**ATOS Group 4(Social Media):**

**Andreas Mihaloianis, Momshad Alvee Dinuri, Chaitanya**

**Agrawal**

**27/03/2015**

**Overview:**

The past two weeks have involved improving the Bag of Words algorithm to

have a greater accuracy of sentiment and also the usability of the whole

system. Improving the usability involved loading more tweets than what was

currently shown, changing color intensity according to the score of the tweet’s

sentiment and adding a star rating to each parameter on the settings page. In

addition to all these improvements, the team also cleaned the code-base,

removed hard-coded entries and thus also implemented the resource-allocation

mechanism in an indirect manner. Each user is able to choose the tweets they

want and this setting allows tweets of to each user’s choice to be displayed.

Each user’s choice has a real time star rating to help the user observe that

choice’s real time performance in an abstract way. Additionally, on the clicking

that star rating the user is taken to 100 tweets relevant to that choice and

responsible for the star rating of that choice. Finally, the team has taken the

time to implement Stanford’s NLP library (http://nlp.stanford.edu/) to perform the

sentiment analysis as an alternative solution. The team has not compared

accuracy of the two systems. However, the execution time of the Stanford’s

NLP library is significantly more than the Bag of Words algorithm.

**Summary of meetings held:**

A client meeting was fixed in the week starting March 16,2015. However, it was

postponed to March 23,2015 after a few email communications between the

team and the client changed requirements specifications for this sprint. The

action points of loading more tweets, improving the Bag of Words algorithm by

adopting a percentage model were successfully presented to the client in the

meeting held on March 23,2015.The client was happy with the work and

proposed another action point of varying color intensity relative to each tweet’s

sentiment score to be showcased in the meeting scheduled for April 1,2015.

**List of tasks completed and estimations:**

The action points of improving Bag of Words algorithm by adding percentage

model, making it possible to load more tweets on the click of a button and

varying color intensity relative to each tweet’s sentiment score were

successfully completed.

Additionally, the Stanford NLP Library was implemented as an alternative

choice to Bag of Words algorithm. A new settings page feature, star rating, was

also added. This displays a real time star rating of each parameter on the

settings page and takes the user to 100 tweets relevant to that choice and

responsible of that choice’s rating on the click of a button.

**Plan for next two weeks:**

The team plans to do a full comparison between Stanford’s NLP Library and

Bag of Words algorithm. The trade-offs with each system will be presented to

the client to make a decision on the choice of use of the algorithm where the

trade-offs would be speed and accuracy. Currently the Bag of Words algorithm

executes under 1 second whereas Stanford’s NLP Library takes about 10-20

seconds. Any action points that might be raised by the client in the upcoming

meeting on April 1,2015 would also need completion.

**Individual Paragraphs:**

**Andreas:**

I was in close contact with both the team and the client in this sprint to sort out

confusions and improvements in implementing the action points for this sprint.

As the project duration is approaching its end, there is not much left to do,

provided that we have finished the most important task of building the Bag of

Words algorithm in the previous sprint. Considering that the actions points were

minor changes to the code-base, I took a decision to allow full liberty to the

team members to play around with Momshad implementing the Stanford NLP

Library and Chaitanya adding a new feature to the settings page. I believe that

the client is happy with our efforts and the project is near its successful end.

**Alvee:**

I mainly focused on improving the Bag of Words algorithm. We are now using a

percentage based scoring system and this appears to be far more accurate

than our previous iteration. Making changes to algorithm was relatively easy

since every part of the code is modularized and hence I could simply replace

few modules with the new code. I removed the hard-coded usernames in

settings page and thus implemented the resource allocation mechanism that we

promised to implement. Also, I helped write the code that allows the user to

reload 100 new tweets. Finally, I packaged the Stanford NLP library into a jar

file and made a new project hosted on an Azure’s Ubuntu virtual machine to use

Stanford’s NLP Library instead of our own. This effectively means that we have

two applications currently and we can use this to compare the two applications

and the client can chose the best one that fits the project’s requirements.

**Chaitanya:**

I was in close contact with Momshad for the loading more tweets feature

implementation where a lot of decisions were cancelled and introduced. We

finally settled on redirecting the user to a brand new reload.php page on the

click of a button rather than loading more tweets dynamically on the same page.

This choice is more efficient as the user can go back to check the previous set

of tweets. I also implemented the action point of varying color intensity of each

tweet according to its sentiment score. A star-rating feature for each parameter

on the settings page was introduced that provides real time performance of that

parameter on Twitter. On clicking that parameter's star rating, the user is taken

to tweets of that choice that are responsible for that particular star rating.