Opinionated Tweet Mining Implicating the Rudimentary 'R' Language

Ram Chatterjee
Dept. Computer Science and Technology, Manav Rachna University (MRU), Faridabad, India.
ram@mru.edu.in

Monika Goyal
Dept. Computer Science and Technology, Manav Rachna University (MRU), Faridabad, India.
monikagoyal27@gmail.com

Abstract

This paper explicates programming technique of extracting twitter data, the micro blog that serves as a rich source for sentiment analysis via the opinions shared by millions of users. The paper primarily focuses on the detailed implementation of 'R' language for twitter data extraction. "Windows 10" has been used as a keyword to extract relevant opinions from the twitter. These extracted tweets on beta release of windows 10 provide an aid during customer feedback loop by the manufacturers.

Keywords: APIs, Opinion Mining, R Language, Sentiment, Tweet, TwitteR.

1. Introduction

Opinion Mining is attributable to the discipline of Natural language processing (NLP) and Information extraction (IE) and connotes to trail the sentiments of people on a particular topic, product or individual. Plethora of opinions, emotions and attitudes about a topic or product in textual form do exist on the web. Using the techniques for opinion mining the relevant information out of voluminous data is extracted in a useful manner. The polarity of the opinions extracted may be attributed as positive, negative or neutral which help the customer in many ways.

To elaborate further, we may consider an instance where a buyer intending to purchase a camera seeks the product comments and reviews, posted by individuals who have just bought the product or have been using it for quite some time. Analogously, the same customer feedback (opinion) on the product camera would benefit the company manufacturing the product; improve its product quality, influencing its marketing. Thus, opinionated mining on the review/feedback data benefits the buyer in forming the right buying decision and the manufacturer to better their products.

The rest of the paper is organized as follows: section 2 discusses about the social media and micro blogging and section 3 discusses significance of twitter in opinion mining. In section 4, we have implemented programming method 'R' whereas section 5 shows various non-programming methods and finally, section 6, provides the conclusions of our study.

2. Theoretical underpinnings of social media and micro -blogging

An important medium, serving as a consolidated and interactive platform for dissemination of information, thoughts and opinions effectively and efficiently, has been the social web [7] in the form of blogs and social networking sites, for creating and exchanging user generated content. Social media has gained abundant popularity in the past decade which is evident by its ubiquity as a powerful source of news updates, networking, entertainment, viral marketing and online collaboration [8]. It is for the readers' interest and knowledge we categorize the most popular social media forms in current use, below.

2.1 Social Networks

The web usage growing abundantly have triggered rapid growth of social networking sites like Facebook, LinkedIn and the likes [6], permitting users to create web pages enabling them to share content, communications and media with their colleagues, friends and other users. This procreates voluminous data to be assimilated for sentiment analysis.

2.2 Web blogs

Blogs as online journals permit people to write about the topics they want to share with others serving as a source of opinion in sentiment analysis [2]. Adding an article to a blog is 'blogging' and these individual articles are known as 'blog post' or 'entries'. A person who posts these entries is known as 'blogger'.

2.3 Micro blogs

Micro blogs [5] are most popular blogging tools that allow users to post short texts of few lines, upload images and videos. Micro blogging promotes content distribution through mobile devices and serves as common medium for posting quick updates. Billions of users share opinions on different aspects of life everyday on Twitter [8].

2.4 Content Communities

Content communities are the platforms attributed towards content, image and video - organization, sharing and commenting. YouTube, Flickr and scribed are some of the famous content communities.

2.5 Review Sites

E-Commerce has been very popular in the present era, providing a great interacting platform between consumers and manufacturers. The product reviews and feedbacks are vital for both the customers intending to buy a product and also to the merchants. A large number of user generated reviews about product, restaurant, mobile, etc. is available on the internet that helps user in making purchasing decisions [3].

3. Significance of Twitter Data in Opinion Mining

The most well-liked micro-blogging platform is twitter which contains a very large number of short messages called tweet. Each tweet is limited to 140 characters and this upper bound make users to construct focused updates. Twitter has over 284 million monthly active user's accounts which share 500 million tweets per day. As we can see the audience of twitter grows every day, this data can be efficiently used in opinion mining and sentiment analysis tasks.

To proceed with the task of opinion mining, twitter micro blogs have been taken into consideration. Readers would appreciate to know the following reasons [1] justifying usage of twitter data for applying sentiment analysis.

- Micro blogging platforms form basis for diverse people to express their opinion about varied topics, lending itself to a valuable source to access people's opinions.
- Twitter is attributed with massive text posts growing rapidly everyday leading to a collection of arbitrarily large corpus.
- The plethora of twitter audiences span from regular users to celebrities, making it possible to collect posts from users belonging to different social and interests groups.
- Twitter users span across several countries, largely comprising users from U.S. owing to which it is possible to collect twitter data in varied languages.

To fulfill the objective of applying sentiment analysis we first need to proceed with collection of tweets by implementing various methods for twitter data extraction as explicated and exemplified in this paper. Following the tweets extraction we proceed with building the corpus of retrieved texts posts from twitter. The corpus is then classified under three sets of texts as: positive emotions, negative emotions or no emotions. Finally, experimental evaluations are carried out to ascertain the polarity of the texts in the given corpus [4].

4. Programming Method of Twitter Data Extraction

Extraction of twitter data can be possible using programming techniques and non-programming techniques. In programming technique we have open source R language working across different platforms such as Windows, Linux, Mac etc. It is viewed as the scripting language for the R environment. R is used for deep statistical analysis in a variety of applications including data mining and visualizations. R language has System and User packages through which we are able to extract information from twitter API for Opinion Mining. Some of the important package of R includes:

- **twitteR**: This package provides an R based twitter client.
- **ROAuth**: This package provides an interface to the OAuth 1.0 specification, allowing users to authenticate via OAuth to the server of their choice.
- **stringr**: stringr is used for strings to make them consistent and simpler to use. It comprise of all functions to deal with zero length character appropriately.
- **httr**: This package provides useful tools for working with HTTP. The API is based around http verbs (GET (), POST (), etc) with pluggable components to control the request (authenticate (), add_headers() and so on).

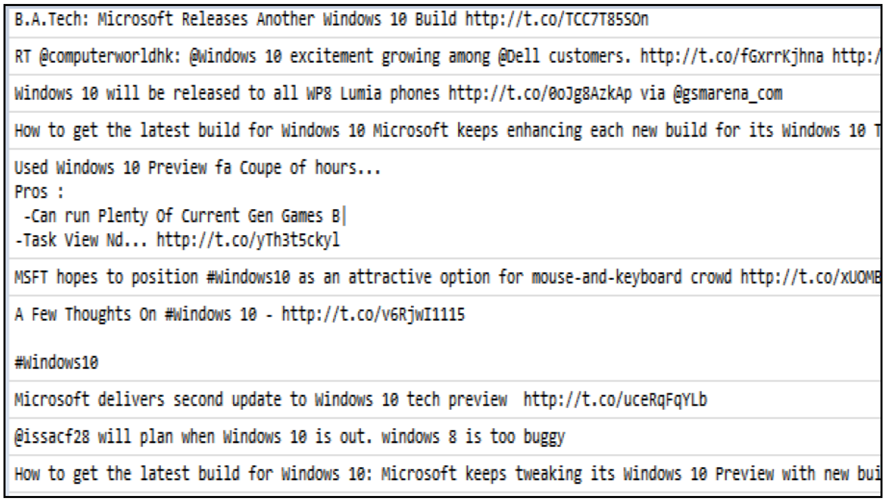
- **rjson**: This package converts R object into JSON objects and vice-versa. JSON is used to retain and display data. JSON connotes to a key-value data store that can map directly to JavaScript objects.
- **jsonlite:** It is a robust and high performance json parser and generator for R
- **bit64**: It provides S3 class for vectors and matrices of 64bit (signed) integers. It helps to perform fast algorithmic operations such as 'match' and 'order' support interactive data manipulation.
- **devtools:** This package contains collection of development tools to make the code easier.

We can classify the twitter programming APIs as Search and Streaming APIs. They both serve for tweet collection and are the indispensable parts of Twitter programming. However, the difference lies between two in the searching mechanism. The search APIs back traces in time and streaming traces forward in time. For the maximum efficiency we need to exercise right judgment on usage of each API at apt time. If we have nothing to start with, the search APIs is the best option. We collect tweets on a subject to fill in the past days tweets. This is often attributed as back-filling. With a database populated up to the present moment, we can then turn on the streaming API and confine all tweets going forward. The examples of twitter search API used for tweets extraction are given below:

4.1 Twitter Search API and R

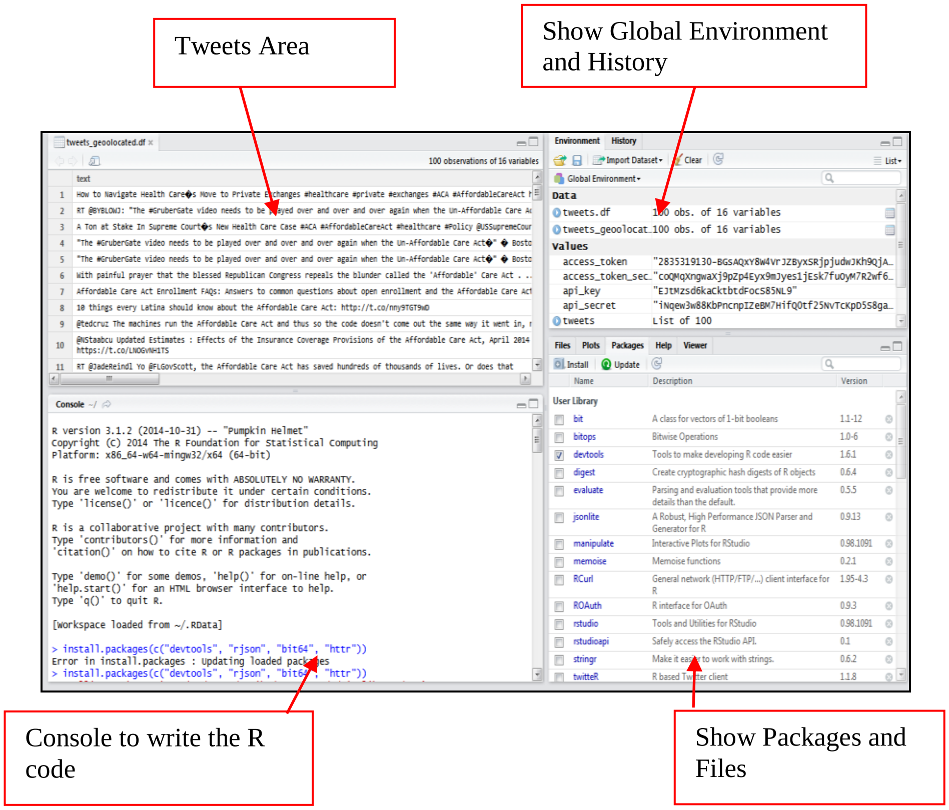
We use R packages to query the Twitter search API in order to collect tweets into R for Opinion Mining. To accomplish this task we firstly, create an application using our twitter account and then Install R and required R packages. Next we will authenticate twitter Search API and lastly, we used to run query using R query structure and save to database.

We established a secure connection between R console and twitter. This requires logging to our respective twitter account in order to create a new twitter app under “My Applications” by clicking “Create New App”. A suitable name, description, and website are provided to the app to access it in order to fetch tweets. Once created, we can also change the permissions of app to allow the application to read, write and access the messages under “modify app permissions”. To fire query using R we also required token actions which can be generated under “Create my access token.” After the twitter app set up, we need to install R and its required packages like devtools, twitteR, httr, ROAuth. We used the Rstudio, an integrated development environment and GUI for R. Then, twitter authentication had done using setup_twitter_oauth() function provided customer key, customer secret key, access token, access token secret as arguments as per the New App Credentials created above. After a successfully authenticated we had searched the twitter for tweets extraction using a function of twitter package as searchTwitter().The search terms, no.of tweets and language of tweets needed are few of its arguments need to be passed. We had used “Windows 10” as a search term to extract twitter data. The extracted tweets using R packages are displayed in Figure1. Moreover, Figure 2 provides a basic outline of Rstudio IDE and its various working areas. The various packages and their dependencies as install during the process is illustrated in Figure 3. The environment variables used during the process are depicted in Figure 4.



```
B.A.Tech: Microsoft Releases Another Windows 10 Build http://t.co/TCC7T85SOn

RT @computerworldhk: @Windows 10 excitement growing among @Dell customers. http://t.co/fGxrrKjhna http:/

Windows 10 will be released to all WP8 Lumia phones http://t.co/0oJg8AzkAp via @gsmarena_com

How to get the latest build for Windows 10 Microsoft keeps enhancing each new build for its Windows 10 T

Used Windows 10 Preview fa Coupe of hours...
Pros :
 -Can run Plenty Of Current Gen Games B|
-Task View Nd... http://t.co/yTh3t5ckyl

MSFT hopes to position #Windows10 as an attractive option for mouse-and-keyboard crowd http://t.co/xUOMB

A Few Thoughts On #Windows 10 - http://t.co/v6RjwI1115

#Windows10

Microsoft delivers second update to Windows 10 tech preview  http://t.co/uceRqFqYLb

@issacf28 will plan when Windows 10 is out. windows 8 is too buggy

How to get the latest build for Windows 10: Microsoft keeps tweaking its Windows 10 Preview with new bui
```

Figure 1: Tweets on Windows 10 as extracted using R packages



100 observations of 16 variables

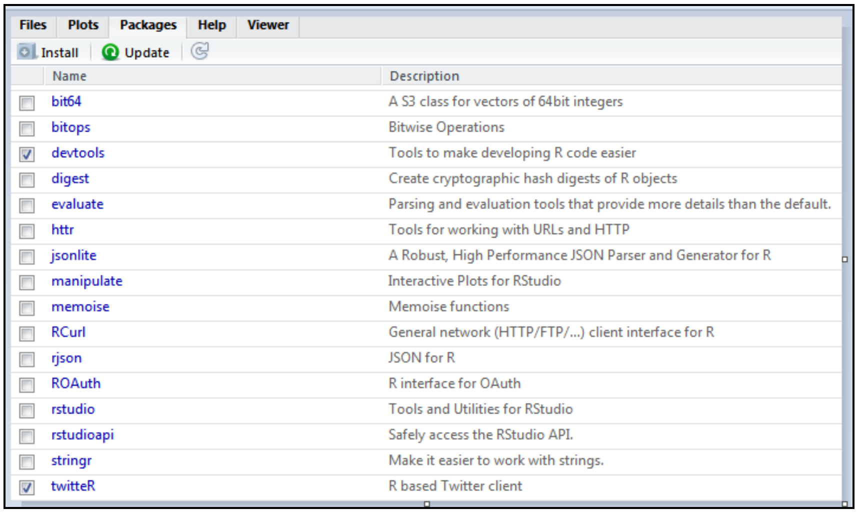
	text
1	How to Navigate Health Care�s Move to Private Exchanges #healthcare #private #exchanges #ACA #AffordableCareAct h
2	RT @BYBLOWJ: "The #GruberGate video needs to be played over and over and over again when the Un-Affordable Care Ac
3	A Ton at Stake In Supreme Court�s New Health Care Case #ACA #AffordableCareAct #healthcare #Policy @USSupremeCour
4	"The #GruberGate video needs to be played over and over and over again when the Un-Affordable Care Act�" � Bosto
5	"The #GruberGate video needs to be played over and over and over again when the Un-Affordable Care Act�" � Bosto
6	With painful prayer that the blessed Republican Congress repeals the blunder called the 'Affordable' Care Act . ..
7	Affordable Care Act Enrollment FAQs: Answers to common questions about open enrollment and the Affordable Care Act
8	10 things every Latina should know about the Affordable Care Act: http://t.co/nny9TGT9WD
9	@tedcruz The machines run the Affordable Care Act and thus so the code doesn't come out the same way it went in, r
10	@NStaabcu Updated Estimates : Effects of the Insurance Coverage Provisions of the Affordable Care Act, April 2014 https://t.co/LNOGvNH1TS
11	RT @JadeReindl Yo @FLGovScott, the Affordable Care Act has saved hundreds of thousands of lives. Or does that

```
R version 3.1.2 (2014-10-31) -- "Pumpkin Helmet"
Copyright (C) 2014 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

> install.packages(c("devtools", "rjson", "bit64", "httr"))
Error in install.packages : Updating loaded packages
> install.packages(c("devtools", "rjson", "bit64", "httr"))
```

```
Data
tweets.df                100 obs. of 16 variables
tweets_geoolocat…        100 obs. of 16 variables
Values
  access_token           "2835319130-BGsAQxY8W4VrJZByxSRjpjudwJKh9QjA…
  access_token_sec…      "coQMqXngwaXj9pZp4Eyx9mJyes1jEsk7fuOyM7R2wf6…
  api_key                "EJtMzsd6kaCktbtdFocS85NL9"
  api_secret             "iNqew3w88KbPncnpIZeBM7HifQ0tf25NvTcKpD5S8ga…
tweets                   List of 100
```

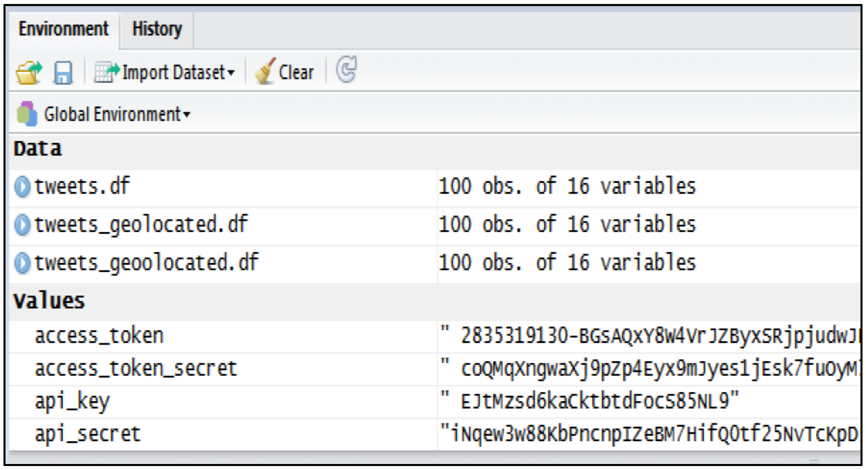
	Name	Description	Version
	User Library		
□	bit	A class for vectors of 1-bit booleans	1.1-12
□	bitops	Bitwise Operations	1.0-6
☑	devtools	Tools to make developing R code easier	1.6.1
□	digest	Create cryptographic hash digests of R objects	0.6.4
□	evaluate	Parsing and evaluation tools that provide more details than the default.	0.5.5
□	jsonlite	A Robust, High Performance JSON Parser and Generator for R	0.9.13
□	manipulate	Interactive Plots for RStudio	0.98.1091
□	memoise	Memoise functions	0.2.1
□	RCurl	General network (HTTP/FTP/...) client interface for R	1.95-4.3
□	ROAuth	R interface for OAuth	0.9.3
□	rstudio	Tools and Utilities for RStudio	0.98.1091
□	rstudioapi	Safely access the RStudio API.	0.1
□	stringr	Make it easier to work with strings.	0.6.2
□	twitteR	R based Twitter client	1.1.8

Figure 2. Rstudio IDE



	Name	Description
□	bit64	A S3 class for vectors of 64bit integers
□	bitops	Bitwise Operations
☑	devtools	Tools to make developing R code easier
□	digest	Create cryptographic hash digests of R objects
□	evaluate	Parsing and evaluation tools that provide more details than the default.
□	httr	Tools for working with URLs and HTTP
□	jsonlite	A Robust, High Performance JSON Parser and Generator for R
□	manipulate	Interactive Plots for RStudio
□	memoise	Memoise functions
□	RCurl	General network (HTTP/FTP/...) client interface for R
□	rjson	JSON for R
□	ROAuth	R interface for OAuth
□	rstudio	Tools and Utilities for RStudio
□	rstudioapi	Safely access the RStudio API.
□	stringr	Make it easier to work with strings.
☑	twitteR	R based Twitter client

Figure 3. User Packages used for twitter extraction in Rstudio IDE



Data	
tweets.df	100 obs. of 16 variables
tweets_geolocated.df	100 obs. of 16 variables
tweets_geoolocated.df	100 obs. of 16 variables
Values	
access_token	" 2835319130-BGsAQxY8W4VrJZByxSRjpjudwJ
access_token_secret	" coQMqXngwaXj9pZp4Eyx9mJyes1jEsk7fuOyM
api_key	" EJtMzsd6kaCktbtdFocS85NL9"
api_secret	"iNqew3w88KbPncnpIZeBM7HifQ0tf25NvTcKpD

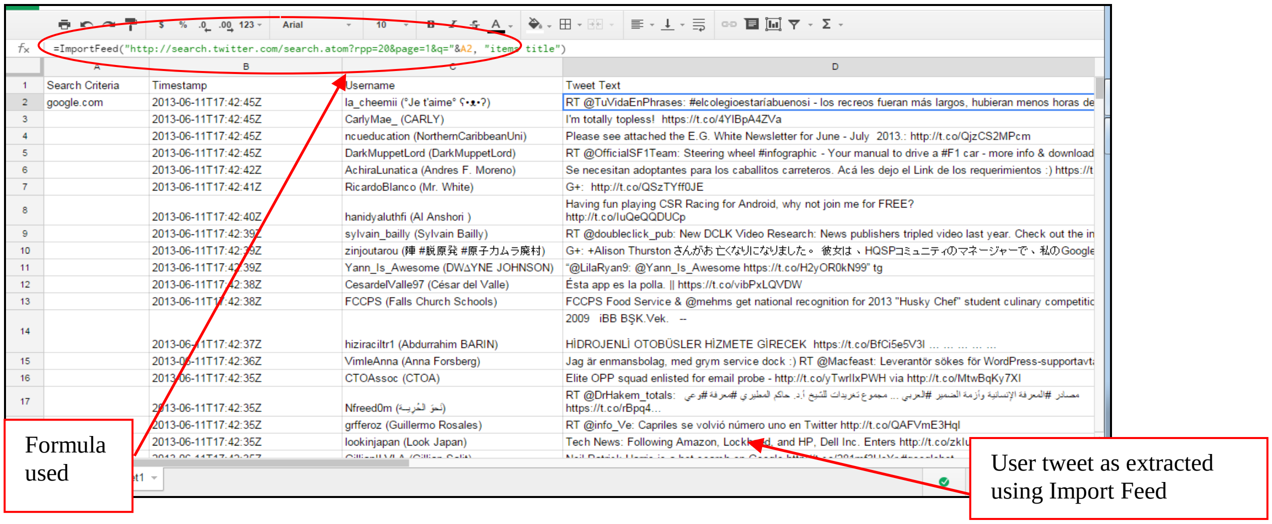
Figure 4. Environment history of twitter data extraction Process

4.2 Twitter Search API with Import Feed

It is the Twitter Search API in concurrence with Import Feed that serves to extract tweets in a structured and easily manageable format like CSV or XLS. To accomplish twitter data extraction, we first proceed to create a Google Spreadsheet with intent to capture twitter data attributed with time, username and text of all tweets corresponding to the links pertaining to a specified page. The Google spreadsheet columns are populated with labels viz. “Search criteria”, “Timestamp”, “Username” and “Tweet text” in order in the cells spanning from A1 to D1. Following this the cells B2, C2, and D2 underneath column labels “Timestamp”, “Username”, “Tweet Text”, the following formulas are entered respectively as displayed in Figure 5 and 6.

```
Import Feed ("http://search.twitter.com/search.atom?rpp=20&page=1&q="&A2, "items created")

Import Feed ("http://search.twitter.com/search.atom?rpp=20&page=1&q="&A2, "items author")

Import Feed ("http://search.twitter.com/search.atom?rpp=20&page=1&q="&A2, "items title")
```

A search query, for instance “Nokia” is then entered into cell A2. On trigger of the enter button, the results relevant to the query gets loaded into the spreadsheet. The extracted tweets are displayed in Figure 7. Readers may note that in order to search for hash tags and usernames, we need to use %23 instead of # and %40 instead of @.



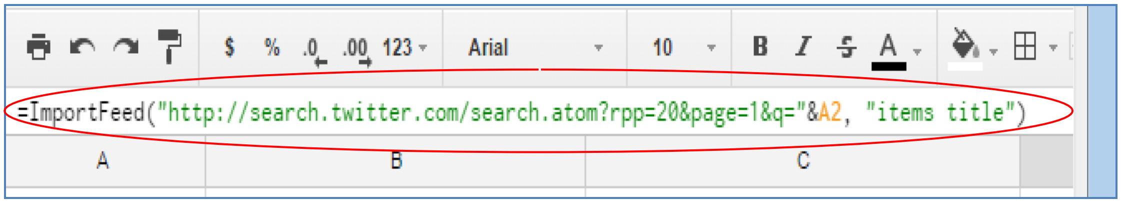
```
=ImportFeed("http://search.twitter.com/search.atom?rpp=20&page=1&q="&A2, "items title")
```

	A	B	C	D
1	Search Criteria	Timestamp	Username	Tweet Text
2	google.com	2013-06-11T17:42:45Z	la_cheemii (°Je t'aime° ʕ•ᴥ•ʔ)	RT @TuVidaEnPhrases: #elcolegioestaríabuenosi - los recreos fueran más largos, hubieran menos horas de
3		2013-06-11T17:42:45Z	CarlyMae_ (CARLY)	I'm totally topless! https://t.co/4YIBpA4ZVa
4		2013-06-11T17:42:45Z	ncueducation (NorthernCaribbeanUni)	Please see attached the E.G. White Newsletter for June - July 2013.: http://t.co/QjzCS2MPcm
5		2013-06-11T17:42:45Z	DarkMuppetLord (DarkMuppetLord)	RT @OfficialSF1Team: Steering wheel #infographic - Your manual to drive a #F1 car - more info & download
6		2013-06-11T17:42:42Z	AchiraLunatica (Andres F. Moreno)	Se necesitan adoptantes para los caballitos carreteros. Acá les dejo el Link de los requerimientos :) https://t
7		2013-06-11T17:42:41Z	RicardoBlanco (Mr. White)	G+: http://t.co/QSzTYff0JE
8		2013-06-11T17:42:40Z	hanidyaluthfi (Al Anshori)	Having fun playing CSR Racing for Android, why not join me for FREE? http://t.co/luQeQQDUCp
9		2013-06-11T17:42:39Z	sylvain_bailly (Sylvain Bailly)	RT @doubleclick_pub: New DCLK Video Research: News publishers tripled video last year. Check out the in
10		2013-06-11T17:42:39Z	zinjoutarou (陣 #脱原発 #原子力ムラ廃村)	G+: +Alison Thurston さんがお亡くなりになりました。 彼女は、HQSPコミュニティのマネージャーで、私のGoogle
11		2013-06-11T17:42:39Z	Yann_Is_Awesome (DW∆YNE JOHNSON)	“@LilaRyan9: @Yann_Is_Awesome https://t.co/H2yOR0kN99” tg
12		2013-06-11T17:42:38Z	CesardelValle97 (César del Valle)	Ésta app es la polla. || https://t.co/vibPxLQVDW
13		2013-06-11T17:42:38Z	FCCPS (Falls Church Schools)	FCCPS Food Service & @mehms get national recognition for 2013 "Husky Chef" student culinary competitic
14		2013-06-11T17:42:37Z	hiziraciltr1 (Abdurrahim BARIN)	2009 iBB BŞK.Vek. -- HİDROJENLİ OTOBÜSLER HİZMETE GİRECEK https://t.co/BfCi5e5V3I ... ...
15		2013-06-11T17:42:36Z	VimleAnna (Anna Forsberg)	Jag är enmansbolag, med grym service dock :) RT @Macfeast: Leverantör sökes för WordPress-supportavt
16		2013-06-11T17:42:35Z	CTOAssoc (CTOA)	Elite OPP squad enlisted for email probe - http://t.co/yTwrlIxPWH via http://t.co/MtwBqKy7XI
17		2013-06-11T17:42:35Z	Nfreed0m (نَحوَ الحُرية)	RT @DrHakem_totals: مصادر #المعرفة الإنسانية وأزمة الضمير #العربي ... مجموع تغريدات للشيخ أ.د. حاكم المطيري #معرفة #وعي https://t.co/rBpq4...
18		2013-06-11T17:42:35Z	grfferoz (Guillermo Rosales)	RT @info_Ve: Capriles se volvió número uno en Twitter http://t.co/QAFVmE3Hql
19		2013-06-11T17:42:35Z	lookinjapan (Look Japan)	Tech News: Following Amazon, Lockheed, and HP, Dell Inc. Enters http://t.co/zklu

Formula used

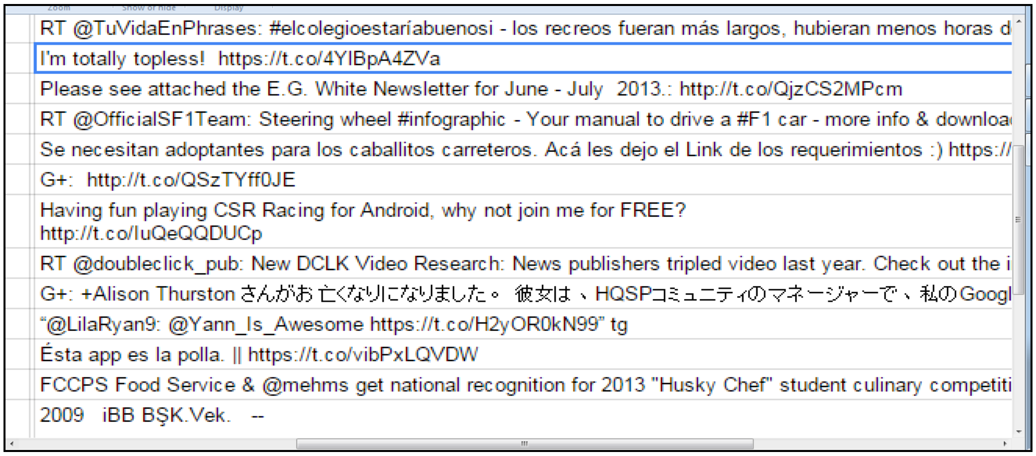
User tweet as extracted using Import Feed

Figure 5. Example of Twitter search API using Google spreadsheet



```
=ImportFeed("http://search.twitter.com/search.atom?rpp=20&page=1&q="&A2, "items title")
```

Figure 6. Depicting the Formula used in Twitter search API



RT @TuVidaEnPhrases: #elcolegioestaríabuenosi - los recreos fueran más largos, hubieran menos horas d

I'm totally topless! https://t.co/4YIBpA4ZVa

Please see attached the E.G. White Newsletter for June - July 2013.: http://t.co/QjzCS2MPcm

RT @OfficialSF1Team: Steering wheel #infographic - Your manual to drive a #F1 car - more info & downloa

Se necesitan adoptantes para los caballitos carreteros. Acá les dejo el Link de los requerimientos :) https://

G+: http://t.co/QSzTYff0JE

Having fun playing CSR Racing for Android, why not join me for FREE?
http://t.co/luQeQQDUCp

RT @doubleclick_pub: New DCLK Video Research: News publishers tripled video last year. Check out the i

G+: +Alison Thurston さんがお亡くなりになりました。 彼女は、HQSPコミュニティのマネージャーで、私のGoogl

"@LilaRyan9: @Yann_Is_Awesome https://t.co/H2yOR0kN99" tg

Ésta app es la polla. || https://t.co/vibPxLQVDW

FCCPS Food Service & @mehms get national recognition for 2013 "Husky Chef" student culinary competiti

2009 iBB BŞK.Vek. --

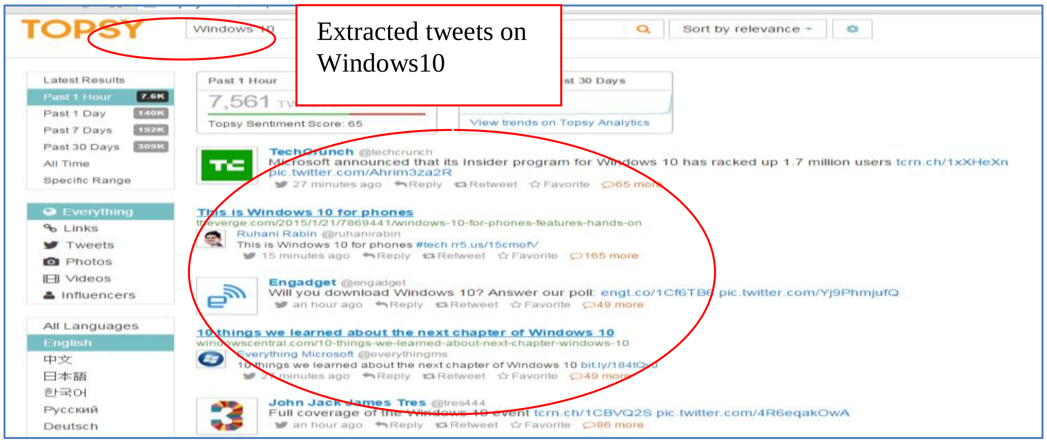
Figure 7. Tweets extracted using Google Spreadsheets on "google.com" as a keyword

The Twitter API programming is advantageous in the way it permits to add value to a collection of tweets as compared to non-programming techniques:

- We can apply quality control rules that filter out false positives for the keywords we are using in our collection query.
- A simple language detection algorithm promotes tweeting for explicit language, eliminating all other languages.
- It is by the virtue of identification of spamming words and blacklisting them, a high percentage of spam tweets can be eradicated.
- We can also exclude accounts that have a spammy profile or an account with recent existence since few days.
- Usage of an influence algorithm, such as follower count or frequency of mentions, proves helpful to select tweets from the most influential users.

5. Non-programming method of Twitter data extraction

We use various sentiment tools as present over web for twitter data extraction under non programming techniques. They are actually a time saver in order to get what users want. User only need to provide with a keyword for which one want to extract tweets and within a minute hundreds of tweets displayed on the user screen. Non-programming methods are easy and flexible to use but create difficulty in building a corpus. Some of the sentiment tools that can be used for the twitter data extraction include topsy in Figure 8, Streamcrab in Figure 9, Sentiment viz (virtualization) in Figure 10, trackur, Sentiment140, Tweet archivist. Topsy is good option to use if user wants large number of tweets on the given keyword while sentiment viz (virtualization) provides various different virtualization techniques to look tweets and user sentiments in better way. Hence, each sentiment tool is different from another to extract tweets and the choice of tool depends on the application.



Extracted tweets on Windows10

TechCrunch @techcrunch
Microsoft announced that its Insider program for Windows 10 has racked up 1.7 million users tcrn.ch/1xXHeXn pic.twitter.com/Ahrim3za2R
27 minutes ago Reply Retweet Favorite 65 more

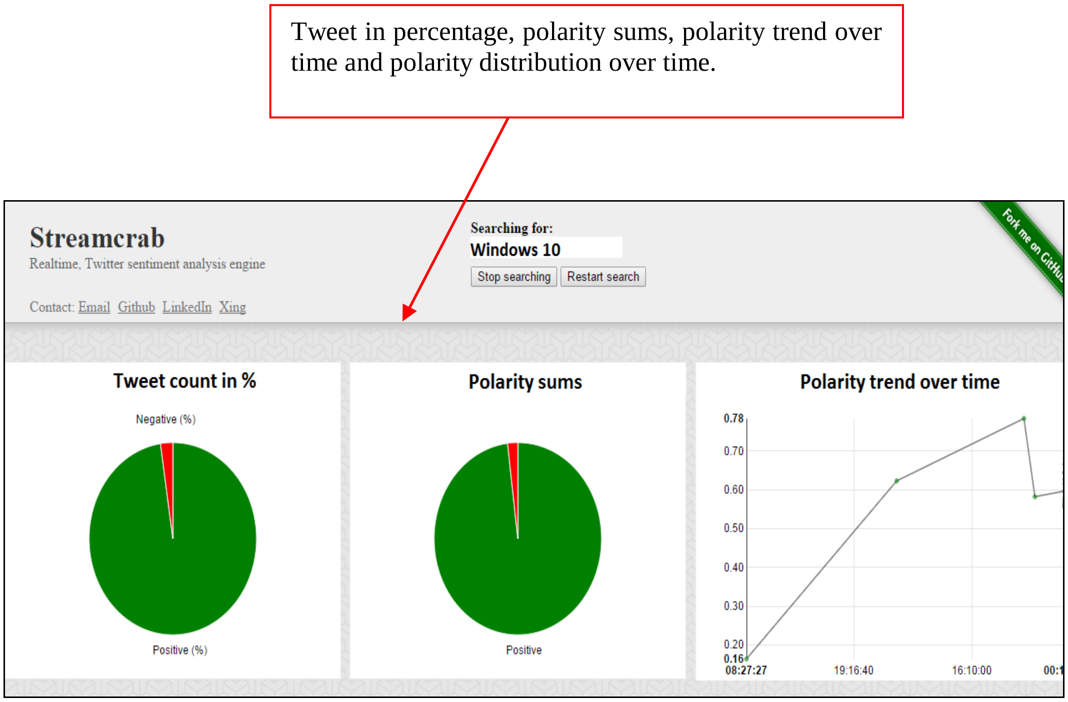
This is Windows 10 for phones
theverge.com/2015/1/21/7869441/windows-10-for-phones-features-hands-on
Ruhani Rabin @ruhanirabin
This is Windows 10 for phones #tech rr5.us/15cmofV
15 minutes ago Reply Retweet Favorite 165 more

Engadget @engadget
Will you download Windows 10? Answer our poll: engt.co/1Cf6TB6 pic.twitter.com/Yj9PhmjufQ
an hour ago Reply Retweet Favorite 49 more

10 things we learned about the next chapter of Windows 10
windowscentral.com/10-things-we-learned-about-next-chapter-windows-10
Everything Microsoft @everythingms
10 things we learned about the next chapter of Windows 10 bit.ly/184tQud
27 minutes ago Reply Retweet Favorite 49 more

John Jack James Tres @tres444
Full coverage of the Windows 10 event tcrn.ch/1CBVQ2S pic.twitter.com/4R6eqakOwA
an hour ago Reply Retweet Favorite 86 more

Figure 8. Topsy

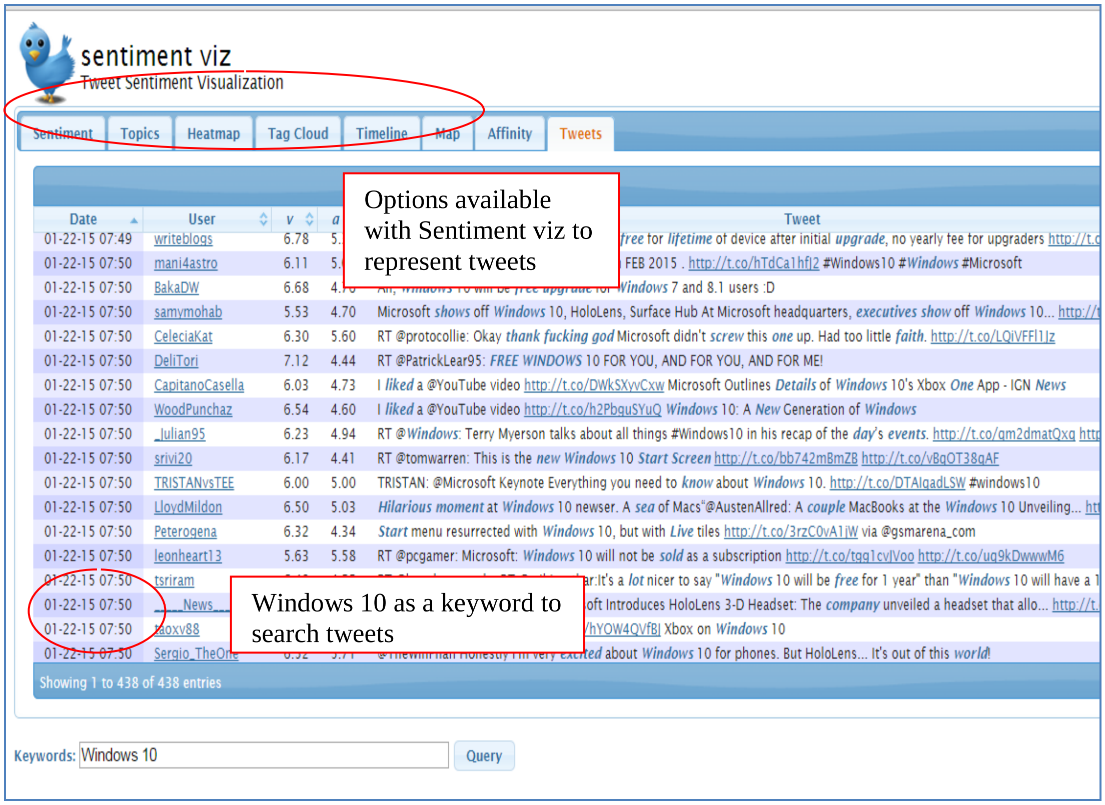

Tweet in percentage, polarity sums, polarity trend over time and polarity distribution over time.

Streamcrab
Realtime, Twitter sentiment analysis engine

Contact: Email Github LinkedIn Xing

Searching for:
Windows 10

Figure 9. Streamcrab



Options available with Sentiment viz to represent tweets

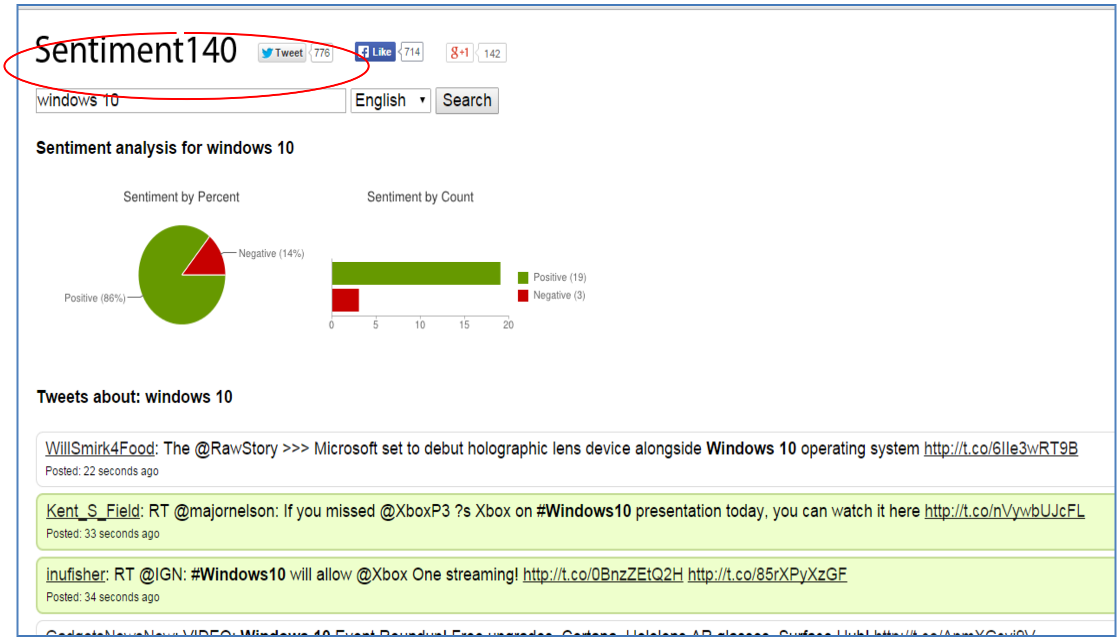
Date	User	v	a	Tweet
01-22-15 07:49	writeblogs	6.78	5.	*free* for *lifetime* of device after initial *upgrade*, no yearly fee for upgraders http://t.c
01-22-15 07:50	mani4astro	6.11	5.	FEB 2015 . http://t.co/hTdCa1hfJ2 #Windows10 #*Windows* #Microsoft
01-22-15 07:50	BakaDW	6.68	4.7	*Windows* 7 and 8.1 users :D
01-22-15 07:50	samymohab	5.53	4.70	Microsoft *shows* off *Windows* 10, HoloLens, Surface Hub At Microsoft headquarters, *executives show* off *Windows* 10... http://t
01-22-15 07:50	CeleciaKat	6.30	5.60	RT @protocollie: Okay *thank fucking god* Microsoft didn't *screw* this *one* up. Had too little *faith*. http://t.co/LQiVFFl1Jz
01-22-15 07:50	DeliTori	7.12	4.44	RT @PatrickLear95: *FREE WINDOWS* 10 FOR YOU, AND FOR YOU, AND FOR ME!
01-22-15 07:50	CapitanoCasella	6.03	4.73	I *liked* a @YouTube video http://t.co/DWkSXyvCxw Microsoft Outlines *Details* of *Windows* 10's Xbox *One* App - IGN *News*
01-22-15 07:50	WoodPunchaz	6.54	4.60	I *liked* a @YouTube video http://t.co/h2PbguSYuQ *Windows* 10: A *New* Generation of *Windows*
01-22-15 07:50	_Julian95	6.23	4.94	RT @*Windows*: Terry Myerson talks about all things #Windows10 in his recap of the *day*'s *events*. http://t.co/qm2dmatQxq http
01-22-15 07:50	srivi20	6.17	4.41	RT @tomwarren: This is the *new Windows* 10 *Start Screen* http://t.co/bb742mBmZB http://t.co/vBqOT38qAF
01-22-15 07:50	TRISTANvsTEE	6.00	5.00	TRISTAN: @Microsoft Keynote Everything you need to *know* about *Windows* 10. http://t.co/DTAIqadLSW #windows10
01-22-15 07:50	LloydMildon	6.50	5.03	*Hilarious moment* at *Windows* 10 newser. A *sea* of Macs"@AustenAllred: A *couple* MacBooks at the *Windows* 10 Unveiling... htt
01-22-15 07:50	Peterogena	6.32	4.34	*Start* menu resurrected with *Windows* 10, but with *Live* tiles http://t.co/3rzC0vA1jW via @gsmarena_com
01-22-15 07:50	leonheart13	5.63	5.58	RT @pcgamer: Microsoft: *Windows* 10 will not be *sold* as a subscription http://t.co/tgg1cvJVoo http://t.co/uq9kDwwwM6
01-22-15 07:50	tsriram			ar:It's a *lot* nicer to say "*Windows* 10 will be *free* for 1 year" than "*Windows* 10 will have a 1
01-22-15 07:50	___News___			oft Introduces HoloLens 3-D Headset: The *company* unveiled a headset that allo... http://t.
01-22-15 07:50	taoxv88			/hYOW4QVfBJ Xbox on *Windows* 10
01-22-15 07:50	Sergio_TheOne			*excited* about *Windows* 10 for phones. But HoloLens... It's out of this *world*!

Showing 1 to 438 of 438 entries

Keywords: Windows 10 Query

Windows 10 as a keyword to search tweets

Figure 10. Sentiment Viz



windows 10

Sentiment analysis for windows 10

Tweets about: windows 10

WillSmirk4Food: The @RawStory >>> Microsoft set to debut holographic lens device alongside **Windows 10** operating system http://t.co/6lIe3wRT9B
Posted: 22 seconds ago

Kent_S_Field: RT @majornelson: If you missed @XboxP3 ?s Xbox on **#Windows10** presentation today, you can watch it here http://t.co/nVywbUJcFL
Posted: 33 seconds ago

inufisher: RT @IGN: **#Windows10** will allow @Xbox One streaming! http://t.co/0BnzZEtQ2H http://t.co/85rXPyXzGF
Posted: 34 seconds ago

Figure 11. Sentiment 140

6. Inference and future work

To perform opinion mining we have chosen Twitter, a popular micro blogging where millions of users share their opinions. In our research we have defined the programming method to extract twitter data using R packages and search API. We have also discussed the various non-programming methods of tweets extraction. We use "Windows 10" as a keyword to extract tweets. The same customer feedback (opinion) on the product viz. windows 10 would benefit the company manufacturing the product; improve its product quality, influencing its marketing. Hence, Opinion Mining aid in customer feedback phase during beta releases of the products. Our experimentation establishes the fact that programming technique proves better than non-programming technique for opinionated tweet mining.

In future, the methodology for proceeding in direction will be to use the extracted tweets via R packages to build corpus. The collected corpus will then be used to perform sentiment analysis using classifiers (Naïve Bayes and Support Vector Machine Classifier). Our classifier will be able to determine the tweets as positive, negative or neutral, thereby, helping the end user to form a decisive opinion on the search query.

References

[1] P. Alexander and P. Patrick, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining", Proceedings of LREC, 2010.

[2] G. Vinodhini and R. M. Chandrasekaran, "Sentiment Analysis and Opinion Mining: A Survey" International Journal of Advanced Research in Computer Science and Software Engineering, vol. 2, no. 6, pp. 282-292, Jun. 2012.

[3] L. Pooja, A. Sachdeva, D. Mahajan, N. Pande and P. Kumar, "An approach towards comprehensive sentimental data analysis and opinion mining" in Proc. Advance Computing Conference, Gurgaon, Feb. 2014.

[4] S. Grigori, et al., "Empirical Study of Machine Learning Based Approach for Opinion Mining in Tweets", Lecture Notes in Computer Science, vol. 7629, pp. 1-14, 2013.

[5] B. Arti, M.B. Chandak and A. Zadgaonkar, "Opinion Mining and Analysis: A Survey" International Journal on Natural Language Computing (IJNLC), vol. 2, no. 3, pp. 39- 48, Jun. 2013.

[6] G. Kathrin and C. I. Chesnevar, "A First Approach Towards Integrating Twitter and Defeasible Argumentation", in Proc. 13th Symposium on Artificial Intelligence (ASAI), Argentina, 2012.

[7] K. Grossel, C. Chesneva, A. Maguitman and E. Estevez, "Empowering an E-Government Platform Through Twitter-Based Arguments", Inteligencia Artificial, vol. 15, no. 50, pp. 46-56, 2012.

[8] K, Akshi and T. M. Sebastian, "Sentiment Analysis on Twitter", International Journal of Computer Science (IJCSI), vol. 9, no. 4(3), Jul. 2012.