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

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

**13.02.2015**

**Overview:**

The team considers this sprint a critical point in the project history. The meaningful and productive work being put in is shaping the final project that is to be delivered. The project has been uploaded on Microsoft Azure Cloud Services:<http://atosrail.azurewebsites.net/cis-ucl-social-media/> after being linked with a MySQL database which takes care of login and personal preferences of users. The team has also begun the transition from use of an API to the use of self-developed algorithms. This will result in much faster and efficient performance along with scope for future improvement. The majority of the tasks ahead will be concerned with the algorithm and user-interface of the application. Having achieved a core template for the application, the team will continue to make further additions to improve the ease, feel and look of using the system.

**Summary of meetings held:**

There was a mid-sprint client meeting on February 5,2015 where previously allocated work was tracked. Chaitanya’s work of uploading the project onto Azure Cloud Services was delayed because of Subscription billing reasons from Microsoft’s end. Although most of the work was not in a deliverable state, the team proposed potential solutions and research that could lead to completion of tasks by the end of this sprint. The client was very comfortable with extending the deadlines to end of sprint after considering the team’s efforts so far. The meeting on February 12,2014 was postponed to February 17,2014 by the client.

**List of tasks completed and estimations:**

The team has managed to complete all the tasks assigned for this sprint. These include uploading the whole project onto Azure Could Services after integrating it with a MySQL database for login and user-preferences purposes, rendering emoticons and improving date display format on the tweets display page, implementing front-end of the settings page and most importantly improving performance of the whole system. The team made a very important decision to switch to a self-developed algorithm from the DatumBox API and has come up with a basic functional algorithm. In the upcoming sprints, the team plans to build upon the already ready basic sentiment analysis algorithm.

**Plan for next two weeks:**

We plan on training the model of sentiment analysis using the Naive Bayes Classifier by using the dataset, which we shall obtain. This should radically improve the accuracy of our sentiment analysis algorithm. Right now access to the pages in our system is hardcoded with the username; we will replace it so that any number of users can access the pages given that they have the correct username and password. We will add new features to the application such as allowing the user to click on the picture of the person who posted the tweet and opening a new tab pointing to that persons home page. The control panel will be updated to provide better control for the user and user-interface of the control panel will be improved as well. Due to shortage of time, emoticons rendering could not be comprehensively tested so it will remain an area of improvement for the upcoming sprint. The most important addition to the project will be our own algorithm to do the sentiment analysis. Since we wont be making calls to a REST API, the performance of the system will be much faster than it currently is. Improving the performance in terms of time and accuracy are the two most important areas we plan to focus on.

**Individual Paragraphs:**

**Andreas:**

During the last two weeks, the team has divided tasks with an aim to fulfill the client’s requirements. I have split these tasks equally within the team and tried to understand Chaitanya’s wish to have a better contribution to the coding aspect of this project by allocating both backend and frontend tasks. Also, as we haven’t managed to do everything in one-week time (when our first meeting with the client took place), I had to discuss with David and rearrange a new meeting the following week so that we could finish all our tasks. In terms of coding, I have updated the displaying of the tweets and the date on the main page, taking into consideration all the feedback we have received from our client. Both the team and the client are happy with the progress so far and I think the project is on the right track at this moment.

**Alvee:**

I have worked on creating an algorithm that is capable doing sentiment analysis. So far I have been able to come up with two algorithms, “Word rank” and “Naive Bayes Classifier”. Comparing the accuracy of the two classifiers the Naive Bayes had a better overall accuracy but it is still not complete as the vast training data required to train a Naive Bayes classifier has not been acquired yet. There was a website which hosted training data for sentiment analysis but for some reason the site has been taken down. I got in touch with the people who commented on the post and requested to send me a copy. The word rank algorithm works by looking at each word in the given string and searching for it. If it is found in the good array (which is an array that holds all positive sentiment) then we give +1 to the overall sentiment and vice versa for negative. If the value is within a given range (currently -1 to +1) then the sentiment is classified as neutral. The Naive Bayes works by fitting a model that is extracted automatically from tens of thousands of pre classified example and hence called he training set. The implementation is almost complete; we are just waiting for the training data.

**Chaitanya:**

The client wanted to test our progress by directly visiting the project website and asked for the project to be put online. I uploaded the project onto Microsoft Azure Cloud Services after creating a MySQL database (for login purpose). Momshad helped me during this task by clearing the confusion between Azure SQL and MySQL approaches to upload projects onto Azure Cloud Services. The team’s Github repository is linked to Azure Cloud Services which means as soon as we commit changes, the client will be able to monitor them in real-time and provide instant feedback.

I was also assigned the job of implementing a simple front-end design for the settings page, which is used to add/remove search tags and hash tags. I also made some tweaks here and there on the project website to enhance the look and feel of it (for example, adding a logout option to every page, adding a menucon on the tweets display page that takes the user to settings page).

Another task was rendering emoticons inside tweets so that they could be displayed onto the tweets display page. This is necessary because tweets containing emoticons often don’t make sense with emoticons. The task was accomplished by finding all the emoticons and replacing them with <img> tags, just like Twitter does.