

***“Parallel Distributing Computing”***

***Group Name: Dynamic Developers***

**Team Members:**

|  |  |  |
| --- | --- | --- |
| **Name** | **Student Id** | **Class Id** |
| Muhammad Ghous Sarwar | 59455 | 100048 |
| Tahreem Feroz | 59464 | 100048 |
| Syed Fasih |  |  |
| Ismail Javery | 60152 |  |

***“Project Name: Find Sentiment of Tweet on Cluster”***

***Introduction:***

We tweet so frequently nowadays making hashtags and many more. To gather them at a place and splitting them into some positivity and negativity is almost impossible in them. Dynamic Developers are here to collect all live tweets and show them negative, positive or neutral on the basis of testing and analysis.

Each tweet is of approximately 140 characters in length and our team has taken the project to summarize all the tweets on the specific topics.

***Problem Statement:***

* Collecting the data set of the tweets.
* Sorting them in different ways
* Evaluating all the purposes using the distributed computing
* Any topic can be tested by using some sort of algorithms and the tweet will be tested as Negative, positive or neutral.

***Methodology:***

**Distribution of Tweets:**

This could be done in two different ways:

* ***Distribution of words:***

where the words are distributed on different machines. In this case, the final polarity of a tweet cannot be decided until the polarity of individual words has been received.

* ***Distribution of tweets:***

in this method, (complete) tweets are distributed on different machines. In this case, polarity of a tweet can be calculated on a machine.

***Steps Required:***

* Pre-process the tweets. Pre-processing includes but is not limited to stemming, passing through a stop-list, replacing URLs with a keyword URL, replacing user names etc.
* Perform tokenization in order to get words, which be uni-grams, bi-grams or tri-grams.
* Distributing either (complete) tweets or words of a tweet on different machines as discussed in Section 1
* Application of Senti WordNet to find the polarity of a tweet