

KIET Group of Institutions, Ghaziabad

Information Technology



Report on

SENTIMENTAL ANALYSIS

Supervised By :- Prof. Navpreet Kaur

Submitted By :- Prabal Katiyar

Roll no. – 1802913106

B.Tech(Information Technology)

5th Semester

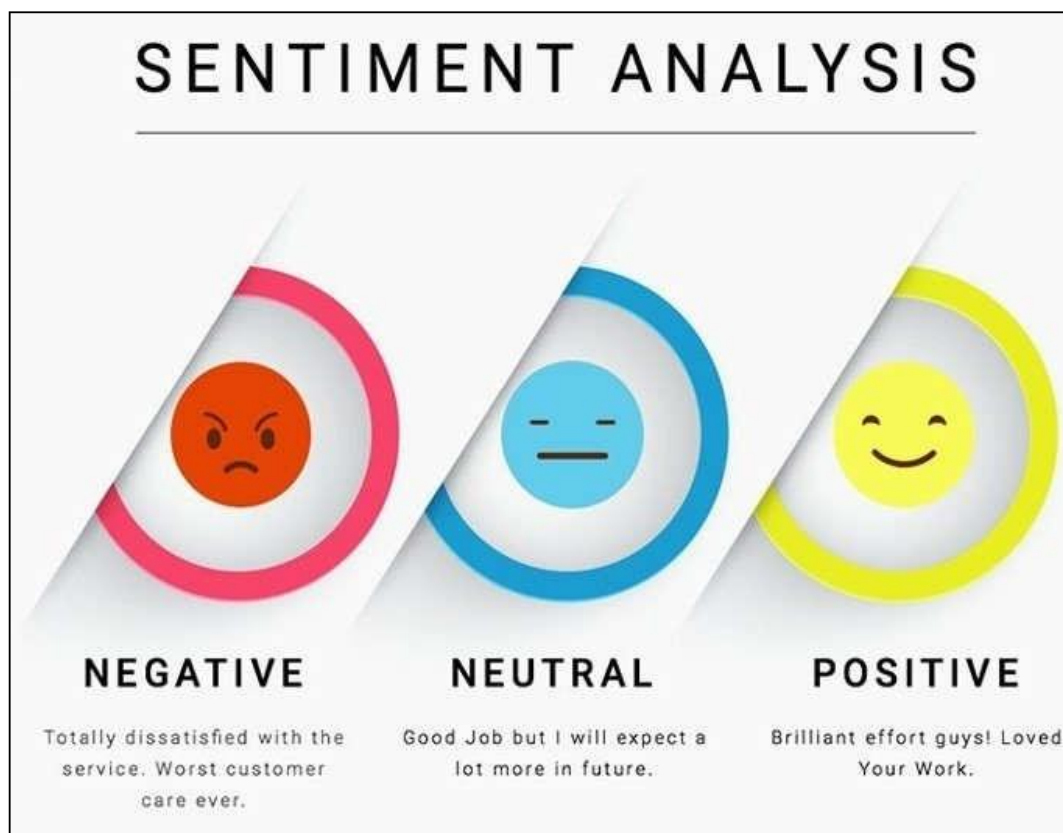
Section – B

Table of Content

1. Introduction
2. Twitter Sentiment Analysis
3. Application of Sentiment Analysis
4. Project Overview
5. Implementation (with snap shots)
6. Conclusion

INTRODUCTION

Generally speaking, sentiment analysis aims to determine the attitude of a speaker, writer, or other subject with respect to some topic or the overall contextual polarity or emotional reaction to a document, interaction, or event. It is a way to evaluate written or spoken language to determine if the expression is favorable, unfavorable, or neutral, and to what degree. The applications of sentiment analysis are broad and powerful. The ability to extract insights from social data is a practice that is being widely adopted by organizations across the world.



TWITTER SENTIMENT ANALYSIS

Twitter is an online news and social networking service that enables users to send and read short 140-character messages called "tweets". Registered users can read and post tweets, but those who are unregistered can only read them.

Hence Twitter is a public platform with a mine of public opinion of people all over the world and of all age categories. As of October 2016, Twitter has more than 315 million monthly active users. Twitter Sentiment Analysis the process of determining the emotional tone behind a series of words, used to gain an understanding of the attitudes, opinions and emotions expressed within an online mention.

WHY TWITTER SENTIMENT ANALYSIS?

Using **sentiment analysis** tools to **analyze** opinions in **Twitter data** can help companies understand how people are talking about their brand. ... This is one of the reasons why social listening — monitoring conversations on social media platforms — has become a crucial process in social media marketing.

Applications of Sentiment Analysis

Sentiment Analysis finds its application in a variety of domains.

A. Online Commerce

The most general use of sentiment analysis is in ecommerce activities. Websites allows their users to submit their experience about shopping and product qualities. They provide summary for the product and different features of the product by assigning ratings or scores. Customers can easily view opinions and recommendation information on whole product as well as specific product features. Graphical summary of the overall product and its features is presented to users. Popular merchant websites like amazon.com

provides review from editors and also from customers with rating information. <http://tripadvisor.in> is a popular website that provides reviews on hotels, travel destinations. They contain 75 millions opinions and reviews worldwide. Sentiment analysis helps such websites by converting dissatisfied customers into promoters by analyzing this huge volume of opinions.

B. Voice of the Market (VOM)

Voice of the Market is about determining what customers are feeling about products or services of competitors. Accurate and timely information from the Voice of the Market helps in gaining competitive advantage and new product development. Detection of such information as early as possible helps in direct and target key marketing campaigns. Sentiment Analysis helps corporate to get customer opinion in real-time. This real-time information helps them to design new marketing strategies, improve product features and can predict chances of product failure. Zhang et al. proposed weakness finder system which can help manufacturers find their product weakness from Chinese reviews by using aspects based sentiment analysis. There are some commercial and free sentiment analysis services are available, Radiant6, Sysomos, Viralheat, Lexalytics, etc. are commercial services.

C. Voice of the Customer (VOC)

Voice of the Customer is concern about what individual customer is saying about products or services. It means analyzing the reviews and feedback of the customers. VOC is a key element of Customer Experience Management. VOC helps in identifying new opportunities for product inventions. Extracting customer opinions also helps identify functional requirements of the products and some non-functional requirements like performance and cost.

D. Brand Reputation Management

Brand Reputation Management is concern about managing your reputation in market. Opinions from customers or any other parties can damage or enhance your reputation. Brand Reputation Management (BRM) is a product and company focused rather than customer. Now, one-to-many conversations are taking place online at a high rate. That creates opportunities for organizations to manage and strengthen brand reputation. Now Brand perception is determined not only by advertising, public relations and corporate messaging.

Brands are now a sum of the conversations about them. Sentiment analysis helps in determining how company's brand, product or service is being perceived by community online.

E. Government

Sentiment analysis helps government in assessing their strength and weaknesses by analyzing opinions from public. For example, "If this is the state, how do you expect truth to come out? The MP who is investigating 2g scam himself is deeply corrupt.". this example clearly shows negative sentiment about government. Whether it is tracking citizens' opinions on a new 108 system, identifying strengths and weaknesses in a recruitment campaign in government job, assessing success of electronic submission of tax returns, or many other areas, we can see the potential for sentiment analysis.

PROJECT OVERVIEW

Tweets are imported using R and the data is cleaned by removing emoticons and URLs. Lexical Analysis is used to predict the sentiment of tweets and subsequently express the opinion graphically through ggplots.

SYSTEM REQUIREMENTS

- Installation of R
- Twitter Authentication to access API

FEATURES

1. Extraction of Tweets

- (i) Setup an app in Twitter: <https://apps.twitter.com>
- (ii) twitteR - Provides an interface to the Twitter web API
- (iii) ROAuth - R Interface for OAuth
- (iv) Create twitter authenticated credential object, It is done using consumer key, consumer secret, access token, access secret.

2. Loading Tweets

3. Cleaning Tweets

The tweets are cleaned in R by removing:

- Extra punctuation
- Stop words (Most commonly used words in a language like the, is, at, which, and, on,.)
- Redundant Blank spaces
- URLs

4. Use **get_nrc_sentiment()** function to analyze sentiment.

5. Use sentiment scores for visualizations with **ggplot()** function.

Implementation (using get_nrc_sentiment function)

1-Loading Libraries:

```
library(twitterR)
library(tm)
library(wordcloud)
library(syuzhet)
library(ggplot2)
library(dplyr)
```

2-Setup twitter oauth using access token:

```
consumer_key<-"iticSogt5Kux9AqXFwh6d3pFI"
consumer_secret<-"iEXSEadVhIXDSxzSFGiwAaJeiAYfmNfQTmgKu8dDeZqDCwbpcQ"
access_token<-"4923758112-7jJIzLRnYGJzUqfTKTca3WurySlevy77tGAm0Yo"
access_secret<-"cCH1sQr68r5mUyItaKK5pffKXhEGqig1MxyM8VSAqHnmP"
setup_twitter_oauth(consumer_key, consumer_secret, access_token,
                    access_secret)
```

3-Loading Tweets on “Lockdown” contain 1000 tweets:

```
searchnew<-searchTwitter ("Lockdown", n=1000, lang="en")
searchnew

search_text_vec=sapply(searchnew, function(x) x$getText())
search_text_vec
```


4-Data cleaning:

```
search_tm=Corpus(VectorSource(search_text_vec))
inspect(search_tm)

#remove Punctuation
search_tm_level1=tm_map(search_tm,removePunctuation)

#remove numbers
search_tm_clean=tm_map(search_tm_level1,removeNumbers)

#remove blank spaces
search_tm_clean=tm_map(search_tm_clean,stripWhitespace)

#remove stopwords
search_tm_clean=tm_map(search_tm_clean,removeWords,stopwords("english"))

#convert tweets to lower
search_tm_clean=tm_map(search_tm_clean,content_transformer(tolower))

#remove URL
removeURL<-function(x) gsub('http[[:a1num:]]*', '',x)
search_tm_clean=tm_map(search_tm_clean,content_transformer(removeURL))
```

5-Converting in textdocument matrix:

```
tdm<-TermDocumentMatrix(search_tm_clean)

tdm
tdm<-as.matrix((tdm))
tdm[1:10,1:20]
```

6-Creating wordcloud:

```
#word cloud
w<-sort(rowSums(tdm),decreasing=TRUE)
set.seed(222)
wordcloud(words=names(w),freq=w,
           random.order = F,min.freq = 3,
           colors=brewer.pal(8,'Dark2'),
           scale=c(5,0.3))
```

7- wordcloud output:



8-Obtain sentiment using “**get_nrc_sentiment**” function built in syuzhet package in R.

```
#obtain Sentiments  
s<-get_nrc_sentiment(search_tm_clean)  
head(s)
```

9-Barplot to visualize the sentiments:

```
#barplot  
barplot(colSums(s),las=2,col=rainbow(10),ylab = "count",  
        main = "Sentimental analysis")
```

Conclusion using get_nrc_sentiment function–

