STUDENT THOUGHT ANALYSIS ON COLLEGE VIA SOCIAL NETWORKING

1saran Ram Kr, 2ms. Vimali J.S, 3dr.G.Rosline Nesakumari, 4dr.S.Maruthuperumal

1U.G Scholar Department of IT, Sathyabama Institute of Science and Technology, Chennai
2Assistant Professor, Sathyabama Institute of Science and Technology
3Professor CSE Shadan women’s college of Engineering and Technology, Hyderabad
4Professor & HOD CSE, NBKR Institute of Science and Technology, AP.

ABSTRACT

In the current generation sharing thoughts and data’s is majorly dominated by social media. Sentimental analysis here is deducing the thoughts and emotions of a person about a source. The process of analyzing tweets with their polarity is sentimental analysis. As twitter is the platform here, we use tweets left by users as the input data to further extract and analyze the polarity of the tweets. These are further classified to improve the particular field or sector.

Keywords: Sentimental analysis, Python

I. INTRODUCTION

In recent years social networking has played a major role in analyzing the sentiments of people. In this process we will be using twitter as the social network platform to extract data about college in this paper. In this case the particular source will be able to analyze the results provided in the survey analysis. These data’s can be vitally used in further improvements and betterments.

II. LITERATURE SURVEY

This survey is about how social media (Twitter) can be healthy utilized in order to obtain social comments or opinion in the form of tweets. These tweets are of high value as they carry the direct sentiment of the user leaving the tweet about the particular objective or topic. These tweets which are gathered can be used in a very optimistic way to enhance the betterment of the concerned objective. These tweets are categorized into different sentiments which can be (Happy, Sad or Neutral). By using several different objectives a huge amount of time and effort can be saved. We have for example an analysis of facebook and twitter usage during the earthquake relief efforts. The devastating Haitian earthquake hit the small Hispoinolaisland, result of which thousands were deceased and damage of property was in billions of dollars. In this point of time the earthquake had started a firestorm social media utilization by various organizations. Framing theory was applied in order to analyze the Facebook posts and tweets sent by media and non-provitale organizations. Social media usage is used to determine the difference between media and noprovable organizations by the social media usage terms. By ] S. Muralidharan, L. Rasmussen, D. Patterson, and J.-H. Shin [2] Also, we have D. L. Cogburnand F. K. Espinoza-Vasquez who were examining the impact of web 2.0 and social media on political participation and civic engagement in the 2008 obama campaign [3] M. Sun, Z. Liu, M. Zhang, and Y. Liu, Chinese Computational Linguistics and Natural Lang. Process. Based on Big data Annotated naturally.[4] Sentiment analysis using support vector machines with diverse information sources.[5] H. Kanayama and T. Nasukawa had created which was a Fully automatic lexicon expansion for domain-oriented sentiment analysis.[6] We also have this applied using a parsimonious rule-based model for sentiment analysis of social media.
III. METHODOLOGY

First in this process of sentimental analysis we have unique methods which are used in order to deduce the input of tweets so that we have the best result. The first method used here is Lexicon-Based Methods which is categorized into 3 sup-approaching methods:

1. Two Approaches where the basic assumption of lexicon-based methods is that its dominant sentiment words are used to determine its sentiments.
2. Basic Lexicon-Based Methods is where automatic classification algorithms are used, for computation purpose special representation is to be used.
3. Linguistic Rules

In this process we have used two different representations which are listed below:

1. Bag-of-words Ngrams
2. Linguistic features.

In evaluation we have used the standard matrices in text categorization to evaluate various classifiers. Suppose we have a set of classification results, in each class of n documents, cij (0 ≤ i ≤ n−1, 0 ≤ i ≤ n−1) denotes the number of instances where a document in the ith class is categorized as belonging to the jth class.

Data collection:

The first process of this sentimental analysis is the collection of data’s. This is the most important part as the data collected here are to be used to process for the sentimental analysis to be made. Collection of data is done with the use of keywords and similarity in the topics.

Data Processing:

In data processing we have the data’s as tweets which are initially preprocessed. The process of preprocessing comprises of several layers. For example the #tags are initially removed as they are no longer required. Further the URL’s are checked and removed if to be. Finally spell checks are made which are very essential for further phases.
Data Filtering:
The tweets which are obtain may contain raw information which may be useful or may not be. Thus, by removing stop words, numbers and punctuations tweets can be filtered.
Feature Extraction:
In the process of data extraction, the data which is obtained from the data filtering or processing part is extracted. These extracted data’s are assorted to be sent on to the next phase where the sentiment of these data are analyzed. The extraction process is carried out with specific methodologies and algorithms.

Sentiment Analysis
Sentimental Analysis is to get the extracted data from the extraction phase and categorize these data into their assorted polarity. This process is done using several methodologies and algorithms. The sentimental analysis provides a result of Positive, Negative and Neutral results to the data.

IV. RESULTS AND DISCUSSIONS
The output is displayed once the review is left on the twitter page. The polarity of the tweet is analyzed and displayed adjacent to the review. Once a user of twitter leaves a review on a particular department, the review along with the polarity of the review will be analyzed and displayed on the Admin user account. These data can be gathered to statistically categorize the fields to be concentrated or focused on.

Fig 1: Input of review from a test account with the location

Fig 2: The review is processed and analyzed with the polarity we were able to deduce and provide the polarity of the review left on a particular department which can be collected as a complete data to analyze the particular departments for further improving them.

V. CONCLUSION
In this paper we have gone through methods for sentimental analysis over twitter data, machine learning based methods, rule-based methods and lexicon-based methods. We faced two problems. [1] Evaluating sentiment classification using precision, accuracy, recall and F1 value legitimately. [2] As the improvement is not significant in sentiment analysis it is baffling to spend effort and time to apply rich linguistic features. In future we plan to apply this concept in several important streams where the sentimental analysis of tweets can be used to determine the actual percentage of each polarity for further improvements and advancements.

REFERENCES


