Mood Visualization

Vishal Rana, Rohit Gupta, Abhishek Raj Sahu, Viijayraj Ghadge

Indian Institute of Technology, Kanpur, Uttar Pradesh, 208016

Abstract- With the boom in information technology, there is so much data available on the internet- 'The Big Data'. The abstraction, analysis and classification of this data is a fast growing field of research. This paper describes an effort at collecting such data from a social networking site and then representing the same visually.

Keywords- Tweepy, Naive-Bayes Algorithm, RaspberryPi 2, NLTK Library, Colours.

I. Introduction

Every sentence expresses an emotion and is a reflection of the mood of the person speaking. Also various colours are associated with different feelings, like red reminds us of anger. So overall emotion of the data collected can be associated with a particular colour and the gradual variation of it over time can be a reflection of the changing emotions of the source of data.

II. Motivation

There are times when there is an outburst of anger across the world in response to some event, while at others people come together to celebrate. The social networking sites are abuzz with activity at such times and the data available there from a myriad of sources reflects the mood of the world.

On a smaller scale, even minor events in a collage campus, like a lecture or seminar can affect the overall mood of the campus. Representing such changes in mood by analyzing the data and presenting it visually appealing way can make this concept apparent to even a novice.

III. Background Theory

Tweepy is twitter API that has been used to stream live tweets having certain keywords and from certain geographical regions. This raw data is pre-processed and classified using machine learning. Naive Bayes Algorithm works on the concept of Bayes Theorm to classify tweets, using a big trraining set consisting of already classified tweets. then statistically, these classified tweets are mapped to a particular emotion and hence to a particular colour.

IV. Implementation

Software:

**NLTK** is a leading platform for building Python programs to work with human language data.

It provides an easy implementation of the Naive Bayes Algorithm.

The overall implementation is divide into the following sections:

1. Training the Classifier: Initially a set of classified tweets is provided to the classifier for training. Bigger the data-set, better is the accuracy of the classifier. The classifier converts the data into pickle file that can be directly loaded and saves it from going through the data repeatedly.

2. Collecting Data: Given the keywords and the data size, the program collects tweets from Twitter in real-time using the Twitter API Tweepy.

3. Pre-Processing the tweets: Then useless parts of the tweets are removed, including urls, username, extra whitespace, letters repeated more than twice(e.g. goood to good), etc.

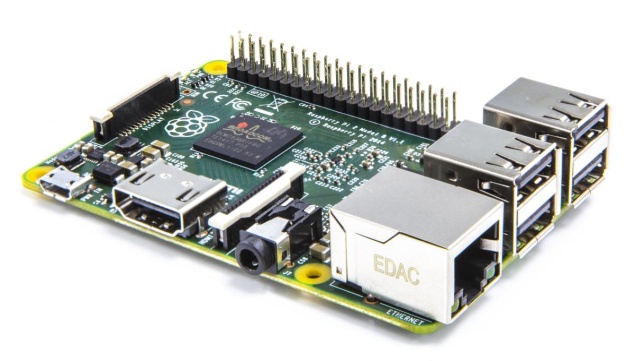
4. Classifying the tweets: Now NLTK implementation of the Naive Bayes Algorithm starts to classify and rate the tweets. The ratings of successive tweets are cumulated to give a final rating to the data collected.

5. Mapping the data: The data is then mapped to various colours representing different emotions according to the rating calculated. This is then sent over serial communication to the hardware to dispatch the requisite amount of sand.

Hardware:

Equipments and devices: Plastic funnels, aluminium rods, colored sand, acrylic sheet, RaspberryPi, ArduinoUno, Servo motors.





The model of this project comprised of four Acrylic sheets supported by four aluminum rods. The rods are 2.5cm\*2.5 cm\*58 cm. The dimension of rectangular sheet is 29cm\*29cm\*0.4cm. The containers of the sand containing sand of five different color, resides on the upper sheet at @@ from center of the sheet symmetrically. Underneath of the each container there is hole of diameter 2cm which is used to drop the sand. That opening is opened and closed with the help of a mover which is connected to the arm of the Servo motor which in turn also embedded in the upper sheet. The servo motor is controlled with help of Arduino. Below that upper sheet there is a middle disc which has holes exactly below of the ones that of top sheet. We also placed small funnels there to make sure uniform flow of the sand. That hole is opened in the glass tube of shape S. The other end of all such five tubes is opened at one point in another glass funnel supported by acrylic sheet. Then at the center of the bottom-most sheet, exactly below the opening of glass funnel, we have placed a transparent glass cylinder. The sand dropped from system described above is collected in that cylinder.



Autodesk Model of the Project.

V. Acknowledgement

The authors would like to thank Prof. Soumya Ranjan Sahoo, IIT Kanpur. Prakhar Jawre, Coordinator Electronics club, IIT Kanpur. Ayush Shakya, Co- ordinator Electronics club, IIT Kanpur.

VI. References

1. Mining Twitter Data with Python.

http://marcobonzanini.com/2015/03/02/mining-twitter-data-with-python-part-1/

2.Text Analysis

http://nealcaren.web.unc.edu/an-introduction-to-text-analysis-with-python-part-1/

3.Choosing a classifier

http://blog.echen.me/2011/04/27/choosing-a-machine-learning-classifier/

4. NLTK Library

http://www.nltk.org/

5. Pickle- Object Serialization

https://docs.python.org/2/library/pickle.html

