

## **First Review Document**

### **Customisable Twitter Bots for Improved QoS in Product Marketing**

Anush Baskaran

13BCE0800

+91-8939007098

anush.baskaran2013@vit.ac.in

Taruni Anand M

13BCE0036

+91-9787119455

tarunianand.m2013@vit.ac.in

Balabhadrapathruni Ramya

13BCE0613

+91-9787119455

bramyasharma@gmail.com

Guide

Jayashree J

Assistant Professor

+91-9790393618

[jayashree.j@vit.ac.in](mailto:jayashree.j@vit.ac.in)

**B.Tech.**

in

**Computer Science and Engineering**

**School of Computer Science & Engineering**



**VIT<sup>®</sup>**

**UNIVERSITY**

(Estd. u/s 3 of UGC Act 1956)

**VELLORE ■ CHENNAI**

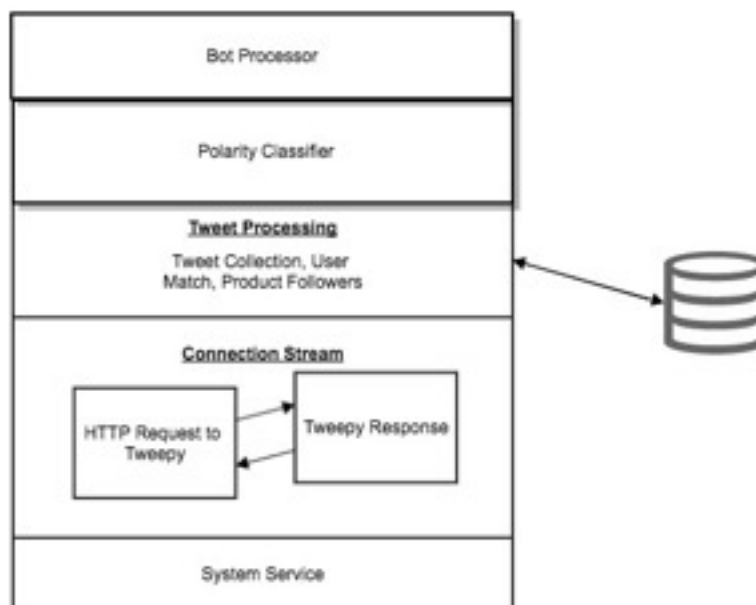


## 1. Abstract:

In recent times, users have started spending a significant amount of their time on social networking platforms (Twitter, Facebook, Orkut), sharing a plethora of personal information and opinions on different aspects of life. Internet users around the world make use of this powerful tool of communication a.k.a microblogging, using Twitter, a web application that provides dual benefits of microblogging and social networking. In this project, we use the open structure of twitter to perform 2 main tasks: (1) analyse data for sentiment analysis and opinion mining of a particular product and classify this data into positive, neutral and negative reviews (2) use this data to create automated programs, known as bots, which are useful for generating benign tweets. We use this double edged sword to provide companies with the opportunity to make use of these automated bots to generate the right tags to advertise products. This enhanced marketing opportunity provides Quality of Service in terms of cost, time and energy.

**Keywords:** *Twitter, microblogging, sentiment analysis, opinion mining, automation, bots, QoS.*

## 2. System Design:



## **2.1. System Service**

The system service component invokes all other components and provides a command line console for twitter to process its search for tweets, based on the search command.

## **2.2. Connection Stream**

Once the HTTP server pings the Tweepy API with the search option, Tweepy induces the generation of the Tweet processing stream. This HTTP connection is open till a single request is served and is open every time a request is made.

## **2.3. Tweet Processing**

The Tweet Processing component is generate during the middle of the Connection Stream process. The collection of new published tweets containing the search term is gathered. It also gathers tweet status id, tweet author id, and tweet author name. Information collected is stored in the database. This component also gathers the product tweeted about's followers.

## **2.4. Polarity Classifier**

The polarity classifier is used to analyse tweets to determine whether a tag is positive, negative or neutral. The results are then written to the database to be used in the next stage. While neutral and negative tweets are ignored, the positive tweets are taken to be used for marketing purposes.

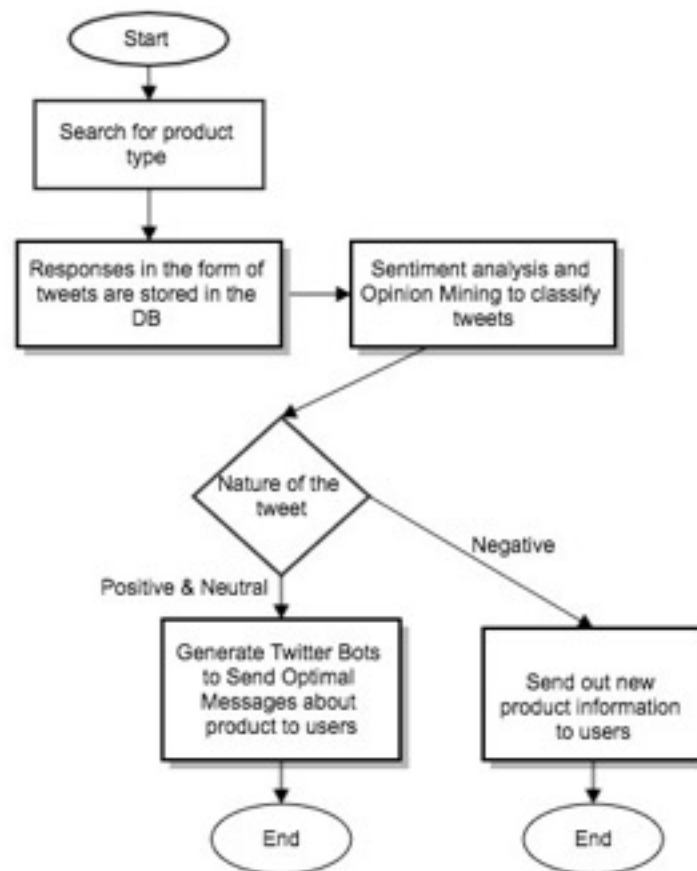
## **2.5. Bot Processor**

The bot processor is responsible for analysis positive reviews to gather the user\_id of the person who tweeted, and tag him in an automated tweet that a company wishes to publish. Also, in order to make sure that the messages do not spam user accounts by continuously tagging them, a counter is set to tag them in optimal number of posts, in order to serve a positive purpose.

## **3. Methodology**

The scenario considers an example use case where the automated twitter bot software being created is being used by a new company to market its product. The product is being searched for at the system service stage. This is analogous to a user requesting for a resource from Twitter. In order for us to extract data from the web application, we are using a Twitter API called Tweepy which sets up a HTTP request to the twitter servers and the request is served in the forms of

tweets relating to the type of the product. Along with the tweets, the user id and specific parts of the tweet are stored in the database for processing at a later stage. The tweet is then analysed at the next stage also known as sentiment analysis and opinion mining using a classifier algorithm to predict whether the tweets are neutral, positive or negative. The users who gave out positive reviews will then be tagged in optimal number of posts that the new company marketing a similar product makes. The negative reviewers will be tagged in new product posts that the company has to make. This is the stage when the company has to take a chance and find out whether they can acquire a new customer. In order to save the companies time, cost and manpower, twitter bots will be used to automate the process of tagging and putting up these posts on behalf of the company. The flow chart given below explains the process described above.



#### **4. Expected Results**

1. Successful approximate classification of tweets on the basis of the input criteria.
2. Complete classification cannot be made due to varying nature/ tone of the tweets. (Refer to point 2.)
3. Analysing the nature of the tweets based on the classification. This is done on the basis of historic statistics resulting from existing research.
4. Bots tweeting to specific target audience on the basis of results from point 3.
5. Overcomes existing limitations (spam) through the users identified as a result of analysis as mentioned above.
6. Application designed for customers to easily customise their bots.

#### **5. Hardware and Software Requirements**

##### **5.1. User Interfaces**

The proposed system is a web application. The user interface will be limited to the types of controls that can be generated using HTML, Javascript, and Cascading Style Sheets. The user interface code will be generated by individual developers, as well as by Anaconda for Python.

##### **5.2. Hardware Interfaces**

The Hardware Interfaces of the system are handled by the Mac OS and the Windows 10 Operating System. No hardware dependent code will be written by the team in Phase 1 of the Customisable Twitter Bots system.

### 5.3. Software Interfaces

- **Operating System**

- The software is being designed to run on Windows 10 and Mac OS. This is specifically in order to have compatible internet browsers.

- **Web Server**

- The software is being designed to run on Internet Web browsers.

- **Database**

- The software will access MongoDB for storing data collected from the analysis phase.

- **Libraries / API**

- The software project utilises the Tweepy module from the list of Twitter API's. The application will be created using the Django Framework.

### **References**

#### Weblinks:

1. <https://marcobonzanini.com/2015/03/02/mining-twitter-data-with-python-part-1/>
2. [http://hanj.cs.illinois.edu/bk3/bk3\\_slidesindex.htm](http://hanj.cs.illinois.edu/bk3/bk3_slidesindex.htm)
3. <http://dmml.asu.edu/smm/slides/>
4. <http://ecs.syr.edu/faculty/reza/tutorials/ICDM13/TutorialICDM13SMM.pdf>
5. <http://blogs.ischool.berkeley.edu/i290-abdt-s12/>

#### Journal/Conference:

1. Ashwin Rajadesingan, Reza Zafarani, Huan Liu “Sarcasm Detection on Twitter: A Behavioural Modelling Approach”, WSDM '15 Proceedings of the Eighth ACM International Conference on Web Search and Data Mining. Pp 97-106.
2. Reza Zafarani “Behaviour Analysis in Social Media”, IEEE Intelligent Systems, Volume 29, Issue 4, 2014, Pp 9-11.

3. S. Haustein, T. D. Bowman, K. Holmberg, A. Tsou, C. R. Sugimoto, and V. Larivière, "Tweets as impact indicators: Examining the implications of automated 'bot' accounts on Twitter," *Journal of the Association for Information Science and Technology*, vol. 67, no. 1, pp. 232–238, May 2015.
4. E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "Title: The rise of social Bots," 2014. [Online]. Available: <https://arxiv.org/abs/1407.5225>. Accessed: Feb. 18, 2017.
5. Jalal S. Alowibdi, Ugo A. Buy, Philip S. Yu, Leon Stenneth, "Detecting deception in Online Social Networks", *Advances in Social Networks Analysis and Mining (ASONAM) 2014 IEEE/ACM International Conference on*, pp. 383-390, 2014.
6. Gregory Maus, "Decoding hacking and optimizing societies: Exploring potential applications of human data analytics in sociological engineering both internally and as offensive weapons", *Science and Information Conference (SAI) 2015*, pp. 538-547, 2015.
7. CM Zhang, V. Paxson, "Detecting and analyzing automated activity on twitter", In *International Conference on Passive and Active Network Measurement 2011*, Mar 20, pp. 102-111, 2011.