­National Institute of Technology Karnataka, Surathkal

Department of Computer Science and Engineering

Project proposal submission (Internet Technology and Applications)

Class: VI Sem. BTech CSE (Jan.-Apr.2018)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Roll No. in the class | Inst. Reg. No. | Name | Mail Id. | Phone No. | CGPA and Role  (Leader/Member) |
| 15CO134 | 156313 | Gunjan Paliwal | [gunjanpaliwal7@gmail.com](mailto:gunjanpaliwal17@gmail.com) | 8208983137 | 7.92 - Leader |
| 15CO137 | 156434 | Rahul Kumar | rahulkumar4871@gmail.com | 7204852903 | 6.65 - Member |
| 15CO219 | 156462 | Jay Prakash | jayprakash1319@gmail.com | 7204335016 | 6.24 - Member |

**Title : Web-Application for Sentimental Analysis of Tweets**

**Abstract :** With the enormous increase in web technologies, number of people expressing their views and opinions via web are increasing. This information is very useful for businesses, governments and individuals. With over 500+ million Tweets (short text messages) per day, Twitter is becoming a major source of information. Many major decisions now requires opinion polling of users on twitter. We propose multiple methods to analyze singular and also bunch of tweets on sentimental basis and classify it as positive or negative. Our project will be a Web-App which will allow you to analyze any user input tweet along with the accuracy results of classifying a group of tweets. Line and bar graphs will be used to represent the statistical analysis of the results.

**Implementation :** Using this social media we plan to build two models for classifying ”tweets”. In the first one, we’ll classify the tweets into positive, negative and neutral classes. We will build models for two classification tasks: a 3-way classification of already demarcated phases in a tweet into positive, negative and neutral classes and another 3-way classification of entire message into positive, negative and neutral classes and experiment with the baseline model and feature based model along with an incremental analysis of the features. We also plan to experiment with a combination of models: combining baseline and feature based model. In the second model, we classify the tweets into subjective sentiments of sad, happy, angry and surprised. We’ll experiment with various hyperparameters and features in a GMM. The GMM gives us a relative score for the tweet instead of an absolute classification which is in line with the subjectivity of any tweet. We are using SVM to classify our tweets.

**Work Distribution :**

Gunjan Paliwal : Working on dataset and preprocessing, Feature Extraction, Dataset model training

Rahul Kumar : GUI(Front-End web page), Statistical Analysis of results and graphical representation

Jay Prakash : Back-End, Interprocess communication between Front-End and Back-End

**Software Requirements :** Operating System : Linux/ OS, Browser Support : Mozilla/Chrome

**Resources and Tools :** Emoticon and Acronym Dictionary, Sklearn, NLTK, Python spell checker