sentimental analysis

Prem Kumar Kamasani

March 28, 2019

pos.words <- read.csv("positive.csv")  
neg.words <- read.csv("negative.csv")  
pos.words <- scan("positive.csv",what = 'character')  
neg.words <- scan("negative.csv",what = 'character')  
apple<-read.csv("apple1101.csv", header = TRUE, sep = ",");

pos.words variable has positive words. neg.words variable has negative words. Data of apple tweets are taken from twitter api.

Our aim is to find sentimental score as well emotion analysis on tweets of Apple. positive word has score of +1 and negative word has score of -1. Range of scores =[-5, +5]

score.sentiment = function(sentences, pos.words, neg.words, .progress='.progress')  
{  
 require(plyr)  
 require(stringr)  
 scores = laply(sentences,   
 function(sentence, pos.words, neg.words)  
 {  
 sentence = gsub('[[:punct:]]', '', sentence)  
 sentence = gsub('[[:cntrl:]]', '', sentence)  
 sentence = gsub('\\d+', '', sentence)  
 sentence = tolower(sentence)  
 word.list = str\_split(sentence, '\\s+')  
 words = unlist(word.list)  
 pos.matches = match(words, pos.words)  
 neg.matches = match(words, neg.words)  
 pos.matches = !is.na(pos.matches)  
 neg.matches = !is.na(neg.matches)  
 score=sum(pos.matches)-sum(neg.matches)  
 return(score)  
 },  
 pos.words,   
 neg.words,  
 .progress='.progress' )  
 scores.df = data.frame(score=scores, text=sentences)  
 return(scores.df)  
}

apple<-as.data.frame(apple)  
result <- score.sentiment(apple$text,pos.words,neg.words)

## Loading required package: plyr

## Warning: package 'plyr' was built under R version 3.5.2

## Loading required package: stringr

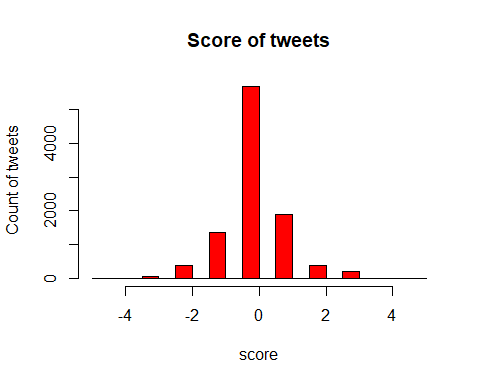
## Warning: package 'stringr' was built under R version 3.5.2

## Warning: Cannot find progress bar progress\_.progress

summary(result$score)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## -5.000 0.000 0.000 0.103 1.000 5.000

hist(result$score,col ="red", main ="Score of tweets",ylab = " Count of tweets", xlab = "score")

 From the above bar graph we can observe that neutral tweets are more.

library(syuzhet)

## Warning: package 'syuzhet' was built under R version 3.5.3

library(lubridate)

## Warning: package 'lubridate' was built under R version 3.5.3

##   
## Attaching package: 'lubridate'

## The following object is masked from 'package:plyr':  
##   
## here

## The following object is masked from 'package:base':  
##   
## date

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.5.3

library(scales)

## Warning: package 'scales' was built under R version 3.5.3

##   
## Attaching package: 'scales'

## The following object is masked from 'package:syuzhet':  
##   
## rescale

library(reshape2)

## Warning: package 'reshape2' was built under R version 3.5.3

library(dplyr)

## Warning: package 'dplyr' was built under R version 3.5.3

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:lubridate':  
##   
## intersect, setdiff, union

## The following objects are masked from 'package:plyr':  
##   
## arrange, count, desc, failwith, id, mutate, rename, summarise,  
## summarize

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

we use these libraries to calculate the scores of emotions for Apple tweets data.

apple<-read.csv("apple1101.csv", header = TRUE, sep = ",")  
tweets<-iconv(apple$text)  
sentiments<-get\_nrc\_sentiment(tweets)  
barplot(colSums(sentiments), las=2, col=rainbow(10),ylab = "count", main = "Sentiments for Apple Tweets")

