# Implementation and analysis

To clearly illustrate the effectiveness of the proposed method, a case study is presented with a sample set of tweets.

## Pre-processing of Tweets

After downloading tweets, we use the pre-processing module for data-cleaning module, primarily to remove hashtags, usernames, hyperlinks, RT symbol, punctuations and non-English characters. But before removal of punctuations and non-English characters, we process the emoticons using Emoji Unicode Table.

## POS Tagging

After pre-processing the tweets, we select only adjectives, adverbs and verbs using POS Tagging. The resultant file is a list of tweets that only has adjectives, verbs and adverbs (in the original order), which are referred to as opinion words.

For example, a tweet: “Cortana in Windows 10 is pretty impressive! #Windows10 #WindowsInsiders” is converted to “quite impressive”.

## Emotion Scoring

Once the POS tagging is done, the words are scored either individually (only an adjective) or as a group (adverbs or verbs followed by an adjective).

For example: the above POS tagged tweet is scored as follows:

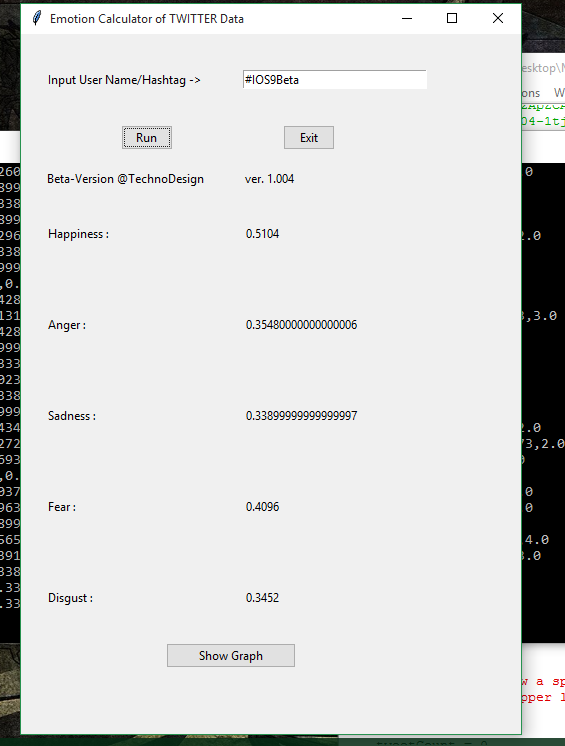
* Here we can see that “pretty” is an adverb and “impressive” is an adjective.
* The adjective emotion values of “impressive” are represented by a vector [4.49, 1.15, 1.15, 1.28, 1.19] ([<Happiness>, <Anger>, <Sadness>, <Fear>, <Disgust>])
* In the list of adverbs we get the values of “pretty” as 0.3
* Now using the algorithm defined in Table III, since the sum of adverbs is 0.3, the emotion values are multiplied by 0.5. And then these values are divided by 5 (only 1 adjective).
* The resultant score of the tweet is given by vector [0.449, 0.115, 0.115, 0.128, 0.119]

Then the tweet is scored by averaging the scores obtained from every adverbs/verbs - adjective group obtained. Then these scores from each tweet are averaged to get a score of the respective hashtag/username. We applied our approach to approx 50,000 tweets for various hashtags/users, for example #*IOS9Beta*, #*WindowsInsiders*, etc. A sample set of 5 tweets with the result of emotion analysis is depicted in the table I below:

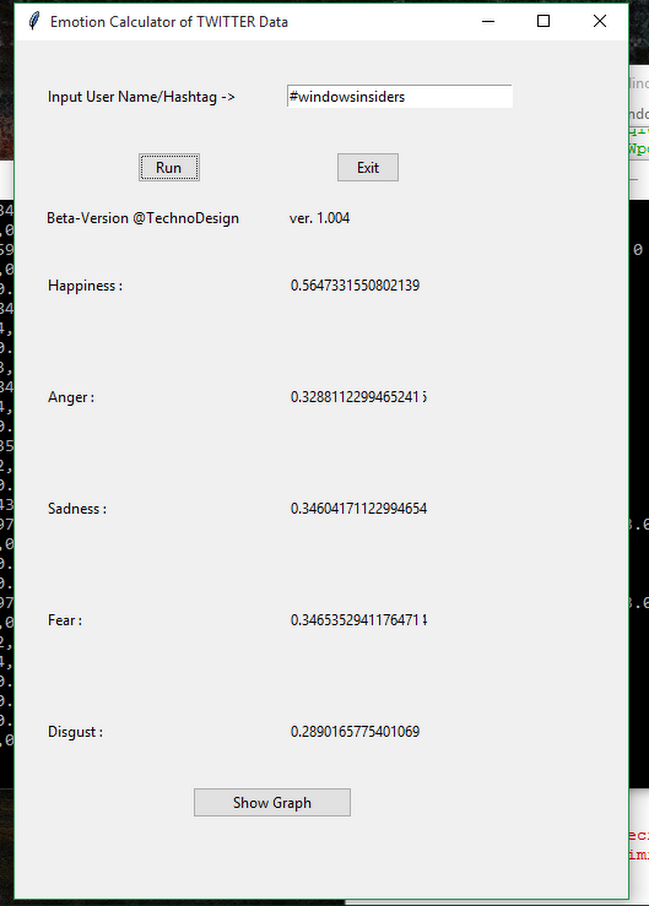
1. Sample Tweet Emotion Values

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Tweet** | **Happiness** | **Anger** | **Sadness** | **Fear** | **Disgust** |
| Great news! [#WindowsInsiders](https://twitter.com/hashtag/WindowsInsiders?src=hash) [#Windows10](https://twitter.com/hashtag/Windows10?src=hash) [#build10532](https://twitter.com/hashtag/build10532?src=hash) | 0.936 | 0.216 | 0.234 | 0.222 | 0.221 |
| Cortana in Windows 10 is pretty impressive! #Windows10 #WindowsInsiders | 0.449 | 0.115 | 0.115 | 0.128 | 0.119 |
| [#Windows10](https://twitter.com/hashtag/Windows10?src=hash) Mail and Calendar Live Tiles don't show any live information. This is so annoying. Is this a known bug? [#windowsinsiders](https://twitter.com/hashtag/windowsinsiders?src=hash) [#wpde](https://twitter.com/hashtag/wpde?src=hash) | 0.214 | 0.798 | 0.398 | 0.212 | 0.196 |
| IOS 9 feels great. #IOS9Beta #IOS9 | 0.812 | 0.226 | 0.211 | 0.201 | 0.191 |
| Woohoo! Today is the day that [@GabeAul](https://twitter.com/GabeAul) pushes the button! Excited [#WindowsInsiders](https://twitter.com/hashtag/WindowsInsiders?src=hash) [#WinPhan](https://twitter.com/hashtag/WinPhan?src=hash) | 0.912 | 0.251 | 0.226 | 0.272 | 0.226 |

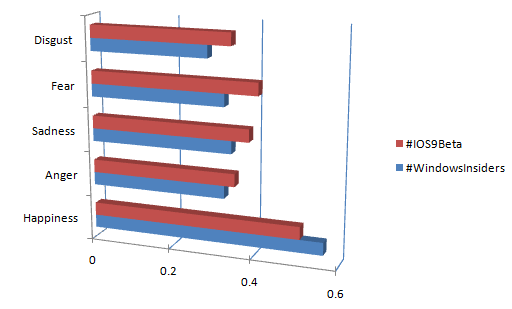
Following fig.1 & fig.2 depict the snapshots of the sample test hashtag *# IOS9Beta* for 246 tweets & # *WindowsInsiders* for 5000 tweets. Fig.3 & fig.4, represent the bar graph and spider chart plots of emotion values calculated displaying the comparisons for both the hashtags*.*

**

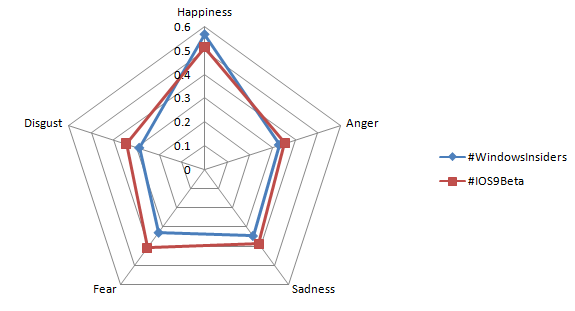
1. For 246 Tweets with hashtag *#IOS9Beta*



1. For 5000 Tweets with hashtag *#WindowsInsiders*

**

1. Emotion Value Bar graph for hashtag *#WindowsInsiders v.s. #IOS9Beta*

**

1. Emotion Value Spider Chart for hashtag *#WindowsInsiders v.s. #IOS9Beta*