Sentiment Analysis on Text data

Uche Kalu

4/17/2022

Sentiment Analysis on Tweet and Tweet-related information of Apple company

Sentiment Analysis or Opinion Mining is a Natural Language Processing (NLP) technique that is performed on textual data to help businesses monitor brand and product sentiment in customer feedback, and understand customer needs.

It is used to determine whether data is positive, negative or neutral

DATA

```
mydata <- read.table("apple.txt", header = T, sep = ',')

ourdata <- iconv(mydata$text, to = "UTF-8")
```

Obtain Sentiment scores for each Tweet

This will classify all the indexed tweets (each tweet will have an index number) among multiple sentiment column that best fits/describes that tweet

Lets see the first 6 tweets and their Sentiment scores/values

```
s <- get_nrc_sentiment(ourdata)
head(s)
```

```
##   anger anticipation disgust fear joy sadness surprise trust negative positive
## 1     0            0       0    0   0       0        0     0        0        1
## 2     0            0       0    0   0       0        0     0        0        1
## 3     0            0       0    0   0       0        1     0        0        0
## 4     1            0       2    2   0       1        0     0        3        0
## 5     0            0       0    0   0       0        0     0        0        0
## 6     0            0       0    0   0       0        0     0        0        0
```

Lets see the 4th tweet that contains all this sentiments

```
ourdata[4]
```

```
## [1] "RT @SylvaCap: Things might get ugly for $aapl with the iphone delay. With $aapl down that means
```

Lets see the 1st tweet that contains all this sentiments

```
ourdata[1]
```

```
## [1] "RT @option_snipper: $AAPL beat on both eps and revenues. SEES 4Q REV. $49B-$52B, EST. $49.1B ht
```

We will classify the stated word - delay (taken from a tweet) among the multiple sentiment column headers that best fits/describes the stated word

```
get_nrc_sentiment('delay')
```

```
##   anger anticipation disgust fear joy sadness surprise trust negative positive
## 1     1            0       1    1   0       1        0     0        1        0
```

We will classify the stated word - ugly (taken from a tweet) among the multiple sentiment column headers that best fits/describes the stated word

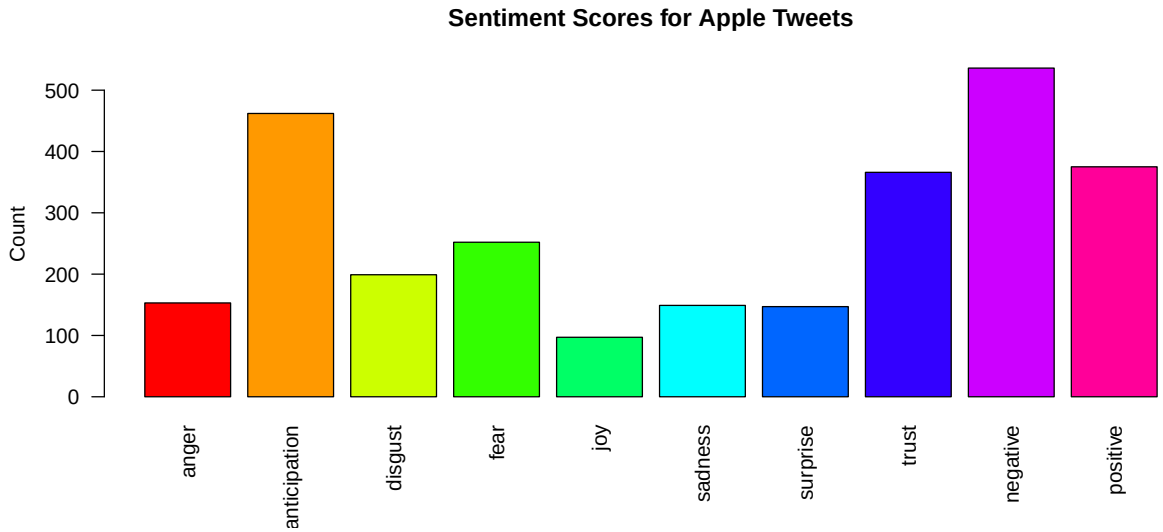
```
get_nrc_sentiment('ugly')
```

```
##   anger anticipation disgust fear joy sadness surprise trust negative positive
## 1     0            0       1    0   0       0        0     0        1        0
```

Visualizing the Sentiments Scores

Create a Bar plot using the Sentiment Scores

```
barplot(colSums(s),
        las = 2,
        col = rainbow(10),
        ylab = 'Count',
        main = 'Sentiment Scores for Apple Tweets')
```



Lets display the Sentiment scores by percentage value

```r
we <- 100*colSums(s)/sum(s)

barplot(we,
        las = 2,
        col = rainbow(10),
        ylab = 'Count',
        main = 'Sentiment Scores for Apple Tweets')
```