent of any other words or grams in the text. So unigrams is just a collection of individual words in the text to be classified, and we assume that the probability of occurrence of one word will not be affected by the presence or absence of any other word in the text. This is a very simplifying assumption but it has been shown to provide rather good performance.. One simple way to use unigrams as features is to assign them with a certain prior polarity, and take the average of the overall polarity of the text, where the overall polarity of the text could simply be calculated by summing the prior polarities of individual unigrams.

Problem Definition

All words are weighted same for example good and best belongs to same Category. The sequence in which words come in test data is neglected. Other issues- Efficiency provided from this implementation is only 50-60

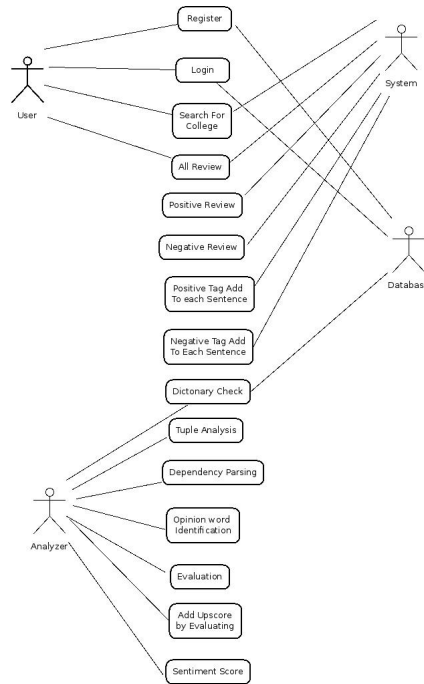


Figure 1: UseCase Diagram

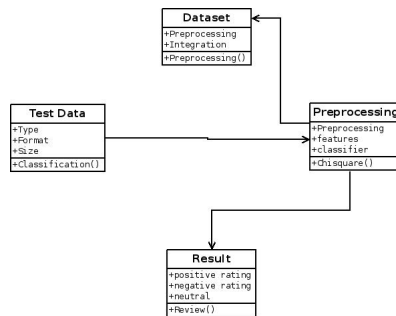


Figure 2: Class Diagram

Proposed System Architecture/Working

To overcome the existing drawbacks we have proposed a system that can extract the college reviews information from website and analyze these reviews using sentiment analysis. Sentiment analysis is a very relevant technique nowadays for analysis. Sentiment analysis or web mining is the process of automatically extracting knowledge from sentiments or reviews of others about some topic or problem. We can identify reviews in a large unstructured/structured data and analyze polarity of reviews. In this proposed system we can use Naive Bayes or chi-square algorithm for analyze college reviews and to tag a given review as positive, negative or neutral. The results can be used for various purposes such as guiding decisions to improve the college system and also the students have a genuine reference of the college before taking the admission.

Advantages of Proposed system

1. Easily gets reviews from various college interacts websites.
2. Sentiment analysis gives a proper result of positive or negative reviews.
3. Using this analysis we can easily get what is our system plus point and which sector require changes.
4. Gives better services for user.

Summary

The work presented in this report is related to College Reviewing System.

References

- [1] Abror Abduvaliyev, Al-Sakib Khan Pathan, Jianying Zhou, Rodrigo Roman and Wai-Choong Wong, “On the vital Areas of Intrusion Detection Systems in Wireless Sensor networks”, IEEE Communications Surveys & Tutorials, Accepted For Publications, 2013-in press.
- [2] H.H. Soliman, et al, “A comparative performance evaluation of intrusion detection techniques for hierarchical wireless sensor networks”, Egyptian Informatics Journal (2012) 13, 225238.
- [3] Giannetsos Athanasios, “Intrusion Detection in Wireless Sensor Networks”, Master THESIS, Carnegie Mellon University, April 8, 2008.
- [4] K.Fall and K.Varadhan, “The NS Manual”, http://www.isi.edu/nsnam/ns/doc/ns_doc.pdf, 1 Feb 2014.
- [5] Jae Chung and Mark Claypool, “NS by Example-Tutorial”, <http://nile.wpi.edu/NS/overview.html>, 1 Feb 2014.
- [6] Network Simulator blog, <http://Mohittahilani.blogspot.com>, 1 Feb 2014.
- [7] AWK Script for NS2, <http://mohit.ueuo.com/AWK-Scripts.html>, 1 Feb 2014.