

SOCIAL NETWORK ANALYSIS AND SENTIMENT ANALYSIS

A PROJECT REPORT

Submitted by

YADAVALLI HARI
VAMSI(17BCE0525)
ABBIREDDY B D V V
GANGARAJU-
(17BCE0039)

**CSE3020
Data Visualization**

Under the guidance of

Dr. Meenakshi S P
Associate Professor, SCOPE
VIT, Vellore.



VIT[®]

Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

**SCHOOL OF COMPUTER SCIENCE AND
ENGINEERING**

NOVEMBER, 2019

ACKNOWLEDGEMENT

We sincerely thank our Chancellor – Dr. G. Viswanathan, VIT University, for giving us the opportunity to pursue this course of Data Visualization and introduce J-component as part of our academic curriculum. We also thank Prof. Meenakshi S P – School of Computer Science and Engineering, VIT University, for giving us the opportunity to do this project. We also thank the Dean and HOD of School of Computer Science and Engineering (SCOPE), for their continued support and encouragement.

PROBLEM STATEMENT

Extraction of online tweets and update the datasets collected dynamically, pre-process (removal of URLs, type casting and removal of stop words) the tweets, and classify the tweets based on 10 parameters

INTRODUCTION

Motivation: Compared to documents text and tweets are smaller. English used will be informal and creative words, punctuations, wrong spellings, different slangs, and new words. To tweets have high information structure about the people who communicate.

Significance: With this anyone can retrieve the data from twitter and dynamically store in database. Also, can be easily visualized using different software. With these visualizations we can analyze the emotions of the tweets (Positive, Negative or Neutral).

Scope and Applications: Sentiment Analysis obtains the sentiment of document or paragraph or sentence or expressions (with in sentence) which is like opinion mining, process of identifying a piece of text according to the tone conveyed by it. Can be utilized by business companies or organizations for enhancing customer experience based on the comments or tweets, this becomes a competitive advantage to the company or organization.

LITERATURE SURVEY

S. NO	TITLE	ABSTRACT	YEAR	AUTHOR/S
1	Large-scale sentiment analysis for news and blogs	Newspapers and blogs express opinions about news organizations (people, locations, things) as they report on recent events. Our system consists of a phase of sentiment recognition, associating opinions shared with each specific entity, and a phase of sentiment aggregation and ranking, scoring each entity in the same category compared to others. Lastly, we determine the importance of our scoring techniques over broad news and blogs corpus.	2007	Namrata Godbole, Manjunath Srinivasaiah, Steven Skiena
2	International sentiment analysis for	Mining opinions are becoming increasingly interested in using methods of sentiment analysis from outlets such as news, forums and product reviews. For English, most of these approaches have been developed and are	2008	Mikhail Bautin, Lohit Vijayarenu, Steven Skiena

	new and blogs	hard to generalize to other languages. We are developing an approach using state-of-the-art machine translation software and performing sentiment analysis on a foreign language text translation in English. Our results indicate that (a) person sentiment scores obtained by our system were statistically significantly associated across nine news source languages and five parallel corpus languages; (b) the performance of our sentiment analysis tool is largely independent of the translator; (C) our agency sentiment scores can be used to make meaningful crosscultural correlations after applying such standardization techniques.		
3	Context-Enhanced Citation Sentiment Detection	Review of nostalgic references in scientific papers and articles is a new and interesting topic that can open up some exciting new applications of bibliographic search and bibliometry. Current work on the analysis of citation sentiment focuses only on the sentence of citation. We address the issue of context-enhanced identification of sentimental citation in this paper. We present a new corpus of citations sentiment that has been annotated to take into account the prevailing feeling in the whole sense of citation. We assume this gold standard is closer to the truth than annotation that only looks at the quotation term itself. When using this context-enhanced gold standard description, we then explore the impact of context windows of different lengths on the quality of a state-of-the-art citation sentiment detection system	2012	Awais Athar, Simone Teufel
4	A Review of Feature Extraction in Sentiment Analysis	Rapid growth of internet users along with increasing power of online review sites and social media has given birth to Sentiment Analysis and Opinion Mining, which helps to decide what other people think and comment on. Sentiments or Opinions include content of goods, services, policies and policies produced by the public. People are usually involved in looking for positive and negative feedback of likes and dislikes posted by users for particular product or service features. Product features or attributes also play an important role in the study of feelings. In addition to enough work being done in text analysis, feature extraction is now becoming an active area of research in sentiment analysis. In sentiment analysis and opinion mining, this review paper addresses current methods and strategies for extraction of data. In this study, we followed a systematic literature review process to identify areas that are well-focused by researchers, as well as highlighting less-targeted areas that provide researchers with an incentive for further work. We also attempted to classify most and least widely used strategies for selecting features to define study gaps for future work.	2014	Muhammad Zubair Asghar, Aurangzeb Khan, Shakeel Ahmad, Fazal Masud Kundi

5	Sentiment analysis using product review data	Study of emotion or opinion mining is one of NLP's (Natural Language Processing) main tasks. Study of opinion has gained a great deal of popularity in recent years. In this article, we intend to address the issue of categorizing sentiment polarity, which is one of the fundamental issues of evaluating sentiment. A general method is proposed for categorizing sentiment polarity with detailed description of the system. Online product reviews obtained from Amazon.com are the information used in this report. Experiments were performed with promising results for both sentence-level categorization and review-level categorization. Ultimately, we also provide insight into our future work on the study of sentiments.	2015	Xing Fang, Justin zhan
6	Citation sentiment analysis in clinical trial papers	Positive references and critical reviews can often be seen in scientific writing in the text that discusses the cited articles, providing useful data on whether or not a thesis can be repeated. In this review, they concentrate on the analysis of citation sentiment, which seeks to evaluate the polarity of sentiment that the meaning of citation contributes to the cited paper. For clinical trial papers, a citation sentiment corpus was first annotated. Using the annotated corpus, the efficacy of n-gram and sentiment lexicon features and problem-specific structure features were then tested for citation sentiment analysis. The combined features of the term n-grams, sentiment lexicons, and structure data achieved the highest Micro F-score of 0.860 and Macro-F score of 0.719, suggesting that in biomedical publications it is feasible to use machine learning methods to evaluate citation sentiment. A detailed comparison was performed between citation sentiment analysis of clinical trial papers and other general domains, further illustrating the unique challenges in this field.	2015	Jun Xu, Yaoyun Zhang, Yonghui Wu, Jingqi Wang, Xiao Dong, Hua Xu
7	Analysis of Various Sentiment Classification Techniques	Analysis of opinion is a continuous area of research in the field of text mining. People post their analysis in the form of unstructured data so that opinion extraction offers overall review opinion so that it does the best job for consumers, individuals, organisations, etc. The main purpose of this paper is to define strategies with good accuracy to generate performance. This paper provides recent updates on articles on the identification of sentiment analysis for various approaches and algorithms that have been introduced. The main contribution of this paper is to give an idea that careful selection of features and current approaches to classification can boost accuracy	2016	Vimal Kumar B. Vaghela, Bhumika M. Jadav
8	The Evolution of	Sentiment analysis is one of computer science's fastest growing research areas, making it difficult to keep track of all computer science activities. We present a	2018	Mika V. Mäntylä, Daniel

	Sentiment Analysis - A Review of Research Topics, Venues, and Top Cited Papers	computer-assisted analysis of literature, using both text mining and qualitative coding, and reviewing 6,996 Scopus articles. We note that the origins of sentiment analysis are in the early 20th century public opinion research studies and in the 1990s text subjectivity analysis carried out by the computational linguistics community. Nevertheless, the computer-based sentiment analysis outbreak occurred only when subjective texts were available on the Web. As a result, after 2004, 99 percent of the papers were released. Sentiment analysis articles are scattered in multiple places of circulation, and the total number of papers in the top-15 places reflects only about. A minimum of 30% of the articles. We are presenting the top-20 cited articles from Google Scholar and Scopus and a research subject taxonomy. Sentiment analysis has changed in recent years from evaluating online product reviews to Twitter and Facebook social media messages. In addition to product reviews such as stock markets, politics, accidents, healthcare, software engineering and cyberbullying, other subjects broaden the use of sentiment analysis		Graziotin, Miikka Kuuttila
9	A survey on sentiment analysis challenges	Writers produce with rapid internet evolution as forums, social networks, blogs, online portals, comments, thoughts, tips, ratings, and feedback. It may be about books, people, hotels, goods, studies, events etc. that this writer created feeling material. Such emotions are becoming very important to corporations, governments and individuals. While this content is intended to be useful, the use of text mining techniques and sentiment analysis requires a large part of the content generated by this author. But the method of sentiment analysis and assessment faces many obstacles. Such problems become obstacles in measuring the strength of feelings and identifying the correct polarity of feelings. Analysis of sentiment is the practice of applying techniques for natural language processing and text analysis to define and extract subjective text data. This paper provides a survey of the problems of sentiment analysis related to their methods and techniques	2016	Doaa Mohey El - Din Mohamed Hussein
10	Sentiment Analysis: A Perspective on its Past, Present and Future	The proliferation of Internet-enabled devices, including desktops, laptops, tablets, and mobile phones, helps people to communicate, engage, and interact in different Web environments, i.e. forums, social networks, blogs. At the same time, the enormous amount of heterogeneous data generated by the users of these groups provides an unparalleled opportunity to create and employ hypotheses and technologies that search for and retrieve relevant data from the enormous amount of information available and mine for	2012	Akshi Kumar, Teeja Mary Sebastian

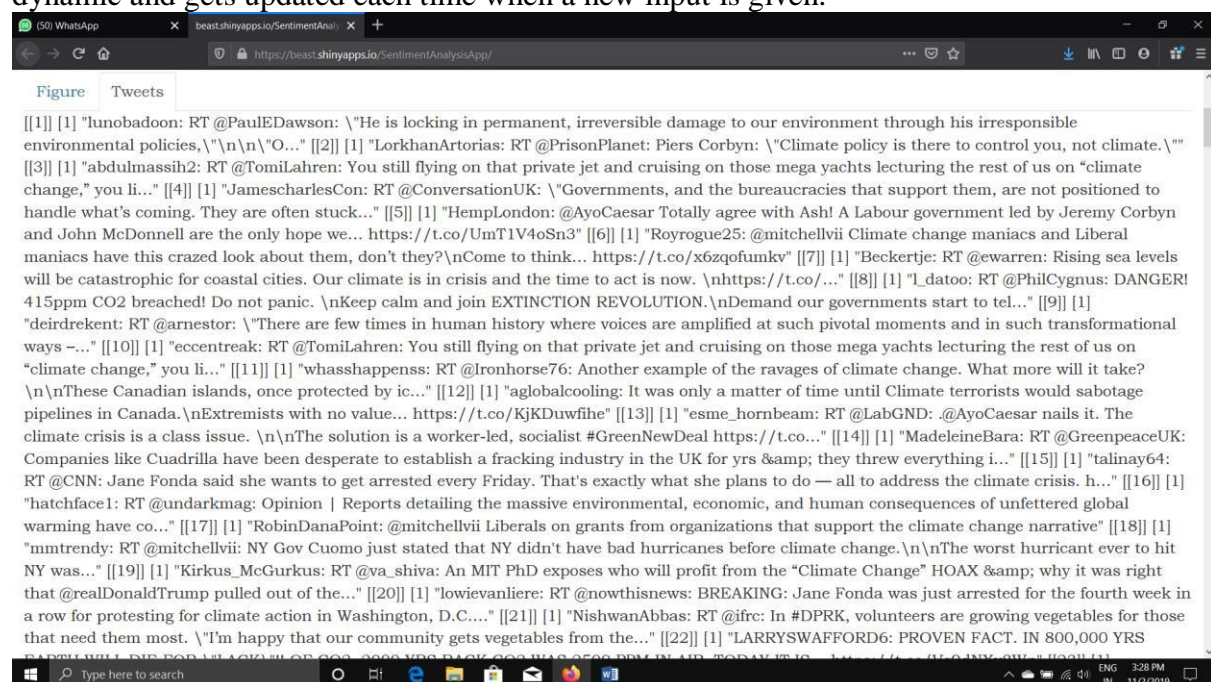
		subsequent opinions. Sentiment Analysis, which automatically extracts and analyses the subjectivities and feelings (or polarities) in written text, has therefore emerged as an active research field. This paper summarizes and reviews substantial research on sentiment analysis, outlining its basic concepts, tasks and levels of granularity. It also gives an overview of the state-of - the-art of some previous attempts to research the psychology of sentiment. It also addresses its realistic and future applications, accompanied by the problems and challenges that will keep the sector competitive and vibrant for years to come.		
11	A Survey on Feature Level Sentiment Analysis	We have been strongly considering friends ' views in recent years, domain experts for making decisions in today's life day. For example, which brand is best for a particular product, whether the current film is good, whether the product delivers better results, etc. Opinion mining, also known as study of sentiment, plays an important role in this process. It's the emotion analysis, i.e. Sentiments, words in natural language. To remove emotions from unstructured data, natural language techniques are applied. Many methods, such as supervised learning, unsupervised learning and hybrid techniques, are used to analyse data. The category of binary information is divided into two subclasses viz. Positive as well as negative. The neutral class expresses a good opinion, and the negative class displays a bad opinion. The paper provides an overview of the approaches proposed and recent advances in these areas and aims to suggest possible directions for research in the field.	2014	Neha S. Joshi, Suhasini A. Itkat
12	Using text mining and sentiment analysis for online forums hotspot detection and forecast	Study of text sentiment, also referred to as computation of emotional polarity, has become a thriving frontier in the field of text mining. This paper uses sentiment analysis and text mining methods to research digital hotspot identification and prediction forums. Next, we create an algorithm for automatically evaluating a text's emotional polarity and for each piece of text to obtain a meaning. Second, to develop unsupervised text mining method, this algorithm is combined with K-means clustering and supporting vector machine (SVM). We use the proposed text mining method to organize the forums into different clusters, each representing a hotspot forum in the current span of time. The data sets used in our empirical studies were collected and compiled from Sina sports forums, covering a total of 31 subject forums and 220,053 posts. Experimental results show that with K-means clustering, SVM forecasting produces highly consistent performance. SVM forecasting's top 10 hotspot forums represent 80	2010	Nan Li, Desheng Dash Wu

		percent of K-means clustering outcomes. The same results were obtained for the top 4 hotspot forums of the year by both SVM and K-means.		
13	Twitter sentiment analysis using hybrid cuckoo search method	Sentiment analysis is one of the influential areas of data mining which explores the detection and evaluation of generally available sentimental material on social media. Twitter is one of those social media that many people use in the form of tweets about certain subjects. Such tweets can be analyzed using clustering-based approaches to find the viewpoints and feelings of the users. Nonetheless, metaheuristic-based clustering approaches outperform traditional methods for sentiment analysis due to the subjective nature of the Twitter datasets. This paper therefore proposes a novel metaheuristic approach (CSK) based on K-means and searching for cuckoo. The suggested approach was used to find the optimal cluster heads from Twitter dataset's sentimental contents. The efficacy of the proposed method was evaluated on various Twitter datasets and contrasted with particle swarm optimization, differential evolution, cuckoo search, enhanced cuckoo search, gauss-based cuckoo search, and two methods of n-grams. Experimental results and statistical analysis verify the current approaches were outperformed by the proposed method. For future research, the proposed method has conceptual implications for analysing the data produced by social media / networks. This approach also has very common practical implications for designing a system that can provide some social issues with definitive feedback.	2017	Avinash Chandra Pandey, Dharmveer Singh Rajpoot, Mukesh Saraswat
14	Sentiment Analysis: An Overview	In response to the availability of informal, opinionated texts such as blog posts and product review websites, a Sentiment Analysis area has developed over the past decade to address the question What do people feel about a particular topic? Bringing together experts in the fields of computer science, computer linguistics, data mining, psychology, and even sociology, Sentiment Analysis extends conventional fact-based text analysis to allow for opinion-oriented information systems. This paper is an overview of Sentiment Analysis, its basic tasks, and the latest techniques developed to tackle the emotionally charged text challenges.	2009	Yelena Mejova
15	Sentiment Analysis and Influence Tracking using Twitter	In social media platforms, an increasing number of consumers are involved. Consumers share their true feelings about a specific brand / product, its functionality, customer service, and how it holds up to competition within these channels. With microblogs booming on the Internet, people on Twitter and other similar services have started to express their opinions	2012	Rushabh Mehta, Dhaval Mehta, Disha Chheda, Charmi Shah, Pramila M. Chawan

	<p>onawiderangeoftopics.Itisimportanttoexaminethe dissemination and impact of data in large-scale networks in an environment where knowledge can distort public opinion. Recent research on social media data to rate users by topical relevance has focused largely on relationships such as "retweet," "follow"and "mention." They also carry out the recorded corpus linguistic analysis and clarify discovered phenomena. We create a feeling classifier using the corpus, which can evaluate positive, negative and neutral feelings for a text. This paper explores how Twitter data is used as acorpusofresearchbyapplying sentimentanalysisand researching different algorithms and methods that help track the effect and impact of an active user / brandon the social network.</p>		
--	---	--	--

IMPLEMENTATION

The Dataset contains tweets retrieved based on the topic the user gives. The dataset is dynamic and gets updated each time when a new input is given.



Analytical methods:

The methodology used is as follows:

1. Retrieve the data from twitter regarding a particulartopic
2. Perform the pre-processing steps on the text data, which involves removing of URLs, typecasting, removing stop wordsetc.

3. Create a term-document matrix to visualize the counts of words.
4. Create a word cloud of the words to represent the dominance of words.
5. Use different libraries in R to perform sentiment analysis on the tweets.
6. Rate the tweets on 10 emotional polarities and represent in the form of a Histogram.

Code:

```
#####Libraries required #####
library(shiny)
library(httr)
library(devtools)
library(base64enc)
library(twitter)
library(syuzhet)
library(lubridate)
library(ggplot2)
library(scales)
library(reshape2)
library(dplyr)
library(tm)
library(wordcloud)
#####Connection to Twitter API#####
consumer_key = "0jgpQihZLGoValA6EOh1xp6vp"
consumer_secret = "mBdajaSFvetDYLJgmH8038vwjWFstQU95PjNVvyvTgnVLLDjx1"
access_token = "1168606699420356608-a6gwYeoHTVx8WvORCVHP7mpYRHLpuk"
access_secret = "5tfsPwcVACvxj8fyponTFpEDBe3oCUDinscfnpSptRHsT"
setup_twitter_oauth(consumer_key, consumer_secret, access_token, access_secret)
#####Front End Code#####
ui <- fluidPage(theme = "to1.css",
  h1("Social Network analysis and Sentiment Analysis"),
  wellPanel(
    fluidRow(
      column(3, textInput(inputId="topic", label="Topic")),
      column(3, numericInput(inputId="ntweet", label="Count", 300)),
      radioButtons(inputId="citwit", label="wanna see the tweets?", c("Yes"="one",
                                                                    "No"="two")),
      selectInput(inputId="choice", label="what do you want today?", c("Histogram of node degree"="two",
                                                                    "Network Diagram"="three",
                                                                    "Community Detection"="four",
                                                                    "wordcloud"="five",
                                                                    "Sentiment Analysis"="six")),
      checkboxInput(inputId="comalgo", "Enable community detection"),
      conditionalPanel(
        condition = "input.comalgo == true",
        selectInput(inputId="choice1", label="which Algorithm do you want to use?", c("Edge Betweenness"="one",
                                                                    "Label Propagation"="two",
                                                                    "Fast-greedy"="three"))
      )
    )
  )
  tabsetPanel(
    tabPanel(title = "Figure",
      plotOutput("mainstuff")
    ),
    tabPanel(title = "Tweets",
      textOutput("tweets")
    )
  )
)
```

```
#####Backend#####  
server function(input,output){  
output$mainstuCC = renderPlot({  
  
    a<-reactive({input5choiceI})  
tweets      searchTwItter{InputStopsC,   nput$Irrcweet,1   'en'}  
weeks df <- wet r+oDF(meets)  
write.csv(tweetsd+,JJJe="appexperimerrcs.csv",rd.nanes=F) data  
read.csv("aoepxperi ments.csv",header=zj}  
  
str  
X euld corpus  
corpus iconvCData%text,to='UTF-8',sub="byte")  
corpus corpus(vectorsource(corpusJJ  
  
E$$$$SeE0e0e0e0e0e####88Y8Y8YSeseSes0e0e0e#### Aclean*eMr ESesEsES0e0e###C#####V8V8#8#8Scseec#####e0 ##  
corpus tm_map(corpus,tolower)  
corpus tm_map(corpus,removePunctua*iou)  
corpus tm_map(corpus,removewumbers)  
c3eanset <- tcrmap(corpus, removeWords, stop^ords('english'))  
RemoveURL<-function(x)gsub('http[:]alnum:JJ","','x)  
Acleanset tm_map(cleanset,content_transformer(removeURL))  
cleanset tm_map(cleanset,  
  
G#####S#S#####G#####T ern documentnatrix##G#####S#S#####f####G#G#####IS##  
com Termocument*aatrix(cleasnetj  
tno as.mat rx (tdu)  
tdm tdmErrowsums(tdm)>10,j  
  
#####Network of terms#####  
library(igraph)  
*[rmd>1] <- 1  
termM <- tdm %%% T(tdm)  
  
g graph.adj.agency(terry\ , weJghted= T , mode ='undlr ected' )  
  
v(g)$label <- v(g)$name  
v(g)$degree <- degree(g)  
#show the tweets  
#####Histogram of node degree#####  
hist(v(g $degree,  
breaks=t0El ,  
col ="red ",  
n = "uistog anofcode oegree", y4 at  
-" frequency",  
x4at "ne9recof verit ces"))  
  
#####Network diagram#####  
set.seed(222)  
  
vertex.col3or='green',  
vertex.sJze=10,  
vertex.3abel.dist=1.5,  
vertex.3abel=NA))  
  
#####Community detection#####  
if(y())== TRUE){  
if(a()=="one"){  
#1ST ALGORITHM  
comm <- cluster_edge_betweenness(g)  
plot(comm,g)]  
e3se if(a()!="two"){  
82ND ML GORE TM  
prop c1ust er_1 abel_prop(g)  
e3se' ifCaC)= ve e hre") f  
  
greeco<- c1ust er_f ast_greedys(as und 1 r eded {g} )  
p4 otr eed.as .undir ected(g)} }  
  
#####YY#####EC#####EEeyveeeY#####wordcloud#EEEE#####YV#####C#####EEeyVeceyY#####  
w sor*(rowsums(tdm), decreasing =TrVE)  
#set seed Tar repeatabily ty  
set.seed(222)  
wordcloud(words nanes(w),  
freq u,  
max. 1 TO,  
random.order=r,m1n.  
freq=s,  
colors brever.pal(8,'oark2 '),  
scale c5,0.3},  
rot.per=0.3)}
```

```
#####sentiment analysis#####
if(x()=="six"){
  data <- read.csv("appexperiments.csv", header = T)
  tweets <- iconv(data$text, to = 'UTF-8')
  # Obtain sentiment scores
  s <- get_nrc_sentiment(tweets)
  # Bar plot
  barplot(colSums(s),
          las = 2,
          col = rainbow(10),
          ylab = 'count',
          main = 'Sentiment Scores for data Tweets'})
}
output$tweets <- renderPrint({
  z<-reactive({input$scitwit})
  if(z()=="one"){
    tweets <- searchTwitter(input$topic,n=input$ntweet,lang='en')
    tweets
  }
})
}
shinyApp(ui =ui,server= server)
```

Execution of the code:

We have Given Long Topic Comments in the code indicating which part has what functionality!

Let us briefly explain each segment:

1. LibrariesRequired

These libraries contain all the functions I used in the code, we need to initialize them in the code in order to use them.

2. Setup connection withTwitter

To connect with Twitter API, we need an account in apps.twitter.com where we will get Tokens and secret keys. We need to initialize them in the code and use them to connect to API using R. The library required is (twitter)

3. Front endcode

We don't need to write any html, css code to develop UI in Rshinyapps, there are inbuilt functions to take care of that. If we run that part of the code in Terminal we can see the html output.

4. Backendcode

Here is where the magic takes place, here contains the retrieving Tweets, preprocessing them and showing the outputs. I added some reactive variables to show only certain output according to the user's wish. You can see that I used "if" conditions to implement that. Detailed explanation for each segment is given below.

5. CleanText

Here many preprocessing steps like Cast typing, removing punctuation, White spaces, removing url'setc are carried out. To make sure the data we are using for the analysis is useful and not of unnecessary text.

6. Term Document matrix

This basically is matrix consisting of the tweets on the columns and terms on the rows. It contains of "0" and "1", 1 means that that term is in that tweet "0" means that the word is not in the tweet.

7. Network Diagram,Histogram

This is formed by the Term-Document-Matrix we created in the previous step. The network diagram consists of nodes(terms) which contains connections based on the

Presence of them in the same tweet. The histogram is then plotted from this network diagram; it shows how many Nodes are present with what degree.

8. CommunityDetection

I have implemented 3 different algorithms for finding communities between the terms namely edge-betweenness, label-propagation, Fast-greedy.

I have included the labels of the terms here so that the user can clearly see what terms are grouped together.

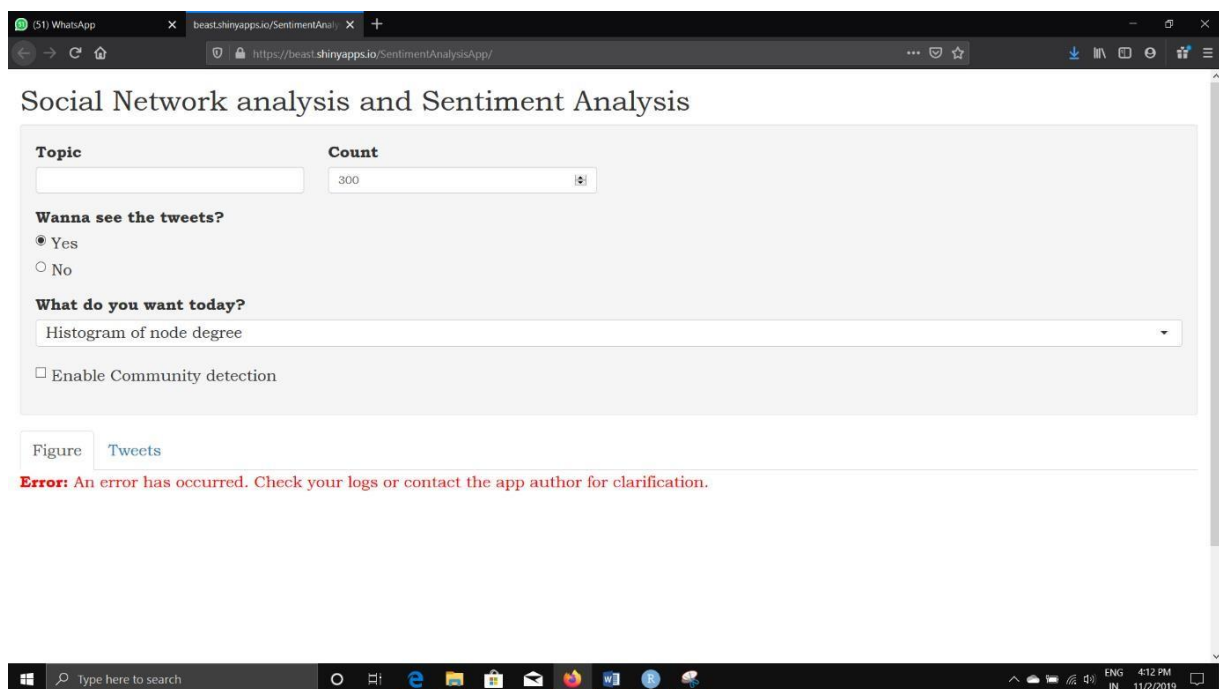
9. Word-cloud

This is classic inference technique where we can see what terms are more dominant in the whole corpus (tweets). The size of the term in the word-cloud is based on the Frequency in the corpus

10. SentimentAnalysis

This uses the Lexicon method. For every word in the English language, the library has categorized into 10 categories (1 or 0), now when we give the tweets, it will count the number of 1's in every category and then show a histogram of the output, we can get a basic inference by taking a look at the histogram. The parameters are "anger", "anticipation", "disgust", "fear", "joy", "sadness", "surprise", "trust", "negative", "positive".

Explanation of vis idiom:

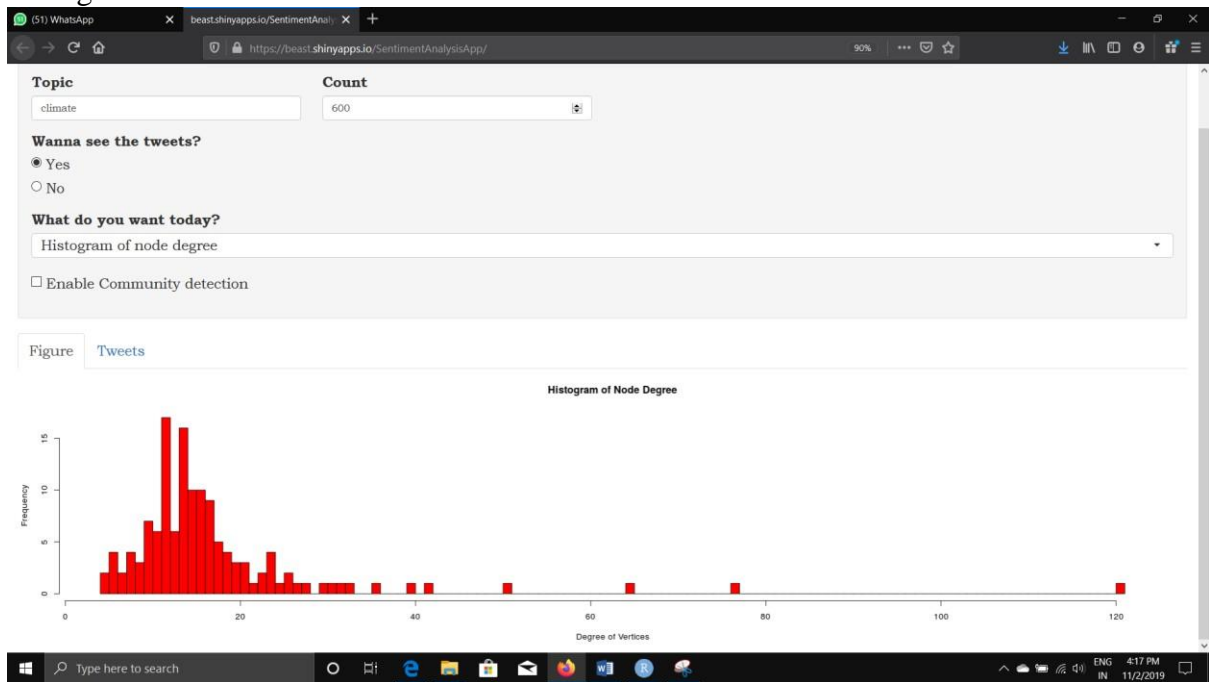


This is the Website which we made. Here we can:

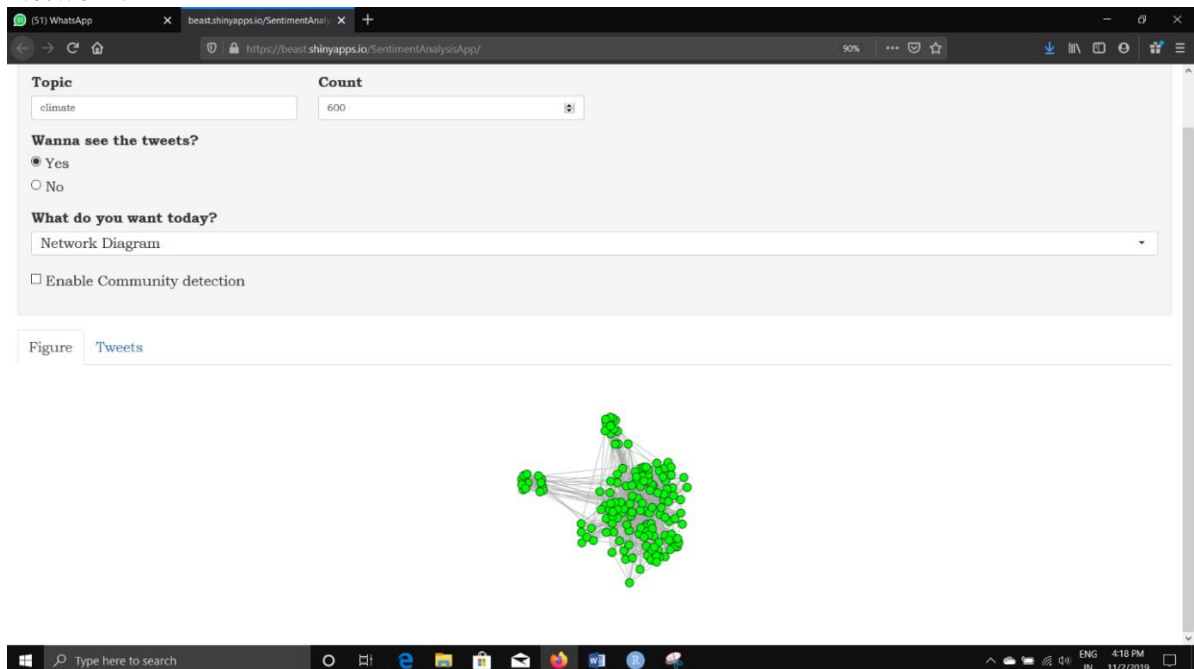
1. Enter a topic on which we want to retrieve the tweets from
2. Enter the number of tweets to extract
3. Select the Visualization technique

RESULT

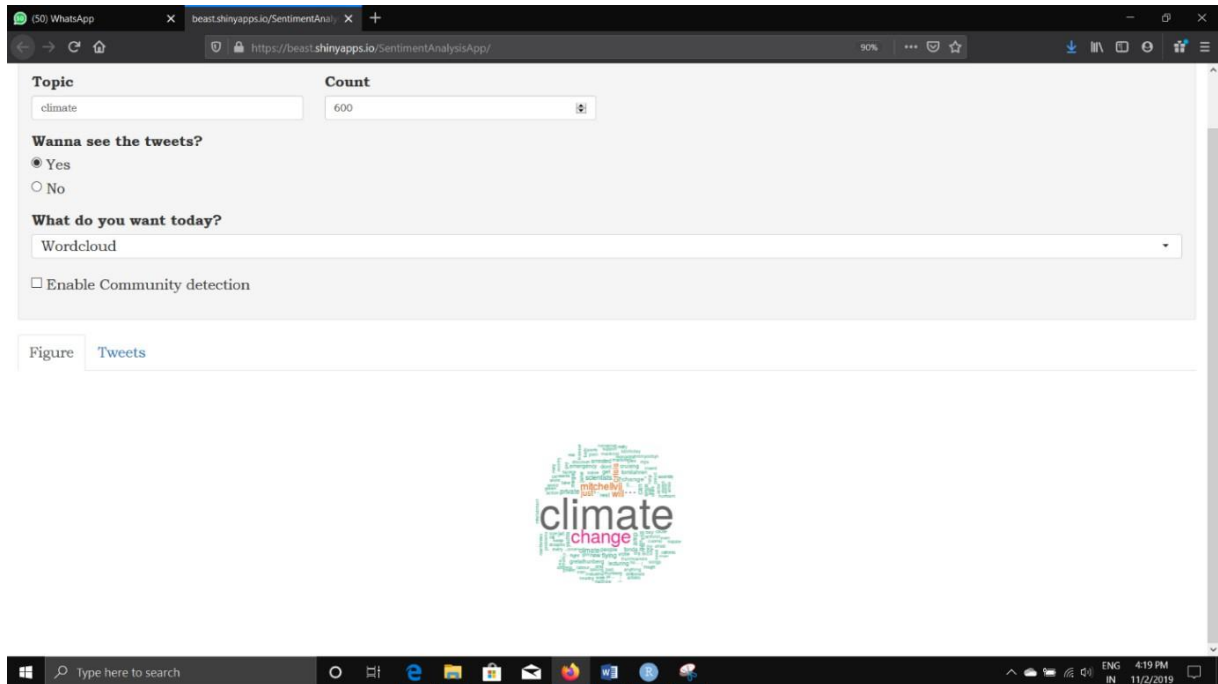
Histogram:



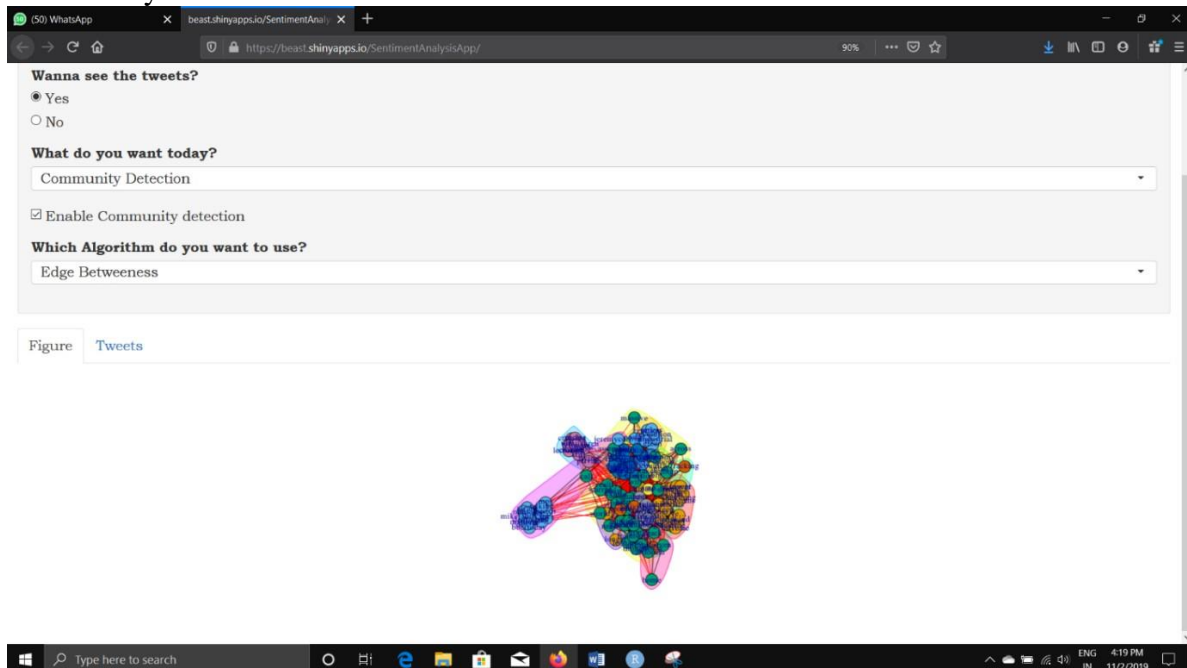
Network:



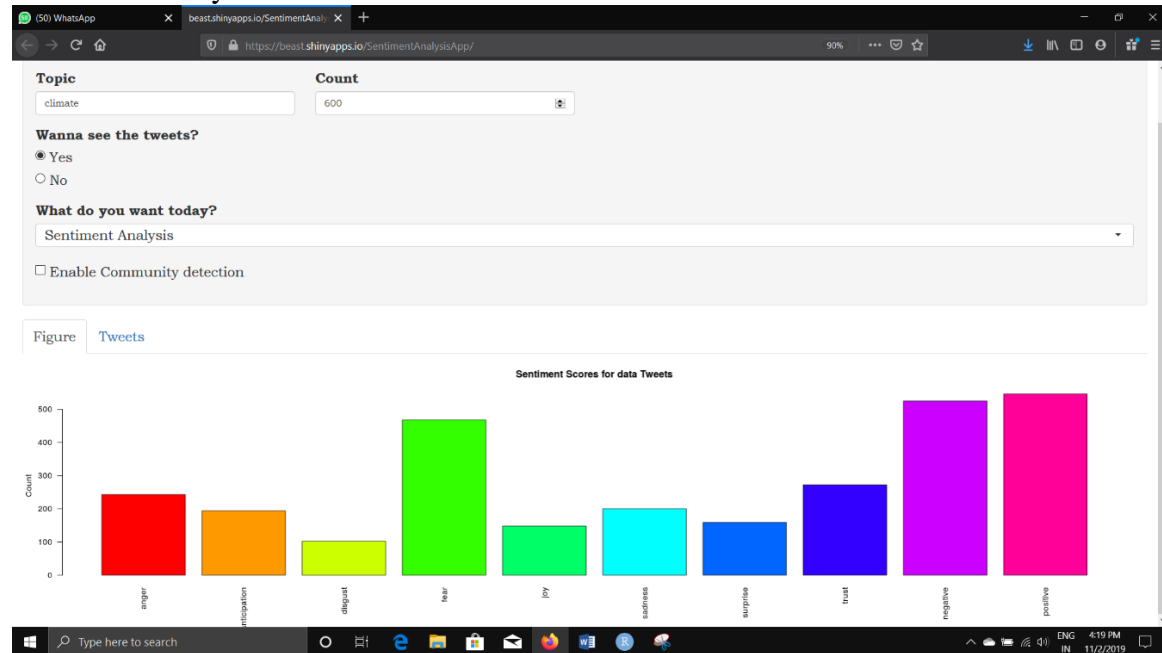
Word cloud:



Community detection:



Sentiment Analysis:



Till now we have seen many different types of visualizations and plots from which we can infer some clues. For example, in the Plot of “Histogram of node degree” we can see that the better the corpus (many number of tweets) the more left shifted the bar plot. So, we can say that the “more left shifted” the histogram looks the more “rich data” we have.

As you can clearly see in the Network diagram in have disabled the node labels as it would be very clumsy to view and infer anything but it is a very important component in the Community detection phase where we use 3 different algorithms to find communities in the network diagram. The 3 Algorithms give different results but we can see almost always the similar set of terms are grouped together. Also it can also be seen from the “word cloud” that it is quite a fancy way of representing data, but there is more to it, the size of the term in the word cloud represents the dominance or the frequency of that particular term which can be seen from the term-document matrix (used to create the Network Diagram).

The sentiment analysis works in a pretty straightforward manner, the function rates every word in the English language on 10 parameters which you can see the plot. For example the word “happy” may have the 0 in the emotion “anger” and 1 in the emotion “excited”, like this they are assigned 0’s or 1 in 10 parameters. The plot basically shows how many terms are there in the corpus which portray the respective emotion.

While the website execution time depends on the user’s internet bandwidth, the processing time is quite slow as the backend runs on the RStudio Shiny apps server.

The final of this project is a live and fully functioning website which can carry out all the functionalities mentioned in the Reviews of this project. The Website can be accessed from here. <https://beast.shinyapps.io/SentimentAnalysisApp/>

FUTURE WORK

Many researchers and business organization are showing lot of in sentiment analysis Sentiment analysis method still now have been used to detect the difference in the thoughts and opinions of all the users who use the social media. Researchers and Businesses are very much interested in understanding the thoughts of people and how they are responding to everything happening around them. Companies use this to evaluate their advertisement campaigns and to improve their products. There is a lot of scope in analysing the video and images on the web. Now a days, with the advent of Facebook, Instagram and Video vines people are expressing their thoughts with pictures and videos along with text. Sentiment analysis can be used for this change.

REFERENCES

- [1] Godbole, N., Srinivasaiah, M., & Skiena, S. (2007). Large-scale sentiment analysis for news and blogs. In *ICWSM 2007 - International Conference on Weblogs and Social Media*.
- [2] Bautin, M., Vijayarenu, L., & Skiena, S. (2008). International sentiment analysis for news and blogs. In *ICWSM 2008 - Proceedings of the 2nd International Conference on Weblogs and Social Media*.
- [3] Athar, A., & Teufel, S. (2012). Context-enhanced citation sentiment detection. In *NAACL HLT 2012 - 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Proceedings of the Conference*.
- [4] Asghar, M. Z., Khan, A., Ahmad, S., & Kundi, F. M. (2014). A Review of Feature Extraction in Sentiment Analysis. *J. Basic. Appl. Sci. Res.*
- [5] Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*. <https://doi.org/10.1186/s40537-015-0015-2>
- [6] Xu, J., Zhang, Y., Wu, Y., Wang, J., Dong, X., & Xu, H. (2015). Citation Sentiment Analysis in Clinical Trial Papers. *AMIA ... Annual Symposium Proceedings. AMIA Symposium*.
- [7] B., V., & M., B. (2016). Analysis of Various Sentiment Classification Techniques. *International Journal of Computer Applications*. <https://doi.org/10.5120/ijca2016909259>
- [8] Mäntylä, M. V., Graziotin, D., & Kuuttila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review*. <https://doi.org/10.1016/j.cosrev.2017.10.002>
- [9] Hussein, D. M. E. D. M. (2018). A survey on sentiment analysis challenges. *Journal of King Saud University - Engineering Sciences*. <https://doi.org/10.1016/j.jksues.2016.04.002>
- [10] Akshi Kumar, T. M. S., Kumar, A., & Sebastian, T. M. (2012). Sentiment Analysis on Twitter. *IJCSI International Journal of Computer Science Issues*.

- [11] Joshi, N., & Itkat, S. (2014). A Survey on Feature Level Sentiment Analysis. *International Journal of Computer Science and Information Technologies*.
- [12] Li, N., & Wu, D. D. (2010). Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Decision Support Systems*.
<https://doi.org/10.1016/j.dss.2009.09.003>
- [13] Chandra Pandey, A., Singh Rajpoot, D., & Saraswat, M. (2017). Twitter sentiment analysis using hybrid cuckoo search method. *Information Processing and Management*.
<https://doi.org/10.1016/j.ipm.2017.02.004>
- [14] Mejova, Y. (2009). Sentiment Analysis : An Overview Comprehensive Exam Paper. Science.
- [15] Anber, H., Salah, A., & El-Aziz, A. A. A. (2016). A Literature Review on Twitter Data Analysis. *International Journal of Computer and Electrical Engineering*.
<https://doi.org/10.17706/ijcee.2016.8.3.241-249>