**Ontology based Sentiment Analysis of Twitter Posts**

**OVERVIEW**

Every day large amounts of twitter posts (or) tweets are getting generated in the web. So we decided to perform an ontology based sentiment analysis of every tweet getting generated with a specific hash tag. We will classify a twitter message either positive or negative based on the range value we have got. We will determine the polarity of a textual corpus (document, sentence, paragraph etc.) and assign a value from range of [0, 9] based on the strength of the sentiment using ontology techniques. If the tweet receives a value between 0 and 5, it is a negative tweet where as if it receives value in between 5 and 10, it is a positive tweet.

**DATA**

For this task, we have two data sources. One is self made data in a text file, where we can provide large number of tweets with a hash tag inside a notepad file and give it as an input. The other data source is Twitter database itself, where we are going to mine the live tweets. For mining this live database we registered in twitter developer API website, where we received a unique key in which we have to use this key in our java code to extract the data.

**Programming APIs, Tools and Libraries used**

We have used Java Programming language to write the logic required to mine textual and live twitter data. We have used the tool ‘Gephi’ to see the results or output in cluster based format. Also we made use of different libraries like Indigo and Twitter 4j to mine the twitter messages. We used NetBeans IDE to code the logic and test the sentiment analysis.

**RESEARCH QUESTIONS**

As part of the testing, we wanted to see the sentiment analysis of all the tweets which were associated with a hash tag ‘apple’. Below are few more questions which we were anticipating for.

**a.** How many positive and negative tweets are getting generated daily in the Twitter?

b. What was the relationship between sentiment and features like timestamps and hash tags?

c. How well the Gephi clustering plot is demonstrating sentiment analysis of tweets.

d. How positive were the thoughts of people on Apple products?

**CODE AND APPLICATION**

The code is in Github and instructions are in the README.md file.

<https://github.com/manjukr57/Section1-Project-Group2>

**PROJECT MANAGEMENT**

Our project team consists of three members, Manjunath Kuruvadi, Chinna Babu Sadam and Ananjay Mishra. Basically we have divided our project work equally among ourselves. Manjunath worked on analysis and extraction of data, also worked on coding part. Chinna Babu worked on setting up the tools and Mishra worked on documentation and testing activities. We are determined to finish our project by collective team work.

**PROJECT TEAM, DELIVERABLES AND CHECKPOINTS**

**TEAM:**

| **Team member** | **Roles and skills** | **Contributions** |
| --- | --- | --- |
| Manjunath Kuruvadi | Team Lead and Code Developer | Formulated the idea of using Twitter sentiment analysis and started analyzing the twitter data. Implemented the sentiment analysis using Java. |
| Chinna Babu Sadam | Installation and Environment setup | Finished installing and setting up the environment required to work in the project. Also worked on transformation of data activity. |
| Ananjay Mishra | Analyzing and Testing | Worked on analysis of Twitter messages and involved in testing activities required to get well formed output |

DELIVERABLES AND CHECKPOINTS

| **Checkpoint date** | **Expected Deliverable** | **Responsible team member(s)** | **Checkpoint results** |
| --- | --- | --- | --- |
| 2/22/2016 | Finished analyzing data | Manjunath | Found some bad data and asked my team mate to remove that data during cleaning process |
| 2/28/2016 | Cleaning of data | Chinna Babu Sadam, Ananjay | Cleaning process finished and some of the bad data has been removed. |
| 3/8/2016 | Installation and setup | Chinna Babu, Manjunath | Worked on setting up the environment and chosen Java, Twitter 4j, Indico and Gephi APIs to implement this project.  Also started coding. |
| 4/11/2016 | Testing | Ananjay, Manjunath | Finished updating project proposal document. Also finished installing and setting up the tools. Finished analysis of data. Initial round of testing the code has also been finished. |
| 4/18/2016 | Code Development and PPT preparation | Manjunath | Finished the coding required to mine the data. Able to see the output correctly in NetBeans IDE. Also tested the cluster based output in Gephi tool. |