Analisis_Sentimen

Naufal Nur Fahriza/Kevin Bimo S

11/29/2021

```
library(e1071)
```

```
## Warning: package 'e1071' was built under R version 4.1.2
```

```
library(caret)
```

```
## Warning: package 'caret' was built under R version 4.1.2
```

```
## Loading required package: ggplot2
```

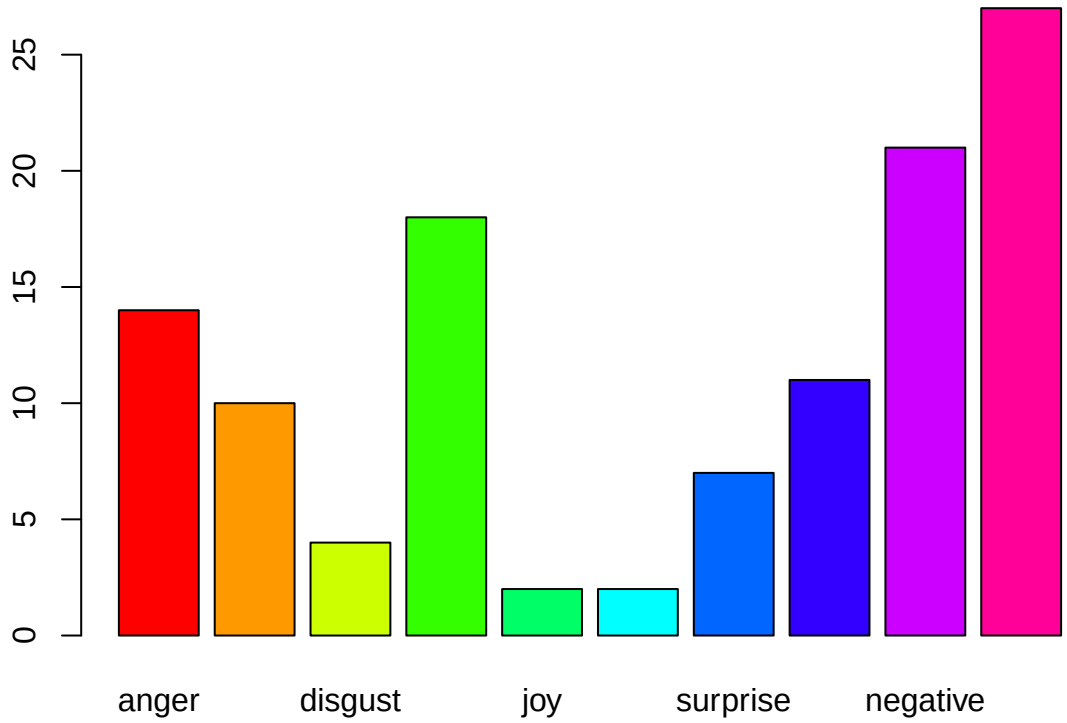
```
## Loading required package: lattice
```

```
library(syuzhet)
```

```
## Warning: package 'syuzhet' was built under R version 4.1.2
```

```
datavaksin<-read.csv("vaksin.csv", stringsAsFactors = FALSE)
review <- as.character(datavaksin$text)
s<-get_nrc_sentiment(review)
```

```
review_combine<-cbind(datavaksin$text, s)
par(mar=rep(3,4))
a<-barplot(colSums(s), col=rainbow(10), ylab='count', main='sentiment analisis')
```




```
brplt<-a
```

```
library(tm)
```

```
## Warning: package 'tm' was built under R version 4.1.2
```

```
## Loading required package: NLP
```

```
##
## Attaching package: 'NLP'
```

```
## The following object is masked from 'package:ggplot2':
##
##     annotate
```

```
library(RTextTools)
```

```
## Loading required package: SparseM
```

```
##
## Attaching package: 'SparseM'
```

```
## The following object is masked from 'package:base':
##
##     backsolve
```

```
library(e1071)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

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

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

```
library(caret)
df<-read.csv("vaksin.csv", stringsAsFactors = FALSE)
glimpse(df)
```

```
## Registered S3 method overwritten by 'cli':
##   method     from
##   print.tree tree
```

```
## Rows: 1,000
## Columns: 2
## $ X    <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19~
## $ text <chr> "hi siapa warga astrazeneca u all dah dapat booster ke \n\nke aku~
```

```
set.seed(20)
df<-df[sample(nrow(df)),]
df<-df[sample(nrow(df)),]
glimpse(df)
```

```
## Rows: 1,000
## Columns: 2
## $ X    <int> 223, 831, 568, 122, 927, 958, 708, 680, 683, 867, 256, 338, 304, ~
## $ text <chr> "InHealthVC AstraZeneca Go Lavanya", "Adapun jenis vaksin yang di~
```

```
corpus<-Corpus(VectorSource(df$text))
corpus
```

```
## <<SimpleCorpus>>
## Metadata:  corpus specific: 1, document level (indexed): 0
## Content:  documents: 1000
```

```r
inspect(corpus[1:10])
```

```
## <<SimpleCorpus>>
## Metadata:  corpus specific: 1, document level (indexed): 0
## Content:  documents: 10
##
##  [1] InHealthVC AstraZeneca Go Lavanya
##  [2] Adapun jenis vaksin yang disuntikan adalah Moderna AstraZeneca dan Coronavac
##  [3]  jenis vaksin dos penggalak disyorkan\nComirnaty Pfizer\nAstraZeneca\nCoronaVac Sinovac\n\nInfo
##  [4]  Warga Diminta Tak Ragu dengan Vaksin Pfizer dan AstraZeneca Menkes Pastikan Aman\nProkes Tangka
##  [5] RT bernamadotcom TWG syor pemberian vaksin Comirnaty AstraZeneca CoronaVac sebagai dos penggalal
##  [6] so dos penggalak ni aku boleh dapat paling awal bulan  huhu \n\nsekurangkurangnya enam bulan se
##  [7] RT TedHilbert Kematian akibat vaksin covid terbukti oleh otopsi StopPaksaVaksin UmatDukungReuni
##  [8] Hallo jangan lupa ya daftar Vaksinasi dosis  di Unikama pendaftaran mulai dari  November \n\nDa
##  [9] Emang sih Aku setelah vaksin astrazeneca malah berasa imun tubuh berkurang dikit dikit cape kura
## [10] RT restasukabumi Assalamualaikum Wr Wb\n\nSelamat siang warga Sukabumi sekitarnya\n\nGerai Vaksi
```

```r
corpus.clean<- corpus %>%
  tm_map(content_transformer(tolower))%>%
  tm_map(removePunctuation)%>%
  tm_map(removeNumbers)%>%
  tm_map(removeWords,stopwords(kind = "en"))%>%
  tm_map(stripWhitespace)
```

```
## Warning in tm_map.SimpleCorpus(., content_transformer(tolower)): transformation
## drops documents
```

```
## Warning in tm_map.SimpleCorpus(., removePunctuation): transformation drops
## documents
```

```
## Warning in tm_map.SimpleCorpus(., removeNumbers): transformation drops documents
```

```
## Warning in tm_map.SimpleCorpus(., removeWords, stopwords(kind = "en")):
## transformation drops documents
```

```
## Warning in tm_map.SimpleCorpus(., stripWhitespace): transformation drops
## documents
```

```r
dtm<-DocumentTermMatrix(corpus.clean)

inspect(dtm[1:10,1:20])
```

```
## <<DocumentTermMatrix (documents: 10, terms: 20)>>
## Non-/sparse entries: 41/159
## Sparsity           : 80%
## Maximal term length: 19
## Weighting          : term frequency (tf)
## Sample             :
##     Terms
## Docs astrazeneca comirnaty coronavac dan dos inhealthvc jenis penggalak pfizer
```

```
##   1            1         0         0   0   0          1     0         0      0
##   10           0         0         0   0   0          0     0         0      0
##   2            1         0         1   1   0          0     1         0      0
##   3            1         1         1   0   1          0     1         1      1
##   4            1         0         0   1   0          0     0         0      1
##   5            1         1         1   0   1          0     0         1      0
##   6            0         0         0   0   2          0     0         1      0
##   7            0         0         0   0   0          0     0         0      0
##   8            0         0         0   0   0          0     0         0      0
##   9            1         0         0   0   0          0     0         0      0
##     Terms
## Docs vaksin
##   1       0
##   10      1
##   2       1
##   3       1
##   4       1
##   5       1
##   6       0
##   7       1
##   8       0
##   9       1
```

```
df.train<-df[1:50,]
df.test<-df[51:100,]

dtm.train<-dtm[1:50,]
dtm.test<-dtm[51:100,]

corpus.clean.train<-corpus.clean[1:50]
corpus.clean.test<-corpus.clean[51:100]

dim(dtm.train)
```

```
## [1]   50 2515
```

```
fivefreq<-findFreqTerms(dtm.train,5)
length(fivefreq)
```

```
## [1] 13
```

```
dtm.train.nb <-DocumentTermMatrix (corpus.clean.test,control = list(dictionary=fivefreq))
dim(dtm.train.nb)
```

```
## [1] 50 13
```

```
dtm.test.nb <-DocumentTermMatrix (corpus.clean.test,control = list(dictionary=fivefreq))
dim(dtm.test.nb)
```

```
## [1] 50 13
```

```
convert_count<-function(x){
    y <-ifelse(x>0,1,0)
    y <-factor(y,levels=c(0,1),labels=c ("no","yes"))
    y
}
trainNB<-apply(dtm.train.nb, 2, convert_count)
testNB<-apply(dtm.test.nb, 1, convert_count)
library(wordcloud)
```

```
## Warning: package 'wordcloud' was built under R version 4.1.2
```

```
## Loading required package: RColorBrewer
```

```
wordcloud(corpus.clean,min.freq = 4,max.words=100,random.order=F,colors=brewer.pal (8,"Dark2"))
```

```
## Warning in strwidth(words[i], cex = size[i], ...): conversion failure on
## 'pedazitosdemí' in 'mbcsToSbcs': dot substituted for <ed>
```

```
## Warning in text.default(x1, y1, words[i], cex = size[i], offset = 0, srt
## = rotWord * : conversion failure on 'pedazitosdemí' in 'mbcsToSbcs': dot
## substituted for <ed>
```

```
## Warning in strwidth(words[i], cex = size[i], ...): conversion failure on
## 'rosalíasotés' in 'mbcsToSbcs': dot substituted for <ed>
```

```
## Warning in strwidth(words[i], cex = size[i], ...): conversion failure on
## 'rosalíasotés' in 'mbcsToSbcs': dot substituted for <e9>
```

```
## Warning in text.default(x1, y1, words[i], cex = size[i], offset = 0, srt
## = rotWord * : conversion failure on 'rosalíasotés' in 'mbcsToSbcs': dot
## substituted for <ed>
```

```
## Warning in text.default(x1, y1, words[i], cex = size[i], offset = 0, srt
## = rotWord * : conversion failure on 'rosalíasotés' in 'mbcsToSbcs': dot
## substituted for <e9>
```


ukassarawak
vaccinespedazitosdem.
langsungapresiasisebagaisetelah
ahyragusatubagijampjpenemu
termasukbelummasihoxfordastrazeneca
banyakkedatanganarabdarahxiv
pimpinandosjutapfizerkembalicarina
prokesinfosaudisakitdengan
astrazeneca
jenis
vaksin
menkes
dose
olehmoderna
comirnaty
tibaakan
ampmereka
globally
disuntikpemerintah
diminta
wargacoronavac
yangdandosis
joeatauterimahari
aman
itu
lagi
indonesia
karenasemuauntukasal
lebihyaitucovidpenerimadiakuimpr
penggalakbisasaatmenerimanamun
barengtahap.umrahiniadavarian
pastikanjemaahsinovaktakdalam
rumahsebanyakvaksinasiomicrontangkal
usedkumparancon
demokratharapanrakyat
astrazenecavacanaotaegui