Analysis of Twitter Data for Understanding Tweeting Activity of Public

Pradeep Kumar G[1]*, Ezhilarasan C[1], and Sandhiya R[1].

[1] School of Computing Science and Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India.

ABSTRACT

The penetration and use of social media services differs from city to city. Here in this paper we analyses the use of Twitter among three major corporate cities in India. We develop a geographical and sentimental analysis of Tweeting activity, and their names, genders and probable ethnicities of Twitter users in these cities. The results will reveal whether the majority of Twitter users are either male or female and will also provide an insight into the ethnic diversity of population and their sentimental analysis in these cities.
Keywords: Twitter, Tweeting activity, Geographical and Sentimental analysis, Ethnic diversity.

*Corresponding author

INTRODUCTION

Social Medias are the computer-aided tools that relay mainly on web-based technologies in order to create more interactive platforms by which the individuals and communities can co-create, share, discuss, and modify the content generated by the user. Every day large amount of data are found to be produced by social media services and it is clearly stated that these data are widely used in commercial sectors and research areas for analyzing the data. Users of Twitter, Flickr, You Tube, Facebook, Orkut and LinkedIn are frequent users of smart phones, iPad and tablet devices, these users may belong to different ethnic and social backgrounds.

Here we mainly focus on Twitter network, which is a well known social media networking site. Twitter is an online social networking service that allows the users to send and read the short messages that are called as "tweets". Twitter was created in March 2006 by Jack Dorsey, Evan Williams, Biz Stone and Noah Glass and introduced during July 2006. Its service increasingly gained the popularity worldwide, which exceeds 100 million users who in 2012 for per day posted 340 million tweets. As of May 2015, it is found that Twitter users are more than 500 million, out of which 302 million are active users. Tweets are publicly available by default, but the senders can limit their message delivery only to their followers. All the users can tweet through the Twitter website, or by Short Message Service that is available in certain countries. Twitter data is a wide data source, where analysis of these data can result in useful information about the locations and environment in which users indulge them in this social media service.

So many different analyses have been carried out using Twitter data, such as dissemination of data across social networks, monitoring community happiness from Tweets, the investigation of behavior of the user, and geographic analysis of the Twitter network. So here we have chosen a different way to analyze the name, age, gender, ethnicity and sentiment of the twitter users. Previous research works have been stated that the name of a person can predict his/her linguistic characteristics, ethnicity and culture. Name and ethnicity of a person can clearly frame the geographical structure of distinct ethnic groups with respect to their particular media services. Gender analysis is made to find which gender uses twitter in more whether male or female. Sentimental analysis is performed on the basis of the tweeting activity of the users. Thus this paper particularly aims at finding the name, age, gender, sentiment and ethnicity of the twitter users among three major corporate cities in India.

Big Data

Massive information investigation alludes to the procedure of sorting, gathering and examining the huge arrangements of information (called huge information) in order to predict the designs and other useful data. Huge information assessment can guide the associations for better understanding of the data that is found inside the information and in the same way it identifies the information that is most important for the business and also helpful in future business planning. Vast information investigators basically require the learning that grows from cracking down the information.

The Challenges of Big Data Analytics

For various combinations, vast information examination is a perfect test. Considering the absolute amount of information and the unique alignment of the information (both unstructured information and organized) that are collected over the whole combination and a wide variety of paths diverse order of information can be combined, differentiated and investigated in order to determine the valuable business data and examples. The major test is in separation of information storehouses in order to get to all information; the associations are stored in better places and regularly in unique designs. A second vast information test is in framing stages that can drag in unstructured information as efficiently as well organized information. This huge volume of information is in general so broad that it's tuff to process programming techniques and utilizing conventional database.

Enormous Data Requires High-Performance Analytics

To divide such a wide amount of information, enormous information assessment is frequently carried out using a particular type of programming devices and applications for perceptive examination, content mining, information mining [1], information enhancement and gauging. These procedures are exact yet

remarkably incorporated essentials of elite investigation. Using programming and wide variety of information apparatuses enhances an association to handle an immense degree massive amount of information that are gathered by a business to frame out which information is appropriate and can be dissected to hold improved business decisions later on.

Benefits of Big Data Analytics

Endeavors are gradually eager in finding important bits of knowledge in their information. Several wide information undertakings start from the requirement to respond certain business queries. With this opportunity several information examination stages an endeavor can improve operations, client administration, guide deals, widen proficiency, and risky administration.

Webopedia guardian institute, QuinStreet, overviewed 540 business enterprise chiefs included in huge information receives to decide which business institute sketches to make use of Big Data examination to improve operations. About the portion of all respondents said that they are applying large information assessment to improve the client maintenance, guide with item enhancement and raise an upper hand. Extremely, the business area receives more consideration similarities with increasing ability and improving operations. In exact, 62 percent of respondents stated that they have used wide information investigation to improve the velocity and reduce multifaceted character.

MongoDB is an archive database that provides high accessibility, elite and simple versatility.

- Archive Database
- Archives (items) delineate to programming dialect data sorts.
- Entrenched account and displays reduced need for joins.
- Polymorphism is made simple by dynamic diagram.
- Elite
- Inserting makes composes and peruses quicker.
- Lists can integrate keys from implanted exhibits and archives.
- Discretionary spilling composes (no affirmations).
- Increased Availability
- Servers that are reproduced with programmed expert failover.
- Linear Scalability
- Programmed sharing circulates collecting information crosswise over machines.
- At the end consistent peruses can appropriated on duplicated servers.
- Drived Operations

With MongoDB Cloud Manager MongoDB underpins an entire reinforcement arrangement and complete organization monitoring.

MongoDB Data Model

A MongoDB arrangement has variety of databases. A database that withholds sorted accumulations. An accumulation withholds sorted archives. An archive is a sorting of key-worth sets. Archives is said to have dynamic blueprint. Dynamic mapping [1] involves that archives in the similar accumulation need not have the similar sorting of structure, and usual fields in a gathering's records may withhold different sorts of information.

RELATED WORK

Twitter.com is a well known smaller scale blogging site; every tweet is 140 characters long. Tweets are habitually used to express a tweeter's feeling on a specific subject. There are firms which survey twitter for investigating feeling on a specific subject. The test is to assemble all such important information, identify and abridge the general conclusion on a theme. Twitter is the essential needs in normal corporate life, a percentage of the examinations have been made as of now utilizing the twitter information, its accumulation of the "mongodb "utilized just for twitter. Utilizing BigData It can actualize for an assortment of groups, the slither tweets delivered by their occupants and get evaluation information of their financial prosperity (Section "Dataset"). All the more unequivocally, slither Twitter accounts whose client determined areas report London

neighborhoods. Geo-referencing of these records is made by changing their area into longitude scope sets. Likewise get the 2007 UK government's Index of Multiple Deprivation (IMD), which is a composite measure of relative successes of 32,482 groups, 78 of which are in London. Examine the feeling of tweets utilizing two calculations (Section "Supposition Analysis"). The primary is the word tally system proposed by Kramer and depends on processing the institutionalized distinction between percent of words that are certain and those that are negative. The second is the Maximum Entropy classifier, which is a condition of the art strategy for grouping estimation of single tweets. Once execute both classifiers and demonstrate that their outcomes relate and are precise about the relationship in the middle of opinion and socio economic well-being (Section "Results").

TWISENT: A multistage System for Analyzing Sentiment in Twitter in Feb 2013 [2].they have exhibited Twisent, an assumption examination framework for twitter. Taking into account the point looked, Twisent gathers positive, negative and objective. Nonetheless, breaking down smaller scale blog entries have numerous characteristic difficulties contrasted with the other content kinds. What's Great and so forth: Learning to characterize the Scope of Negation for Improved Sentiment Analysis in July 2010 [3].Presents a refutation identification framework taking into account a restrictive arbitrary field displayed utilizing components from an English reliance parser, the extent of nullification discovery is constrained to express as opposed to inferred invalidations inside of single sentences. Twitter informal organization in light of clients' geo locations and researched how clients' geo locations sway their cooperation in Twitter, their network with different clients, and the data they trade with tweets delivered in Twitter. In this way, monetary lopsided characteristics in the disconnected from the net world are misrepresented in the online world. The information has been taken from twitter and it made for examination, by taking his profile name and made a few tweets whatever he made in his record. The figured assumption scores crosswise over diverse groups to make two evaluations. Initially, we expected to pick up an understanding into the geographic circulation of opinion crosswise over London neighborhoods by overlaying "Gross Community Happiness" (GCH) scores onto a city guide utilizing the Google Maps Javascript API [4].

The guide freely shows that the North West is the happiest quadrant of the city, whilst South East London is genuinely flooded with blue, demonstrating miserable estimation. In reality, ordinary intelligence among Londoners holds that "By and large London has a North West/South East slope of rich to unpleasant territories. It's less 'West end' versus 'East end' as 'North West' versus 'South East'. Large portions of the truly dodgy parts of London are in the South East". Research is carried out on miniaturized scale blogging, examination on client profiling utilizing web pursuit and desktop conduct, research on utilizing online networking for client data, and exploration on applying client profiles for personalization and grouping of social hunt. Twitter is a well known micro blogging administration, and despite the fact that it is genuinely new, it has been concentrated on by a significant number of specialists as a result of a certain kind qualities of short messages, unidirectional connections among clients, and successive real time redesigns through versatile web [5] was one to start with, and it exhibits their initial discoveries about general use of tweets and the informal organization on Twitter. Later it is also presented that an examination of tweets as per diverse classifications, and in light of those classes, order of the Twitter clients into meformers and sources [6]. The exploration on the Twitter patterns is carried out to recover aggregate knowledge utilizing tweets and the demographics of the creators of those tweets [7]. Other than these, there is a developing writing on Twitter studies, and an all around overhauled rundown is at danah boyd's website3. There is great former examination on construing about clients by following their web and desktop practices. Considering the gander at a client's Navigate is conducted to construe what it is occupied with, Word Count [8]. Utilize a lexicon called "Semantic Inquiry Word Count". LIWC is a standard lexicon of 2,300 English words that catch 80% of the words utilized as a part of ordinary discussions and mirror individuals' passionate and intellectual discernments. In the wake of expelling prevent words from tweets, check, for every profile, the quantity of words that are certain and those that are negative (words coordinating the two classifications of 'positive feelings' and 'negative feelings' as characterized in LIWC) and total both checks to create the "Word Count Sentiment" score, which is like the score proposed by Kramer [9].

Overview of Twitter Data Analysis

The first step is to create a Twitter application for you. Go to https: //twitter.com/apps/new and log in. After filling in the basic info, go to the "Settings" tab and select" Read, Write and Access direct messages". Make sure to click on the save button after doing this. In the "Details" tab, take note of your consumer secret

and consumer key. In your R session, you have to do the followings with appropriate values from the web page:

```
> setup_twitter_oauth ("API key", "API secret")
```

This will authenticate via http, I recommend looking at that package's Token man page for more information in order to manage the caching processes and authentication. If you are in a headless environment or otherwise if you don't want to deal with the browser based authentication dance, then you can get your secret and access token from your apps webpage and call setup_twitter_oauth a bit differently:

```
> setup_twitter_oauth ("API key", "API secret", "Access token", "Access secret")
```

For more in depth examples I recommend exploring the mailing list or Stack Overflow. I've also included some links to examples of twitteR being used in the real word at the end.

```
> library(twitteR)
> setup_twitter_oauth("API key", "API secret")
```

Searching Twitter

The search Twitter function is used in searching the tweets that matches a preferred term. Example searches towards such things like hash tags, basic Boolean logic such as AND and OR. It is worthy to look at https://dev.twitter.com/docs/ using-search for an example of what can and cannot be performed here. The n argument can be used to define a variety of tweets to return, defaulting to 25. Note that while searchTwitter will enfold an arbitrary number of real search calls to give the number of tweets that are requested, the Twitter API has some limitations on how much it will return in actual and it is clear that you can only go back a handful of days worth of tweets. 2

```
> tweets = searchTwitter('#rstats', n=50)
> head(tweets)
```

If the Twitter user runs the command getUser, it will only work correctly with the users who have their profiles in public, or if you're authenticated with access. You can also see things like user's followers, who they follow, retweets and more. The getUser function can also return a user object, which is then be polled for further information.

```
> crantastic = getUser('crantastic')
> crantastic$getDescription()
> crantastic$getFollowersCount()
> crantastic$getFriends(n=5)
```

METHODOLOGY

- **Twitter data**: Twitter Streaming API is used to download geo-tagged Tweets.
- **Gender Analysis:** Fore name of a person is used to perform gender analysis.

Expected outcome: Presenting a sentimental analysis 10, specifically according to their location, name, ethnicity, gender and age of the Twitter users over three major corporate cities in India.

Conversion to data.frames

There are times when it is convenient to display the object lists as a data.frame structure. To do this, every class is found to have a reference method to DataFrame as well as a corresponding S4 method as.data.frame that is said to work in the traditional sense. Transforming a single object will typically not be particularly useful by itself but there is an easy method to transform an entire list, twListToDF which considers a list of objects from a single twitteR class:

```
> df = twListToDF(tweets)
```

> head(df)

Database Persistence

A question that is often asked is how to recover data from the past, generally people are performing analysis on some major event that have already happened (e.g. Arab Spring, an election, etc.). It is impossible using the Twitter API because you can only track back a little amount. However, if you have the ability to view forward, it is easy to allow a prospective study by gathering data and automatically persist it to the database. This will then permits you to load everything into a later R session that includes using tools such as dplyr. A full writeup of this functionality is given at:

http://geoffjentry.blogspot.com/2014/ 02/twitter-now-supports-database.html.

Here's a brief example for it:
> sql_lite_file = tempfile()
> register_sqlite_backend(sql_lite_file)
> store_tweets_db(tweets)
> from_db = load_tweets_db()
> head(from_db)

EXPERIMENT AND RESULT ANALYSIS

Data set:

>setup_twitter_oauth(consumer_key=   "z5Z4m0yDgElVOyvTQm5NXPrag",consumer_secret=
"21ReeOR479LvbBw1F7QEooV1sig8VGZNlqEIb8iQWhnCcWWNXC",access_token   =   "2363672953-
NQfWuAhBbJqWoDRuxFQVOZgZrMSJoAd7JPrJJNy",access_secret   =
"9OYJKZ9bPeLmpkW4AS0wPDGV3lKaDS9566caMdpV2m1kb")
>tweets<- searchTwitter('nadigarsangam', lang ='en')
>tweets
[[1]]
[1] "arvind_mohanraj: Not bothered about nadigarsangam but the way in which @realsarathkumar handled the situation is what you expect of a leader @realradikaa"

[[2]]
[1] "UNKHTrends: Retweeted ऋद्धिमा (@BeingRidhima):\n\nHe spoke only fr welfare of Tamil Cine & NadigarSangam without paying heed... https://t.co/4Klbj420Mc"

[[3]]
[1] "UNKHTrends: Retweeted ऋद्धिमा (@BeingRidhima):\n\nHe spoke only fr welfare of Tamil Cine & NadigarSangam without paying heed... https://t.co/isbnhI3iob"

[[4]]
[1] "UNKHOurPride: RT @BeingRidhima: He spoke only fr welfare of Tamil Cine & NadigarSangam without paying heed to comments against him - #KamalHaasan..Resp…"

[[5]]
[1] "nakkheeranweb: SarathKumar Breaks Down in Tears - NadigarSangam Issue Continues... https://t.co/CRMV0PT0oT"

[[6]]
[1] "edgei_tech: LIVE - NadigarSangam Election Results 2015 – In India\n\nhttps://t.co/rNQt0QGZJ6\n\nMedia ads by Edgei https://t.co/uhg0AigXnR"

[[7]]

[1] "Sree926: This NadigarSangam elections were so similar to our MAA elections, but with a better script and enough masala."

[[8]]
[1] "JeyenthV: Dei, please have nadigarsangamreelection da. So, boring. :D\n#NadigarSangamElection"

[[9]]
[1] "Raju5rnr: I liked a @YouTube video from @galattadotcom https://t.co/UFsQ6Jv5pE Let's pay way for the youngsters for this NadigarSangam"

[[10]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to Junior Balaiah - NadigarSangam Committee Member https://t.co/f3K8mJgM01"

[[11]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to KuttyPadmini - NadigarSangam Committee Member https://t.co/PyFkaIm8dB"

[[12]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to Nandha - NadigarSangam Committee Member https://t.co/y9LF2R3d6N"

[[13]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to @Prasanna_actor - NadigarSangam Committee Member https://t.co/gC5row6SK8"

[[14]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to Rajesh - NadigarSangam Committee Member https://t.co/nNb6BYz8BA"

[[15]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to Ramana - NadigarSangam Committee Member https://t.co/PZ0xnqzmO4"

[[16]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to Sriman - NadigarSangam Committee Member https://t.co/pBOvyDIDOf"

[[17]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to @sangithakrish - NadigarSangam Committee Member https://t.co/Cvqtb04jCs"

[[18]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to Thapathy Dinesh - NadigarSangam Committee Member https://t.co/X0uQrQmhRU"

[[19]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to @ACTOR_UDHAYAA - NadigarSangam Committee Member https://t.co/BquLsqANiL"

[[20]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to Vignesh - NadigarSangam Committee Member https://t.co/yJ6NGkWMri"

[[21]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to Pasupathy - NadigarSangam Committee Member https://t.co/1JrZ44RDwG"

[[22]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to Nirosha- NadigarSangam Committee Member https://t.co/VP8s8gAIWF"

[[23]]
[1] "vith286: RT @onlynikil: Hearty Congratulations to Nalini - NadigarSangam Committee Member https://t.co/31ifL4cDEX"

[[24]]
[1] "ramvikh: @itisprashanthi really like the way you review the films, gudwrk. Surprised that no video about nadigarsangamelection , your take on it."

[[25]]
[1] "devnullHun: RT @VisvyaSolutions: #Retweet Naser Win InNadigarSangam Election | Tamil Cinema News | Updates https://t.co/owxFK22SCn #India #News #… ht…"

>tweets<- twListToDF(tweets)
>tweets
text
1   Not bothered about nadigarsangam but the way in which @realsarathkumar handled the situation is what you expect of a leader @realradikaa
2 Retweeted<U+090B><U+0926><U+094D><U+0927><U+093F><U+092E><U+093E> (@BeingRidhima):\n\nHe spoke only fr welfare of Tamil Cine & NadigarSangam without paying heed... https://t.co/4Klbj420Mc
3   Retweeted<U+090B><U+0926><U+094D><U+0927><U+093F><U+092E><U+093E> (@BeingRidhima):\n\nHe spoke only fr welfare of Tamil Cine & NadigarSangam without paying heed... https://t.co/isbnhI3iob
4   RT @BeingRidhima: He spoke only fr welfare of Tamil Cine & NadigarSangam without paying heed to comments against him - #KamalHaasan..Resp…
5   SarathKumar Breaks Down in Tears - NadigarSangam Issue Continues... https://t.co/CRMV0PT0oT
6   LIVE - NadigarSangam Election Results 2015 – In India\n\nhttps://t.co/rNQt0QGZJ6\n\nMedia ads by Edgei https://t.co/uhg0AigXnR
7   This NadigarSangam elections were so similar to our MAA elections, but with a better script and enough masala.
8   Dei, please have nadigarsangamreelection da. So, boring. :D\n#NadigarSangamElection
9   I liked a @YouTube video from @galattadotcom https://t.co/UFsQ6Jv5pE Let's pay way for the youngsters for this NadigarSangam
10   RT @onlynikil: Hearty Congratulations to Junior Balaiah - NadigarSangam Committee Member https://t.co/f3K8mJgM01
11   RT @onlynikil: Hearty Congratulations to KuttyPadmini - NadigarSangam Committee Member https://t.co/PyFkaIm8dB
12   RT @onlynikil: Hearty Congratulations to Nandha - NadigarSangam Committee Member https://t.co/y9LF2R3d6N
13   RT @onlynikil: Hearty Congratulations to @Prasanna_actor - NadigarSangam Committee Member https://t.co/gC5row6SK8
14   RT @onlynikil: Hearty Congratulations to Rajesh - NadigarSangam Committee Member https://t.co/nNb6BYz8BA
15   RT @onlynikil: Hearty Congratulations to Ramana - NadigarSangam Committee Member https://t.co/PZ0xnqzmO4
16   RT @onlynikil: Hearty Congratulations to Sriman - NadigarSangam Committee Member https://t.co/pBOvyDIDOf
17   RT @onlynikil: Hearty Congratulations to @sangithakrish - NadigarSangam Committee Member https://t.co/Cvqtb04jCs
18   RT @onlynikil: Hearty Congratulations to Thapathy Dinesh - NadigarSangam Committee Member https://t.co/X0uQrQmhRU

19   RT @onlynikil: Hearty Congratulations to @ACTOR_UDHAYAA - NadigarSangam Committee Member https://t.co/BquLsqANiL

20   RT @onlynikil: Hearty Congratulations to Vignesh - NadigarSangam Committee Member https://t.co/yJ6NGkWMri

21   RT @onlynikil: Hearty Congratulations to Pasupathy - NadigarSangam Committee Member https://t.co/1JrZ44RDwG

22   RT @onlynikil: Hearty Congratulations to Nirosha- NadigarSangam Committee Member https://t.co/VP8s8gAIWF

23   RT @onlynikil: Hearty Congratulations to Nalini - NadigarSangam Committee Member https://t.co/31ifL4cDEX

24 @itisprashanthi really like the way you review the films, gudwrk. Surprised that no video about nadigarsangamelection , your take on it.

25 RT @VisvyaSolutions: #Retweet Naser Win In NadigarSangam Election | Tamil Cinema News | Updates https://t.co/owxFK22SCn #India #News #… ht…

statusSource

1 Twitter for iPhone
2 Facebook
3 Facebook
4 Twitter for Android
5 Facebook
6 Facebook
7 Twitter for iPhone
8 Twitter for Android
9 Google
10 Twitter for Android
11 Twitter for Android
12 Twitter for Android
13 Twitter for Android
14 Twitter for Android
15 Twitter for Android
16 Twitter for Android
17 Twitter for Android
18 Twitter for Android
19 Twitter for Android
20 Twitter for Android
21 Twitter for Android
22 Twitter for Android
23 Twitter for Android
24 Twitter Web Client
25 winterpostgenerator

screenNameretweetCountisRetweet retweeted longitude latitude

	screenName	retweetCount	isRetweet	retweeted	longitude	latitude
1	arvind_mohanraj	0	FALSE	FALSE	NA	NA
2	UNKHTrends	0	FALSE	FALSE	NA	NA
3	UNKHTrends	0	FALSE	FALSE	NA	NA
4	UNKHOurPride	25	TRUE	FALSE	NA	NA
5	nakkheeranweb	0	FALSE	FALSE	NA	NA
6	edgei_tech	0	FALSE	FALSE	NA	NA
7	Sree926	0	FALSE	FALSE	NA	NA
8	JeyenthV	0	FALSE	FALSE	NA	NA
9	Raju5rnr	0	FALSE	FALSE	NA	NA
10	vith286	3	TRUE	FALSE	NA	NA
11	vith286	3	TRUE	FALSE	NA	NA

12	vith286	3	TRUE	FALSE	NA	NA
13	vith286	7	TRUE	FALSE	NA	NA
14	vith286	4	TRUE	FALSE	NA	NA
15	vith286	4	TRUE	FALSE	NA	NA
16	vith286	4	TRUE	FALSE	NA	NA
17	vith286	5	TRUE	FALSE	NA	NA
18	vith286	4	TRUE	FALSE	NA	NA
19	vith286	4	TRUE	FALSE	NA	NA
20	vith286	3	TRUE	FALSE	NA	NA
21	vith286	3	TRUE	FALSE	NA	NA
22	vith286	2	TRUE	FALSE	NA	NA
23	vith286	2	TRUE	FALSE	NA	NA
24	ramvikh	0	FALSE	FALSE	NA	NA
25	devnullHun	2	TRUE	FALSE	NA	NA

Graph

```
>barplot.default(grep('data',tweets,ignore.case = TRUE))
```
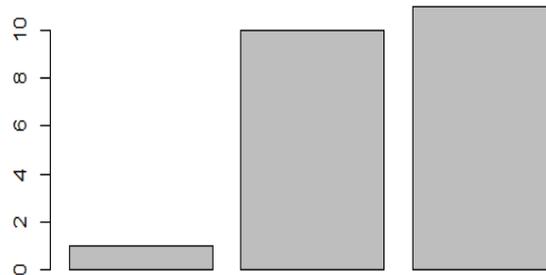

CONCLUSION

In this paper, twitter data for "Then india Nadigar sangam" –Election which is held on OCT 19th 2015 has been classified and the expected outcome has achieved using the tweets. Here, we endeavored to address the question: does disconnected from the net geology still matters in online informal organizations? To end this, we analyzed the Twitter informal organization taking into account clients' geo locations and researched how clients' geo locations bend their hold in Twitter, their network with various clients, and the data they deal with them.

Our inside and outside assessments clearly states that that, geology seriously alters the all parts of Twitter interpersonal organization. In particular, we find that in spite of the fact that clients specially border what's more, trade data with different clients from their respective nation, more than 33% of all connections and tweets are found traded crosswise exceeding national boundaries. Such transnational associations and co operations happen among clients in topographically what's a more, semantically proximal nation. Our findings is said to have possible applications in foreseeing or signifying social associations for the client and also in understanding worldwide distribution of data.

REFERENCES

[1]   Alexander Pak, Patrick Paroubek, LREC 2010.
[2]   Bahrainian S A, Dengel A, IEEE 2013; pp. 26- 29.
[3]   Alec Go, Lei Huang, Richa Bhayani, CS224N Stanford 2009; pp. 1-12.
[4]   Dean J and Ghemawat S, Communications of the ACM 2008; 1: 1107-113.

[5]   Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, Byers A H, McKinsey & Company 2011.
[6]   Shvachko K, Kuang H, Radia S, and Chansler R, IEEE 2010; pp. 1-10.
[7]   Ghemawat S, Gobioff H, Leung S T, IEEE 2003; 5: 29-43.
[8]   Ambler S W, Relational databases 101: Looking at the whole picture, AgileData 2009.
[9]   Shafter J, Rixner S, Cox A L, IEEE 2010; pp. 122-133.
[10]   Tang H, Tan S, Cheng X, Elsevier 2009; 7: 10760-10773.