

LAMPIRAN

Lampiran 1. *Packages* yang digunakan

```
library(devtools)
library(plyr)
library(tm)
library(wordcloud)
library(RColorBrewer)
library(ggplot2)
library(e1071)
library(RTextTools)
```

Lampiran 2. Sintaks R Preprocessing Data dengan Text Mining

```
#Import data tweet
datashopee <- read.csv("d:/SKRIPSI
RIRI/data2klasifikasi.csv", sep=";")
class(datashopee)
datashopeeCorpus <-
Corpus(VectorSource(datashopee[,2]))
class(datashopeeCorpus)
inspect(datashopeeCorpus[1])

#Preprocessing
#Menghapus URL, rt, username
removeURL <- function(x) gsub("http[^[:space:]]*", "",
x)
datashopee_clean <- tm_map(datashopeeCorpus, removeURL)
removeRT <- function(y) gsub("RT ", "", y)
datashopee_clean <- tm_map(datashopee_clean, removeRT)
removeUN <- function(z) gsub("@\\w+", "", z)
datashopee_clean <- tm_map(datashopee_clean, removeUN)
#Remove punctuation, to lower
datashopee_clean <- tm_map(datashopee_clean,
removePunctuation)
datashopee_clean <- tm_map(datashopee_clean, tolower)

#Menghapus stopwords
file_stop <- file("d:/SKRIPSI RIRI/stopword.txt", open =
"r")
id_stopwords <- readLines(file_stop)
close(file_stop)
id_stopwords = c(id_stopwords, "amp")
datashopee_clean <- tm_map(datashopee_clean,
removeWords, id_stopwords)
#Menghapus nomor, strip white space
datashopee_clean <- tm_map(datashopee_clean,
removeNumbers)
datashopee_clean <- tm_map(datashopee_clean,
stripWhitespace)
datashopee_clean <- tm_map(datashopee_clean,
PlainTextDocument)
inspect(datashopee_clean[1:5])

datashopee_clean <- tm_map(datashopee_clean,
removeWords,
c("shopee", "shopeeid", "tolong", "min", "admin", "gak", "nya
"))
```

```
#TOKENIZING#
tokenizing <- function(x) strsplit(as.character(x),
";")
shopee_clean <- tm_map(datashopee_clean, tokenizing)
shopee_clean[[1]]

#Stemming
stem_text <- function(text,mc.cores=1) {
  #stem each word in a block of text
  stem_string <- function(str)
  {
    str <- tokenize(x=str)
    str <- sapply(str, katadasaR)
    str <- paste(str, collapse="")
    return(str)
  }
  #stem each text block in turn
  x <- mclapply(X=text, FUN=stem_string,
mc.cores=mc.cores)
  #return stemmed text blocks
  return(unlist(x))
}
shopee_clean <- tm_map(shopee_clean, stem_text)
class(shopee_clean)
```

Lampiran 3. *Build Term Document Matrix and Document Term Matrix*

```
#Build TDM and DTM
shopee_clean <- tm_map(shopee_clean, PlainTextDocument)
dtm <- DocumentTermMatrix(shopee_clean)
inspect(dtm[1:5, 20:25]) #inspeksi matriks baris, kolom
findFreqTerms(dtm, 5) #mencari kata yang mempunyai
frekuensi x
tdm <- TermDocumentMatrix(shopee_clean)
inspect(tdm[1:5, 20:25])

shopee_clean[[1]]
```

Lampiran 4. Classification, Wordcloud, dan Association

```
#####Klasifikasi SVM#####
container <- create_container(dtm,
as.numeric(as.factor(datashopee[,1])), trainSize=1:192,
testSize=193:368, virgin=FALSE)
# train a SVM Model
models <- train_models(container, algorithms="SVM")
results <- classify_models(container, models)
SVM_table =
table(as.numeric(as.factor(datashopee[193:368, 1])),
results[, "SVM_LABEL"])
SVM_accuracy =
recall_accuracy(as.numeric(as.factor(datashopee[193:368
, 1])), results[, "SVM_LABEL"])

#Wordcloud
wordcloud(datashopee_clean, random.order=F, min.freq=10,
colors=brewer.pal(8, "Dark2"))

#ASSOCIATION#
v<-as.list(findAssocs(dtm, terms =c("barang",),
corlimit =
c(0.15,0.15,0.15,0.15,0.15,0.15,0.15,0.15,0.15,0.15)))
v
```