EMOTION DETECTION USING TWITTER ANALYSIS

MINOR PROJECT II

Submitted by:

Ayush Govil (9914103159)
Saurabh Sharma (9914103166)
Karthik Venkataraman (9914103168)
Rohan Gupta (9914103181)

Under the supervision of:
Mr. Ravinder Ahuja

Department of CSE/IT
Jaypee Institute of Information Technology University, Noida

MARCH 2017

ACKNOWLEDGEMENT

I would like to place on record my deep sense of gratitude to **Mr Ravinder Ahuja**, Assistant Professor, Jaypee Institute of Information Technology, India for his generous guidance, help and useful suggestions.

I also wish to extend my thanks to my seniors and other classmates for their insightful comments and constructive suggestions to improve the quality of this project work.

Signature(s) of Students

Ayush Govil (9914103159)
Saurabh Shamra (9914103166)
Karthik Venkataraman (9914103168)
Rohan Gupta (9914103181)

ABSTRACT

Social media and microblog tools are increasingly used by individuals to express their feelings and opinions in the form of short text messages. Detecting emotions in text has a wide range of applications including identifying anxiety or depression of individuals and measuring well-being or public mood of a community. In this project, we propose a new approach for automatically classifying text messages of individuals to infer their emotional states. To model emotional states, we utilize the well-established Circumplex model that characterizes affective experience along two dimensions: valence and arousal. We select Twitter messages as input data set, as they provide a very large, diverse and freely available ensemble of emotions. Using hash-tags as labels, our methodology trains supervised classifiers to detect multiple classes of emotion on potentially huge data sets with no manual effort. We investigate the utility of several features for emotion detection, including unigrams, emoticons, negations and punctuations. To tackle the problem of sparse and high dimensional feature vectors of messages, we utilize a lexicon of emotions. We have compared the accuracy of several machine learning algorithms, including SVM, KNN, Decision Tree, and Naive Bayes for classifying Twitter messages. Our technique has an accuracy of over 90%, while demonstrating robustness across learning algorithms.

Gantt chart

SYNOPSIS						
LITERATURE SURVEY						
OPTIMIZATION CODE						
DEBUGGING						
VALIDATION AND DOCUMENTATION						
DATE	**MARCH 2,2017**	**MARCH 4,2017**	**MARCH 7,2016**	**MARCH 13,2017**	**MARCH 15,2017**	**MARCH 19,2017**

TABLE OF CONTENTS

LIST OF FIGURES