

**LAPORAN PROYEK AKHIR  
PRAKTIKUM DATA SCIENCE**

**ANALISIS SENTIMEN MASYARAKAT  
TERHADAP VAKSINASI COVID-19 DI  
INDONESIA PADA MEDIA SOSIAL TWITTER**



RADEN RARA LYDIA D.S	123190086
LISA ANIS SAFITRI	123190089

**PROGRAM STUDI INFORMATIKA  
JURUSAN TEKNIK INFORMATIKA  
FAKULTAS TEKNIK INDUSTRI  
UNIVERSITAS PEMBANGUNAN NASIONAL “VETERAN”  
YOGYAKARTA  
2021**

## 1. PENDAHULUAN

COVID-19 (*Coronavirus Disease-2019*) merupakan wabah penyakit menular yang berbahaya karena tingkat penyebarannya sangat cepat dan meluas di seluruh dunia terutama di Indonesia. Berdasarkan data WHO (World Health Organization) hingga tanggal 27 Oktober 2021, tercatat 259.502.031 kasus COVID-19 yang terkonfirmasi dan terdapat 5.183.003 kasus kematian di dunia. Sementara itu, Indonesia sendiri menempati posisi ke-14 sebagai kasus COVID-19 terbanyak di dunia dengan jumlah kasus yang terkonfirmasi sebanyak 4.255.268 dan kasus kematian sebanyak 143.796 jiwa [1].

Pemerintah menerapkan berbagai upaya untuk meminimalisir penyebaran COVID-19 di Indonesia, salah satunya yaitu tindakan vaksinasi. Ada beberapa jenis vaksin yang digunakan oleh Pemerintah Indonesia, yaitu vaksin Sinovac, vaksin AstraZeneca, vaksin Moderna, vaksin Pfizer, vaksin Bio Farma, vaksin Sinopharm, dan beberapa jenis vaksin lainnya. Tindakan vaksinasi pada masyarakat Indonesia menimbulkan banyak tanggapan dan pendapat dari berbagai kalangan. Tanggapan dari masyarakat Indonesia sendiri terkait adanya tindakan vaksinasi yaitu banyak tanggapan positif dan membangun, tetapi ada pula tanggapan yang negatif dan menentang terkait adanya tindakan vaksinasi.

Twitter sebagai salah satu platform media sosial bagi masyarakat untuk mencurahkan dan menyampaikan pendapatnya. Twitter adalah salah satu media sosial yang banyak digunakan oleh masyarakat Indonesia. Pengguna Twitter dapat membagikan kicauan (tweet) berupa tulisan, foto, video kepada publik. Masyarakat dengan mudah memberikan pendapat dan opininya ke media sosial Twitter melalui tweet tersebut. Demikian pula dengan tindakan vaksinasi, banyak pengguna Twitter yang menyampaikan tanggapan dan pendapatnya mengenai pemberlakuan vaksinasi di Indonesia. Ada beberapa pengguna Twitter yang mendukung tindakan tersebut, adapula pengguna Twitter yang menolak dan tidak setuju terkait tindakan vaksinasi.

Terdapat beberapa metode yang dapat digunakan untuk menganalisis pendapat masyarakat Indonesia terkait adanya tindakan vaksinasi di media sosial Twitter, salah satunya yaitu analisis sentimen. Analisis Sentimen yaitu suatu proses mengekstraksi dan mengolah data tekstual untuk mendapatkan informasi, mengidentifikasi kecenderungan hal, memperoleh gambaran umum persepsi masyarakat terhadap suatu pelayanan atau tindakan [2].

Pada penelitian ini, dilakukan analisis sentimen terhadap opini masyarakat Indonesia terkait tindakan vaksinasi pada media sosial Twitter. Pengumpulan data pada penelitian ini dilakukan dengan pengambilan data set dari Twitter langsung menggunakan cara crawling data. Data set tersebut adalah kicauan (tweet) dari masyarakat Indonesia yang menyampaikan pendapatnya mengenai vaksinasi di Indonesia. Data set tersebut selanjutnya dikelompokkan dengan label positif, negatif, dan netral dengan metode Naive Bayes Classifier. Pada akhirnya, tujuan dari

penelitian analisis sentimen ini yaitu untuk menentukan apakah tindakan vaksinasi di Indonesia diterima oleh masyarakat atau tidak.

## 2. METODE

Peneliti menggunakan beberapa langkah dalam mengambil, mengolah, dan menyimpulkan data tweet masyarakat Indonesia terkait dengan kegiatan vaksinasi Covid19. Adapun metode penelitian yang dilakukan adalah crawling data tweet, *preprocessing* data, pengolahan data, dan penampilan hasil pengolahan data. Penelitian dilakukan dengan menggunakan bahasa R sebagai bahasa pengkodean dan kegiatan pengolahan data dilakukan menggunakan aplikasi R Studio.

### A. Crawling data melalui media sosial Twitter

Crawling data adalah suatu langkah untuk mengambil data yang dibutuhkan. Penelitian ini menggunakan menggunakan library twitter, rtweet, tm, tidyverse, vroom, ggplot2, plyr, shiny, wordcloud, dan RColorBrewer. Crawling data melalui media sosial Twitter didahului dengan menghubungkan RStudio dengan API Twitter. Autentikasi Twitter dilakukan dengan menghubungkan RStudio dengan API Twitter menggunakan API key, *password* API, token akses, dan token akses rahasia.

Setelah melakukan autentikasi Twitter dengan RStudio, dilakukan proses crawling data. Crawling data dilakukan dengan melakukan pencarian tweet menggunakan kata kunci. Kata kunci yang digunakan dalam penelitian ini adalah vaksinasi masyarakat. Tweet yang digunakan sebagai data bahan penelitian ini dibatasi dengan jumlah 3000 tweet dengan bahasa Indonesia.

### B. Preprocessing Data

Data tweet yang sudah diambil kemudian dilakukan beberapa tahap penghapusan karakter yang tidak digunakan. Adapun karakter yang dihapus dari tweet adalah alamat url, spasi perpindahan baris, tanda baca: koma; titik dua; titik koma. Selanjutnya dilakukan penghapusan terhadap titik tiga pada tweet, kata 'RT', karakter ampersand, username, spasi, tanda baca lain. Setelah dilakukan penghapusan karakter tweet, selanjutnya tweet diubah sehingga seluruh isi tweet tersimpan dengan huruf kecil. Data tweet yang sudah mengalami tahap preprocessing disimpan ke dalam dataframe. Data frame yang sudah terbuat selanjutnya disimpan dalam format csv sebagai data jadi.

### C. Data Processing

Data jadi hasil preprocessing kemudian diolah dalam beberapa tahapan. Adapun tahapan yang dilakukan, yaitu sentimen analisis menggunakan logika Naive Bayes, klasifikasi nilai sentimen, pembuatan word cloud, dan penampilan barplot kata yang sering muncul.

#### D. Menampilkan Hasil

Hasil data processing ditampilkan menggunakan aplikasi web sederhana. Hasil pemrosesan data ditampilkan menggunakan library Shiny. Data yang ditampilkan berupa tabel sumber data, barplot yang menampilkan sentimen analisis, barplot frekuensi kata yang sering muncul, dan word cloud dari kata yang sering muncul.

### 3. HASIL DAN PEMBAHASAN

#### A. Menyiapkan Library

```
```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```
```

**Tabel 1** Kode Pemrograman Peng-*install*-an *Set Up R Markdown*

Potongan kode pertama di bawah ini digunakan untuk menentukan pengaturan global apapun yang akan diterapkan ke script R Markdown. Contohnya menetapkan seluruh potongan kode chunks sebagai “echo = TRUE”, yang berarti akan disertakan pada versi final yang dirender.

```
```{r}
library(tidyverse)
library(twitteR)
library(rtweet)
library(tm)
library(ggplot2)
library(plyr)
library(shiny)
library(wordcloud) #library untuk word-cloud
library(RColorBrewer) #library untuk menggunakan palet warna
```
```

**Tabel 2** Kode Pemrograman Pengaktifan *Library Packages*

Sebelum memulai tahapan pemrosesan dan analisis data diperlukan penginstalan packages dan mengaktifkannya dengan mengetik script library(...). Pada penelitian ini digunakan beberapa packages yaitu:

1. library(tidyverse) berfungsi untuk melakukan pengolahan data pada aplikasi R seperti import, export, visualisasi, dan pemodelan data.
2. library(twitteR) digunakan untuk mendapatkan (crawling) data dari Twitter dengan R untuk berinteraksi dengan REST API.
3. library(rtweet) berfungsi untuk mengakses API Twitter melalui R.
4. library(tm) digunakan untuk cleaning data text di RStudio
5. library(ggplot2) berfungsi membuat grafik yang merepresentasikan dan memvisualisasikan data dengan menerapkan warna, simbol, dan lainnya.
6. library(plyr) berfungsi untuk memanipulasi data.
7. library(shiny) paket dalam R yang digunakan untuk membangun web apps yang interaktif.

8. library(wordcloud) adalah library untuk word-cloud
9. library(RColorBrewer) adalah package yang digunakan untuk menggunakan palet warna

## B. Proses Crawling Data Twitter

```
```{r}
## Crawling Data Twitter menggunakan API
api_key<- "D5zz4dTl6xjWxi5PfwWPBVPy7"
api_secret<-
"4vzip8PPUbXftEJcdTOdENygdK7Eqgnkp5T4px2bw1Ma7dc5oj"
access_token<-
"1060059384-0RmpWSJXqYXU5aXZ1Qshr7OG41U4Mth37ehKfgW"
access_token_secret<-
"E7F8j8tRh4ASFJOeOuefzOFD5ejIjSIidibRXQ2aCBsHjI"
setup_twitter_oauth(api_key,api_secret,access_token,access_token_secret)

tweet = searchTwitter('vaksinasi covid',
                      n = 3000,
                      retryOnRateLimit = 10e5, lang = "id")
saveRDS(tweet,file = 'tweetVaksinasi.rds')
```
```

**Tabel 3** Kode Pemrograman Crawling Data Twitter

Crawling data di Twitter adalah proses mengambil data dari server Twitter dengan bantuan Application Programming Interface (API) Twitter. Untuk mengakses API Twitter diperlukan akun Twitter, aplikasi (application) dan token akses (access token). Dalam membuat aplikasi harus menggunakan *developer account*. Proses selanjutnya yaitu otentikasi (*authentication*) agar R melalui rtweet bisa berinteraksi dengan API Twitter. Setelah proses otentikasi berhasil, dapat dilakukan penambangan data (crawling) dari Twitter. Crawling data dilakukan dengan melakukan pencarian tweet menggunakan kata kunci “vaksinasi masyarakat”. Tweet yang digunakan sebagai penelitian ini dibatasi dengan jumlah 3000 tweet dengan bahasa Indonesia dan disimpan pada tweetVaksinasi.rds

## C. Data Cleansing

```
```{r load dataset}
tweet <- readRDS('tweetVaksinasi.rds')
datatweet = twListToDF(tweet) #convert twitter list to data

tweetVaksinasi <- datatweet$text
tweetVaksinasi_c <- Corpus(VectorSource(tweetVaksinasi))

##menghapus URL
removeURL <- function(x) gsub("http[^[:space:]]*", "", x)
tweetclean <- tm_map(tweetVaksinasi_c, removeURL)

##menghapus NewLine
```

```

removeNL <- function(y) gsub("\n", " ", y)
tweetclean <- tm_map(tweetclean, removeNL)

##menghapus koma
replacecomma <- function(y) gsub(",", "", y)
tweetclean <- tm_map(tweetclean, replacecomma)

##menghapus titik2
removetitik2 <- function(y) gsub(":", "", y)
tweetclean <- tm_map(tweetclean, removetitik2)

##menghapus titik koma
removetitikkoma <- function(y) gsub(";", " ", y)
tweetclean <- tm_map(tweetclean, removetitikkoma)

##menghapus titik3
removetitik3 <- function(y) gsub("p..", "", y)
tweetclean <- tm_map(tweetclean, removetitik3)

##menghapus RT(retweet)
removeRT <- function(y) gsub("RT", "", y)
tweetclean <- tm_map(tweetclean, removeRT)

##menghapus &
removeamp <- function(y) gsub("&", "", y)
tweetclean <- tm_map(tweetclean, removeamp)

##menghapus username
removeUN <- function(z) gsub("@\\w+", "", z)
tweetclean <- tm_map(tweetclean, removeUN)

##menghapus space dan lainnya
remove.all <- function(xy) gsub("[^[:alpha:][:space:]]*", "", xy)
tweetclean <- tm_map(tweetclean, remove.all)
##menghapus tanda baca
tweetclean <- tm_map(tweetclean, removePunctuation)
##mengubah ke huruf kecil
tweetclean <- tm_map(tweetclean, tolower)

myStopwords = readLines("stopwords-id.txt")

tweetclean <- tm_map(tweetclean, removeWords, myStopwords)
```

```

**Tabel 4** Kode Pemrograman Data Cleansing

Data hasil *crawling* Twitter yang digunakan pada penelitian ini sebelumnya diolah dengan melakukan *data cleansing*, yaitu penghapusan karakter yang dianggap tidak dibutuhkan dalam kegiatan penelitian. Data yang akan dibersihkan, sebelumnya disimpan sebagai corpus. Proses *data cleansing* pada data yang digunakan adalah sebagai berikut.

1. Penghapusan URL.
2. Penghapusan Baris Baru.
3. Penghapusan Koma.

4. Penghapusan Titik Dua.
5. Penghapusan Titik Koma.
6. Penghapusan Titik Tiga.
7. Penghapusan RT (*Retweet*).
8. Penghapusan Tanda Ampersand.
9. Penghapusan Username.
10. Penghapusan Spasi.
11. Penghapusan Tanda Baca.
12. Pengubahan *Tweet* menjadi Huruf Kecil.

```
```{r}
##Menyimpan data yang sudah bersih ke dataframe dan disimpan
pada .csv
dataframe<-data.frame(text=unlist(sapply(tweetclean, `[`)),
stringsAsFactors=F)

write.csv(dataframe,file = 'tweetclean_Vaksinasi.csv')
```
```

**Tabel 5** Kode Pemrograman Penyimpanan Data Clean

Setelah dilakukan *data cleansing*, data yang sudah bersih atau *data clean* disimpan ke dalam *data frame*. Selanjutnya, data frame diubah menjadi format *Comma Separated Value* (CSV) agar lebih mudah diolah.

#### D. Data Processing

```
```{r}
## Naive Bayes
library(e1071) #Library untuk Naive Bayes
library(caret) #Library untuk Klasifikasi Data
library(syuzhet) #Library untuk membaca fungsi get_nrc
(sentiment analysis)

#Membaca file csv yang sudah menjalani proses cleaning data
vaksinasi_data
<-read.csv("tweetclean_Vaksinasi.csv",stringsAsFactors =
FALSE)

#Mengubah text menjadi char
tweets <- as.character(vaksinasi_data$text)

#Memanggil sentimen dictionary untuk menghitung presentasi
dari beberapa emotion dan mengubahnya ke dalam text file
#get_nrc_sentiment('happy')
#get_nrc_sentiment('excitement')
sentimen<-get_nrc_sentiment(tweets)

tweet_sentiment<-cbind(vaksinasi_data$text,sentimen)
par(mar=rep(3,4))
barplot(
  colSums(sentimen),
  col=rainbow(10),
  ylab='count',
```

```

main='Sentiment Analysis'
)
```

```

**Tabel 6** Kode Pemrograman Analisis Sentimen dengan Naive Bayes

Proses selanjutnya yaitu analisis sentimen menggunakan metode Naive Bayes. Digunakan beberapa library penting untuk analisis sentimen menggunakan Naive Bayes yaitu library(caret) untuk klasifikasi data dan library(syuzhet) membaca fungsi get\_nrc (sentiment analysis). Selanjutnya Membaca file csv yang sudah menjalani proses cleaning data dan mengubahnya dari format text menjadi char. Selanjutnya memanggil sentimen dictionary untuk menghitung persentase dari beberapa emoticon dan mengubahnya ke dalam text file. Fungsi get\_nrc\_sentimen berguna untuk mengekstrak sentimen dari Twitter. Fungsi ini bekerja dengan memanggil kamus sentimen NRC untuk menghitung adanya emosi yang berbeda dan valensi yang sesuai dalam file text. Selanjutnya analisis sentimen ditampilkan dengan barplot.

```

```{r}
## Mengklasifikasikan sentimen positif, negatif, dan neutral
class_sentiment <-
data.frame(negative=sentimen$negative,positive=sentimen$positi
ve)
klasifikasi <- mutate(class_sentiment, text_sentiment =
ifelse((class_sentiment$negative != class_sentiment$positive),
ifelse(class_sentiment$negative!=0,print("negative"),
print("positive")),print("neutral")))

data_vaksinasi <-
data.frame(text=tweets,sentimen=klasifikasi$text_sentiment)

```

```

**Tabel 7** Kode Pemrograman Pengklasifikasian Sentimen

Data frame hasil ekstraksi sentimen Twitter selanjutnya dimasukkan ke dalam sebuah variabel dengan masukkan data berupa data sentimen negatif dan sentimen positif. Selanjutnya dilakukan klasifikasi data menggunakan fungsi 'mutate' pada variabel 'class\_sentiment'. Hasil klasifikasi sentimen selanjutnya disimpan pada variabel 'data\_vaksinasi'.

```

```{r}
#Menampilkan barplot representasi 10 kata yang sering muncul
data_vaksin = as.factor(vaksinasi_data$text)
corpus = Corpus(VectorSource(data_vaksin))
tdm <- TermDocumentMatrix(corpus)

```



```
tdm <- removeSparseTerms(tdm, sparse = 0.98)
tdm <- as.matrix(tdm)

w = sort(rowSums(tdm), decreasing = T)
barplot(w[1:5],
        las=2,
        main = "Frequency of Words",
        col= rainbow(15))
```

```

**Tabel 8** Kode Pemrograman Menampilkan Kata yang Sering Muncul menggunakan Bar Plot

Hasil data clean disimpan sebagai *faktor* menggunakan fungsi '*as.factor*'. Fungsi '*as.factor*' berguna untuk mengubah kolom data numerik atau data frame menjadi bentuk faktor. Data berbentuk faktor selanjutnya disimpan sebagai '*corpus*'. Data pada corpus selanjutnya diolah dengan *function* '*TermDocumentMatrix*' dalam mencari frekuensi kemunculan kata dan disimpan sebagai matrix. Matrix hasil diurutkan dengan fungsi *sort* dan ditampilkan dengan menggunakan *barplot*.

```
```{r}
#Menampilkan kemunculan kata yang sama
getwordcloud <- function(x){
  set.seed(12345)
  wordcloud(words = x$word, freq = x$freq, min.freq = 3,
            max.words=100, random.order=FALSE, rot.per=0.40,
            colors=brewer.pal(8, "Dark2"))
}
getwordcloud(freq_words)
```

```

**Tabel 9** Kode Pemrograman Menampilkan Kata yang Sering Muncul menggunakan Get Word Cloud

Penampilan hasil frekuensi kemunculan kata juga dilakukan dengan menggunakan *function* '*WordCloud*'. *Function* '*WordCloud*' mengolah setiap kata pada data dan mencari jumlah kemunculan kata dengan batasan minimal kemunculan sejumlah tiga dan jumlah kata maksimal yaitu 100.

```
```{r ui}
library(markdown)
library(DT)

ui <- fluidPage(
  titlePanel("Tweets Vaksinasi Covid-19 Masyarakat
Indonesia"), #halaman judul
  mainPanel(
    tabsetPanel(type = "tabs",
      tabPanel("Data Twitter",
DT::dataTableOutput('data')),
      tabPanel("Sentiment Analysis",

```

```

DT::dataTableOutput('sentiment')),
      tabPanel("Scatterplot",
plotOutput('scatterplot')),
      tabPanel("Frequency Words",
plotOutput('freqword')),
      tabPanel("Wordcloud", plotOutput('wordcloud'))
    )
  )
)
```

```

**Tabel 10** Kode Pemrograman UI Web Shiny

Shiny adalah sebuah *package* dalam R yang dapat membangun sebuah web apps yang interaktif. Shiny menggabungkan antara kekuatan komputasi statistika R dan interaksinya dengan web modern. Pada proyek penelitian ini Shiny web apps menampilkan hasil analisis sentimen terkait vaksinasi masyarakat di Indonesia berupa data mentah twitter, data sentiment analysis, scatterplot, frequency words, dan worcloud. Shiny dibagi menjadi 3 komponen yaitu UI (User Interface), Server, dan ShinyApp. Pada kode program diatas merupakan bagian code shiny dari *user interface* yang mendefinisikan tampilan web dari aplikasi. Fungsi UI adalah untuk memuat seluruh input dan output yang akan ditampilkan dalam app.

```

```{r server}
server <- function(input, output) {
  output$data <- DT::renderDataTable({
    DT::datatable(vaksinasi_data, options = list(lengthChange
= FALSE))
  })

  output$sentiment <- DT::renderDataTable({
    DT::datatable(data_vaksinasi, options = list(lengthChange
= FALSE))
  })

  output$scatterplot <- renderPlot({
    barplot(
      colSums(sentimen),
      col=rainbow(10),
      ylab='count',
      main='Sentiment Analysis'
    )
  })

  output$freqword<- renderPlot({
    barplot(w[1:5],
      las=2,
      main = "Frequency of Words",
      col= rainbow(15))
  })
}
```

```

```

output$wordcloud<- renderPlot({
  getwordcloud(freq_words)
})

}
```

```

**Tabel 11** Kode Pemrograman Server Web Shiny

Pada kode program diatas merupakan bagian code shiny dari *server* yang mendefinisikan logika kerja analisis dari sisi server pada aplikasi.

```

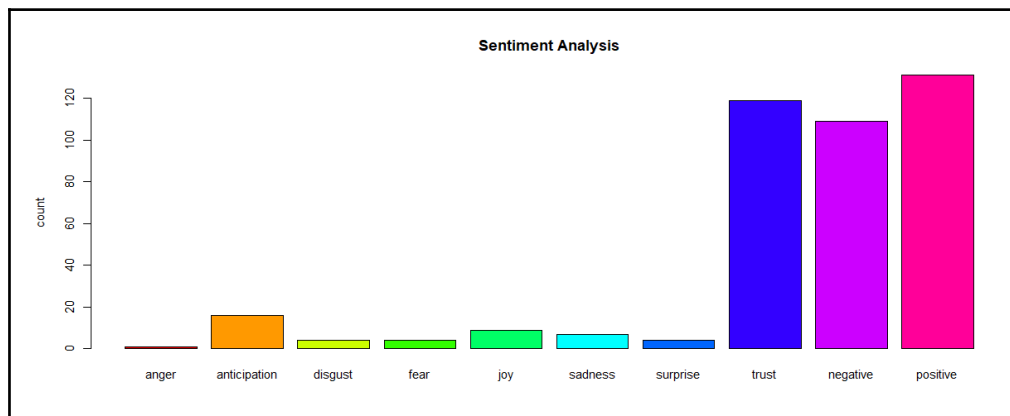
```{r run-opp}
shinyApp(ui = ui, server = server)
```

```

**Tabel 12** Kode Pemrograman Menampilkan Web Shiny

Pada kode program diatas merupakan bagian code shiny dari ShinyApp yang merupakan fungsi dari aplikasi yang memanggil UI dan Server untuk menjalankan aplikasi.

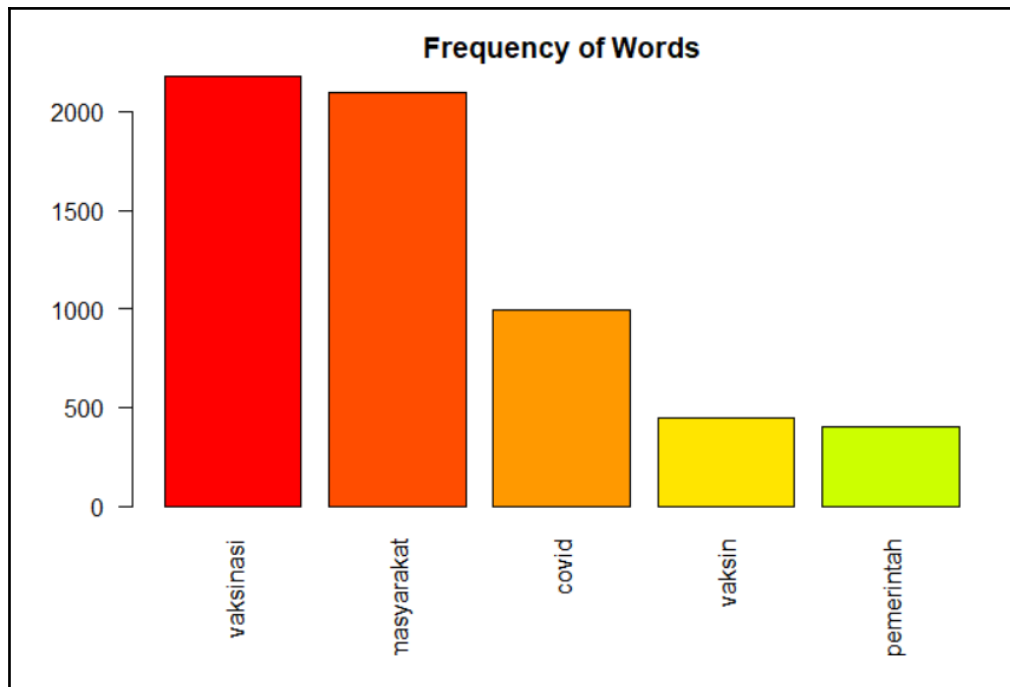
#### E. Hasil Kode Pemrograman



**Gambar 1** Tampilan Scatterplot Sentimen

Berdasarkan hasil analisis sentimen menggunakan metode *Naïve Bayes Classifier* dapat ditentukan emosi dibalik serangkaian kata melalui curahan tweet yang diunggah pada Twitter. Didapatkan hasil sentimen yang diklasifikasikan menjadi beberapa kategori yaitu emosi *anger*, *anticipation*, *disgust*, *fear*, *joy*, *sadness*, *surprise*, *trust*, *negative*, dan *positive*.

Twitter Sentiment Analysis yang menentukan emosi terkait pendapat masyarakat tentang vaksinasi Covid-19 di Indonesia divisualisasikan ke dalam bentuk scatterplot. Analisis sentimen kategori *positive* menduduki posisi tertinggi dari kategori lainnya. Hal ini menunjukkan bahwa sikap, pendapat, dan emosi yang diungkapkan oleh masyarakat pada Twitter terkait vaksinasi Covid-19 di Indonesia mengandung dan memiliki sikap *positive*.



**Gambar 2** Tampilan frekuensi 5 kata paling banyak dengan Barplot

Berdasarkan fungsi sort pada data, didapatkan hasil frekuensi kemunculan kata yang ditampilkan dalam bentuk *barplot*. Barplot frekuensi kemunculan kata merepresentasikan lima tingkat frekuensi kemunculan kata, yaitu ‘vaksinasi’, ‘masyarakat’, ‘covid’, ‘vaksin’, dan ‘pemerintah’.



**Gambar 3** Tampilan frekuensi 5 kata paling banyak dengan *Wordcloud*

Frekuensi kemunculan kata juga ditampilkan dalam bentuk *Wordcloud*. Pada tampilan *Wordcloud*, kata yang paling banyak muncul ditampilkan dengan ukuran yang paling besar dan berwarna gelap. Selanjutnya pada kata keempat dan kelima frekuensi kemunculan kata diberi warna lain agar dapat memudahkan *user* dalam mengetahui kata yang paling sering muncul.

Kata ‘masyarakat’ ditampilkan dengan ukuran paling besar dan berwarna gelap sebagai penanda bahwa kata masyarakat paling sering muncul. Kata kedua yang sering muncul pada *tweet* adalah kata ‘vaksinasi’ yang memiliki ukuran besar, tetapi tidak lebih besar dari kata ‘masyarakat’ dan memiliki warna gelap. Kata ‘covid’ ditampilkan dengan warna gelap dan

ukuran besar, tetapi tidak lebih gelap dan tidak lebih besar dibandingkan dengan kata ‘masyarakat’ dan ‘vaksinasi’. Kata ‘pemerintah’ dan ‘vaksin’ ditampilkan dengan menggunakan warna oranye sebagai pembeda dan ukuran sedikit lebih besar dibandingkan dengan kata lain yang tidak terdapat dalam lima kata yang paling sering muncul dalam *tweet*.

**Tweets Vaksinasi Covid-19 Masyarakat Indonesia**

Data Twitter   Sentiment Analysis   Scatterplot   Frequency Words   Wordcloud

Search:

|    | X  | text                                                                                                   |
|----|----|--------------------------------------------------------------------------------------------------------|
| 1  | 1  | trakindo menghadirkan sentra vaksinasi menyediakan masingmasing dosis vaksincovidbagi masyarakat m     |
| 2  | 2  | trakindo menghadirkan sentra vaksinasi menyediakan masingmasing dosis vaksincovidbagi masyarakat m     |
| 3  | 3  | pemkot kota bandung targetkan persen vaksinasi vivid masyarakat desember                               |
| 4  | 4  | wonogiri mencapai angka persen cakupan vaksinasi masyarakat lakuka                                     |
| 5  | 5  | indonesia lampau target who capaian vaksinasi covid utk keamanan masyarakat i                          |
| 6  | 6  | yes indonesia kedatangan vaksin astrazeneca pfizer langkah pemerintah wujudkan st                      |
| 7  | 7  | vaksinasi covid capai persen masyarakat diimbau disiplinproses                                         |
| 8  | 8  | percepat program vaksinasi nasional polresta bandar lampung buka gerai vaksin presisi masyarakat melin |
| 9  | 9  | sukseskan vaksinasi covid masyarakat                                                                   |
| 10 | 10 | himbauan masyarakat menjaga keselamatan berkendara ajakan vaksinasi                                    |

Showing 1 to 10 of 3,000 entries

Previous 1 2 3 4 5 ... 300 Next

**Gambar 4** Tampilan Shiny Web Apps Data Twitter

Shiny Web Apps menampilkan hasil analisis sentimen pendapat masyarakat terkait vaksinasi Covid-19 di Indonesia berupa data twitter, data sentiment analysis, scatterplot terkait hasil analisis sentimen emosi dibalik tweet masyarakat Indonesia, frequency words, dan worcloud.

Data tweet yang telah dilakukan proses *cleaning* data ditampilkan pada Shiny Web Apps dengan tabPanel “Data Twitter” dalam bentuk tabel. Data ini memiliki dua *field* yaitu ‘X’ dan ‘text’ yang merupakan data tweet. Tweet masyarakat Indonesia mengenai vaksinasi ini sebanyak 3000 tweet dalam

Bahasa Indonesia. Pada panel ini juga terdapat fitur pencarian data atau *searching* sehingga akan memudahkan pengguna untuk mencari data tweet.

**Tweets Vaksinasi Covid-19 Masyarakat Indonesia**

Search:

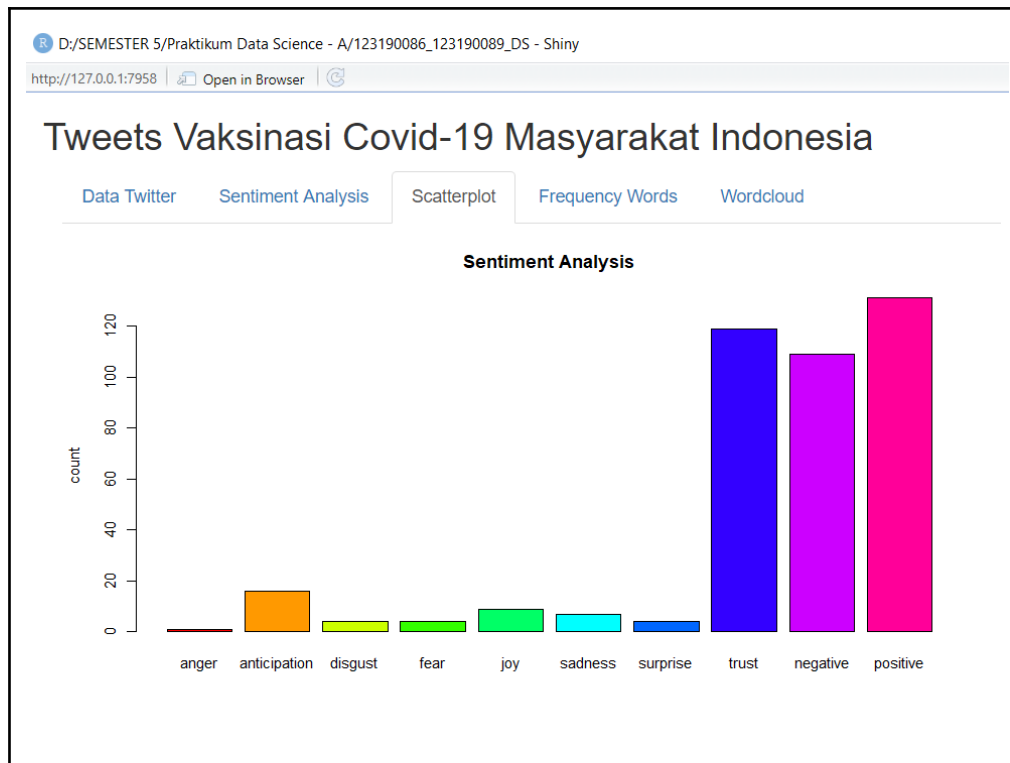
|    | text                                                                                                   | sentimen |
|----|--------------------------------------------------------------------------------------------------------|----------|
| 1  | trakindo menghadirkan sentra vaksinasi menyediakan masingmasing dosis vaksincovidbagi masyarakat m     | neutral  |
| 2  | trakindo menghadirkan sentra vaksinasi menyediakan masingmasing dosis vaksincovidbagi masyarakat m     | neutral  |
| 3  | pemkot kota bandung targetkan persen vaksinasi vivid masyarakat desember                               | positive |
| 4  | wonogiri mencapai angka persen cakupan vaksinasi masyarakat lakuka                                     | neutral  |
| 5  | indonesia lampau target who capaian vaksinasi covid utk keamanan masyarakat i                          | neutral  |
| 6  | yes indonesia kedatangan vaksin astrazeneca pfizer langkah pemerintah wujudkan st                      | neutral  |
| 7  | vaksinasi covid capai persen masyarakat diimbau disiplinprokes                                         | neutral  |
| 8  | percepat program vaksinasi nasional polresta bandar lampung buka gerai vaksin presisi masyarakat melin | neutral  |
| 9  | sukseskan vaksinasi covid masyarakat                                                                   | neutral  |
| 10 | himbauan masyarakat menjaga keselamatan berkendara ajakan vaksinasi                                    | neutral  |

Showing 1 to 10 of 3,000 entries

Previous 1 2 3 4 5 ... 300 Next

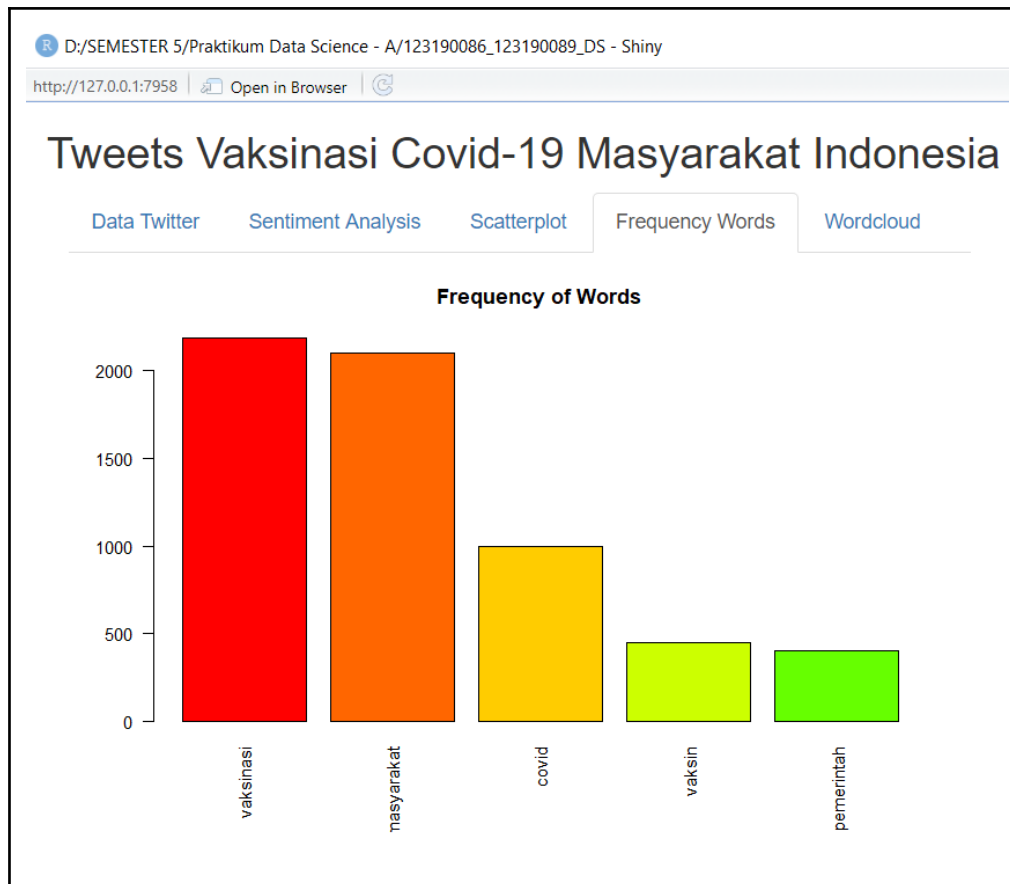
**Gambar 5** Tampilan Shiny Web Apps Sentiment Analysis

Field ‘teks’ dari hasil *crawling data* yang selanjutnya dilakukan pemrosesan *cleaning* data dilakukan pengklasifikasian analisis sentimen *positive*, *negative*, dan *neutral* menggunakan metode *Naïve Bayes Classifier*. Hasil pengklasifikasian sentimen tersebut ditampilkan pada Shiny Web Apps panel “*Sentiment Analysis*” dalam bentuk tabel yang memberikan keterangan sentimen pada setiap teks data tweet masyarakat Indonesia mengenai adanya tindakan vaksinasi.



**Gambar 6** Tampilan Shiny Web Apps Scatterplot

Sentimen diklasifikasikan menjadi beberapa kategori yaitu emosi *anger*, *anticipation*, *disgust*, *fear*, *joy*, *sadness*, *surprise*, *trust*, *negative*, dan *positive*. Twitter Sentiment Analysis yang menentukan emosi terkait pendapat masyarakat tentang vaksinasi Covid-19 di Indonesia divisualisasikan ke dalam bentuk scatterplot dan ditampilkan pada Shiny Web Apps panel “*Scatterplot*”. Dapat dilihat bahwa analisis sentimen kategori *positive* menduduki posisi tertinggi dari kategori lainnya. Hal ini menunjukkan bahwa sikap, pendapat, dan emosi yang diungkapkan oleh masyarakat di media sosial Twitter tentang vaksinasi Covid-19 di Indonesia memiliki sikap *positive*.



**Gambar 7** Tampilan Shiny Web Apps Frequency Words

Berdasarkan fungsi sort pada data, didapatkan hasil frekuensi kemunculan kata yang ditampilkan dalam bentuk *barplot*. Barplot frekuensi kemunculan kata merepresentasikan lima tingkat frekuensi kemunculan kata, yaitu ‘vaksinasi’, ‘masyarakat’, ‘covid’, ‘vaksin’, dan ‘pemerintah’. Hasil visualisasi tersebut ditampilkan pada Shiny Web Apps panel “*Frequency Words*”. Dapat dilihat bahwa dari 5 kata paling sering muncul pada tweets, kemunculan kata terbanyak yaitu kata ‘vaksinasi’.





dengan menggunakan *Web ShinyApp*. Program ini berhasil membagi data *tweet* berdasarkan hasil sentimen menjadi sentimen *positive*, *negative*, ataupun *neutral*. Berdasarkan analisis, dapat diamati bahwa sebagian besar *tweet* menunjukkan sikap, pendapat, dan emosi yang diungkapkan oleh masyarakat tentang vaksinasi Covid-19 di Indonesia mengandung respon *positive*.

Penelitian yang dilakukan pada proyek Analisis Sentimen Masyarakat Terhadap Vaksinasi Covid-19 di Indonesia Pada Media Sosial *Twitter* juga menampilkan kata yang muncul dengan frekuensi terbanyak menggunakan *function* ‘sort’ menggunakan *barplot*. Kata dengan frekuensi kemunculan terbanyak juga ditampilkan dengan menggunakan metode *WordCloud*. Lima kata dengan frekuensi terbanyak yang ditampilkan menggunakan metode *barplot* dan *WordCloud* adalah ‘masyarakat’, ‘vaksinasi’, ‘covid’, ‘vaksin’, dan ‘pemerintah’.

## DAFTAR PUSTAKA

- [1] World Health Organization (WHO). <https://covid19.who.int/table> diakses pada 27 November 2021.
- [2] Ipmawati, J. (2016). Komparasi teknik klasifikasi teks mining pada analisis sentimen. *IJNS-Indonesian Journal on Networking and Security*, 6(1).