

ANALISA SENTIMEN TENTANG REUNI 212 PADA DOKUMEN TWITTER BERBAHASA INDONESIA MENGGUNAKAN R STUDIO

Hudi Kusuma Bharata

Program Studi Teknik Informatika, STMIK Bani Saleh, hudi@stmik.banisaleh.ac.id

HS. Sulistyowati

Program Studi Teknik Informatika, STMIK Bani Saleh, hs.sulistyowati@gmail.com

ABSTRAK

Analisa sentimen masyarakat mengenai Reuni 212 Tahun 2018 yang diekspresikan lewat tweet dari Twitter merupakan fenomena yang menarik untuk diteliti. Dikarenakan peristiwa Reuni 212 2018 merupakan peristiwa yang menjadi sorotan sebagian besar rakyat Indonesia dan otomatis akan memancing opini pro, kontra atau netral.

Adapun metode penelitian yang digunakan adalah *one-shot experiment* dengan Twitter sebagai sumber data penelitian. Tweet periode tertentu dikumpulkan ke folder lokal dalam bentuk file CSV, dibaca, diolah dan kemudian diklasifikasikan dengan membandingkan terhadap kepastakaan kata positif dan negatif. Masing-masing kelompok kata dihitung skor klasifikasi positif, netral dan negatif kemudian divisualisasikan dalam bentuk grafik histogram dan wordcloud. Sebelum kalkulasi aritmatik dijalankan, data tweet yang diperoleh dibersihkan dari karakter yang tidak digunakan seperti tanda ^, @, http:// dan dengan dukungan kepastakaan *stopword* bahasa Indonesia maka kalimat yang ada akan dipotong sesuai dengan kaidah KBBI.

Hasil yang diharapkan adalah gambaran analisa sentimen masyarakat mengenai peristiwa Reuni 212 tahun 2018 yang diperoleh dari jejaring sosial dan diwakili oleh komunitas pengguna Twitter dan diolah dengan pendekatan ilmiah.

Kata Kunci: analisa sentimen, reuni 212, API Twitter, R studio, big data

Sentiment Analysis regarding Reunion 212 year 2018 which expressed through tweets is an interesting phenomenon to be studied. Reunion 212 year 2018 is an event that is in the spotlight of most Indonesian people and will automatically provoke pro, contra or neutral opinions.

The research method used is a one-shot experiment with Twitter as a research data source. Specific period tweets are then collected into local folders in the form of CSV files, followed by read, processed and classification by comparing each single word with the positive and negative Indonesian words library. Each word group was calculated as a positive, neutral and negative classification will then scored and visualized in histogram and wordcloud graphs. Prior to arithmetic calculation is carried out, the tweet data obtained is cleared from unused characters such as ^, @, http:// sign and with the support of the Indonesian stopwords library, the sentences will be decomposed under KBBI rules.

The expected results are illustrations of a people sentiment analysis regarding the Reunion 212 happened in year 2018 obtained from social networks represented by the Twitter community and processed with a scientific approach.

Key word: sentiment analysis, Reunion 212, API Twitter, R Studio, Big Data

PENDAHULUAN

Big Data bukanlah sebuah teknologi, teknik, maupun inisiatif yang berdiri sendiri. Big Data adalah suatu trend yang mencakup area yang luas

dalam dunia bisnis dan teknologi. Big Data menunjuk pada teknologi dan inisiatif yang melibatkan data yang begitu beragam, cepat berubah, atau berukuran super besar sehingga terlalu sulit bagi teknologi, keahlian, maupun infrastruktur

konvensional untuk dapat menanganinya secara efektif. Implementasi *Big Data* di berbagai situs jejaring sosial termasuk *Twitter*, *Facebook*, dan *LinkedIn* merupakan sumber data yang dapat dimanfaatkan untuk berbagai keperluan. Semua sumber data yang terdiri dari data teks, gambar dan video baik terstruktur atau tidak terstruktur dapat diolah dengan *data mining* untuk memperoleh kinerja tata kelola sistem yang diinginkan. Saat ini wawasan perencanaan dan intelegen bisnis yang mengarah ke pengambilan keputusan strategis banyak memperoleh masukan dari media sosial. *Big Data* pada awalnya memiliki dimensi 3V namun saat ini berkembang menjadi 6V atau bahkan 7V. Pada awalnya metrik “*bigness*” adalah *Volume*, *Velocity* dan *Variety*. Namun metrik tersebut saat ini sudah bertambah dengan *Value*, *Variability*, *Veracity* dan *Visualization*. Berkembangnya *Big Data* membuat disiplin ilmu *data science* semakin banyak berperan dalam analisa, perencanaan arah dan strategi perusahaan serta dalam pengambilan keputusan bisnis penting, dikenal dengan *Data-driven decision making*.

Analisa Big Data

Pengolahan sekumpulan data dalam jumlah besar baik untuk analisa deksriptif untuk memahami hubungan antar data atau analisa prediktif untuk menggali pola baru trend terkini pasar. Berbagai pola data tersebut kemudian dikonversi menjadi pengetahuan yang mampu ditindaklanjuti dan digunakan dalam pengambilan keputusan. Analisa *Big Data* dapat membantu organisasi dalam memahami informasi yang terkandung dalam data secara lebih baik serta membantu identifikasi data penting yang dibutuhkan bisnis dan keputusan bisnis dimasa yang akan datang.

Untuk menganalisa data bervolume besar diperlukan aplikasi alat bantu berupa software untuk analisa prediktif, *data mining*, *text mining*, *forecasting* dan optimalisasi data. Software *Big Data* seperti **Clojure**, **Scala**, **Python**, **Hadoop**, **Java for NLP** (*Natural Language Processing*), *Text Mining*, **R**, dan **MATLAB** dapat digunakan untuk analisa data.

Jurnal ini akan menganalisa sentimen pengguna *Twitter* yang diperoleh dari kumpulan data *Tweet* mengenai peristiwa Reuni 212 Tahun 2018.

Twitter

Perkembangan akses teknologi informasi yang semakin pesat pada era globalisasi saat ini dapat

memberikan kemudahan untuk berkomunikasi lebih efektif dan efisien, contohnya adalah internet. Seperti halnya yang telah diungkapkan oleh Kementerian Komunikasi dan Informatika (Kemenkominfo, 2013) bahwa pengguna Internet di Indonesia saat ini mencapai 63 juta orang. Dari angka tersebut, terdapat 95 Persennya menggunakan internet untuk mengakses media sosial. Salah satu media sosial yang diakses masyarakat adalah *Twitter*.

Media sosial *twitter* merupakan salah satu situs *Microblogging* yang memungkinkan penggunanya untuk menulis tentang berbagai topik dan membahas isu-isu yang sedang terjadi. *Twitter* memberikan layanan kepada penggunanya untuk mengirim atau membaca *Tweets* yang telah dibagikan dengan membatasi karakter maksimal 140. Adanya layanan tersebut menyebabkan masyarakat lebih memilih menuangkan opininya melalui media sosial daripada menyampaikannya secara langsung. Opini-opini tersebut dapat dimanfaatkan untuk mencari sebuah informasi.

Namun, dalam pemanfaatannya membutuhkan analisis yang tepat sehingga informasi yang dihasilkan dapat membantu banyak pihak Untuk mendukung suatu keputusan atau pilihan. Pesan yang dibuat menggunakan aplikasi *Twitter* disebut *Tweet* dan merupakan data publik. Data ini dapat diperoleh dalam bentuk data mentah yang kemudian dapat diekstraksi opininya untuk tujuan antara lain analisa kepuasan pelanggan, dan analisa sentimen.

Reuni 212

Peristiwa berkumpulnya ummat Islam Indonesia di lapangan Monas dan sekitarnya yang terjadi pada hari Ahad, 2 Desember 2018 atau 24 *Rabbiul Awwal* 1440 H di kawasan Tugu Monumen Nasional, kota Jakarta. Hadirin melaksanakan Sholat Tahajjud, Sholat Shubuh berjamaah, mendengarkan tausiyah dari beberapa ulama serta diakhiri dengan Sholat Dhuhur berjamaah dikenal dengan nama Reuni 212 2018.

Penyelenggara Reuni 212 menyatakan bahwa acara tersebut bukan merupakan agenda politik suatu partai tertentu, namun dekatnya kontestasi pemilihan presiden menimbulkan polemik yang merambah dunia maya antara pendukung dengan penentang perhelatan Reuni 212 tahun 2018. Peristiwa ini juga terrekam dalam *Tweet* pengguna *Twitter* dari pihak pro maupun kontra serta yang memberikan komentar netral.

Analisa Sentimen

Text mining adalah sebuah proses pengambilan informasi dari data tekstual yang memiliki kualitas tinggi serta dapat mengetahui permasalahan dalam teks dari sebuah topik tertentu. Text mining dalam analisa sentimen mampu mengidentifikasi emosional terhadap suatu pernyataan (Zhang, et al, 2015). Semenjak penelitian dari (Pang, Lee, 2008), banyak ditemukan studi pembelajaran tentang analisa sentimen. Analisa sentimen sering juga dikenal dengan opinion mining adalah studi omputasi dari pendapat, sentimen, sikap, serta emosi yang disajikan dalam sebuah teks (Liu, B., 2012). Seperti pada penelitian (Patela, T., Undavia, J., & Patela, A., 2015) melakukan analisa sentimen terhadap feedback yang diberikan oleh orang tua murid dengan membagi sentimen menjadi 3 kategori yaitu sentimen positif, negatif dan netral.

Analisa sentimen dapat digunakan oleh pelanggan untuk mengetahui kualitas produk atau jasa yang akan dibeli. Produsen dapat menggunakan analisa sentimen sebagai feedback dari pelanggan terhadap produknya sehingga dapat diukur tingkat kepuasan pelanggan serta langkah-langkah peningkatan kualitas bila diperlukan. Analisa sentimen saat ini merupakan salah satu riset populer di area komputasi karena adanya ledakan informasi sentimen dari website, forum online serta blog. Analisa sentimen dapat dilakukan dengan melihat respon positif, netral atau negatif yang dihasilkan dari pengelompokan Tweet dengan metode klasifikasi *Bayesian*.

Adalah perhitungan sentimen masyarakat yang berisi opini atas suatu peristiwa tertentu. Analisa ini akan mengelompokkan review positif, negatif dan netral dan kemudian membandingkan banyaknya masing-masing kelompok opini tersebut serta digambarkan dalam visualisasi grafik histogram.

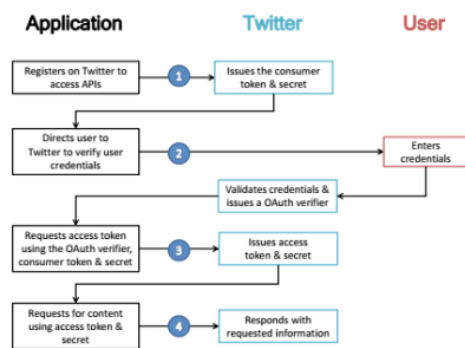
TEORI

Twitter Environment

Twitter adalah aplikasi jejaring sosial yang menyediakan fasilitas komunikasi antar pengguna. Twitter diciptakan oleh Jack Dorsey pada tahun 2006 dan pertama kali diperkenalkan di dunia maya pada bulan Juli 2006 (<http://www.twitter.com>). Pengguna bisa mulai membahas suatu topik (mention), mengirim (Tweet), dan meneruskan (Retweet) pesan yang diinginkan. Pemanfaatan Twitter sudah merambah ke berbagai bidang, karena Twitter menyediakan fasilitas untuk membahas topik yang sedang terjadi didunia serta memfasilitasi

komunikasi antar pengguna. Hampir setiap peristiwa akan dibahas dan ditanggapi lewat Twitter, terutama menjelang Pilpres 2019 di Indonesia. Peristiwa politik maupun non politik akan selalu ditanggapi pro dan kontra, sehingga bisa dijadikan bahan analisa kecenderungan respon masyarakat terhadap isu yang terjadi. Namun pemanfaatan data yang tersimpan di Twitter sebelumnya harus melalui tahapan registrasi dan setelah memperoleh kode API Twitter maka pengguna bisa mengambil data Tweet dari environment Twitter. Setiap kali terjadi interaksi antara aplikasi pengguna dengan Twitter selalu diawali dengan proses Otentifikasi. Proses Otentifikasi terhadap API Twitter dikerjakan oleh paket Oauth dari R. Gambar dibawah menggambarkan tahapan penggunaan Oauth dalam mengakses API Twitter. Perlu dibuat sebuah aplikasi untuk menjalankan API Twitter, dalam hal ini aplikasi yang dibuat dengan bahasa R.

1. Untuk memperoleh API Twitter, user harus registrasi ke Twitter yang kemudian akan diberikan sejumlah key API setelah disetujui oleh Twitter. Pengguna harus menjelaskan rencana penggunaan API beserta palikasi yang akan memanfaatkannya karena key API dan kelengkapannya harus dimasukkan kedalam aplikasi yang memerlukan otentifikasi.



Gambar 1 Arsitektur Twitter

2. Key dari Twitter akan digunakan untuk memulai link ke Twitter agar otentifikasi segera dapat dimulai.
3. Setelah proses otentifikasi API Twitter Access Token dan Access Secret untuk masing-masing user selesai dilakukan, maka pengguna dapat melanjutkan ke proses berikutnya.
4. Proses selanjutnya adalah mulai mengumpulkan Tweet yang dikehendaki dengan memberikan keyword serta jumlah Tweet yang diinginkan. Twitter memiliki kebijakan untuk membatasi jumlah Tweet yang diberikan kepada pengguna apabila tidak secara spesifik ditentukan.

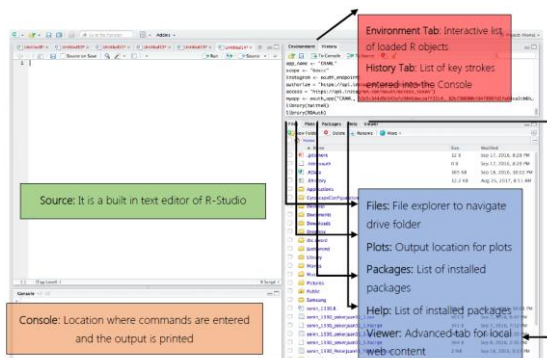
R-Studio

R adalah sebuah perangkat lunak gratis untuk komputasi statistik dan grafik (Faisal, 2016). R dapat diekstensikan dengan mudah terhadap lebih dari 4.728 paket yang tersedia di CRAN. Selain di CRAN, banyak packages lain yang tersedia seperti di Github.

R-Studio digunakan secara luas, baik di bidang Akademis dan di bidang Industri. R-studio mendapatkan peringkat 1 pada polling KD Nuggets 2013 dalam kategori “Top Languages for Analytics, Data Mining, and Data Science”.

Cara yang paling mudah untuk dapat segera menggunakan R adalah dengan melakukan instalasi R dan menggunakan R Studio sebagai IDE (Integrated development environment). R sendiri adalah suatu bahasa pemrograman yang memang difokuskan untuk melakukan kegiatan yang berkaitan dengan statistika dan analisa grafik.

Untuk dapat melakukan pemrograman pada R sebenarnya kita tidak perlu menggunakan R Studio namun tanpa adanya tools tambahan yang membantu kita dalam mengelola project R akan menambah pekerjaan yang harus kita lakukan. Sehingga R Studio berfungsi untuk menujung pekerjaan kita dan diharapkan lebih cepat dalam menyelesaikan pekerjaan. Kita tidak perlu pusing dalam konfigurasi R dan fokus ke script untuk melakukan kegiatan yang berkaitan dengan statistika dan analisa grafik.



Gambar 2. Layar R-Studio

Layar R Studio memiliki 4 window environment, yaitu a.l:

1. Area kiri bawah adalah work area tempat implementasi script R.
2. Area kiri atas adalah tempat menuliskan script R yang eksekusinya dapat dikerjakan per baris atau per kelompok perintah yang kemudian hasilnya dapat dilihat di work area.

3. Area kanan atas adalah Global Environment yang menampilkan definisi variable dan data yang diperoleh dari implementasi perintah R.
4. Area kanan bawah adalah area untuk menampilkan data dalam bentuk grafik. Tampilan grafik di area tersebut bisa diperbesar dengan perintah zoom yang akan menampilkan window khusus tampilan grafik.

Paket R

Paket R adalah kumpulan fungsi-fungsi R dalam bentuk coding yang sudah dicompile. Fungsi-fungsi ini disimpan di library R dimana aplikasi tersebut diinstall. Belum semua paket yang dibutuhkan terpasang pada saat instalasi awal, namun bila terdapat proses yang membutuhkan paket tertentu akan dengan mudah ditambahkan kedalam instalasi kepustakaan R (Paradis, 2005). Berikut ini paket-paket yang dibutuhkan untuk implementasi analisa sentimen sesuai judul jurnal ini.

1. **twitterR**: Digunakan untuk interface ke API Twitter
2. **ROAuth**: User dapat melakukan otentifikasi langsung ke server menggunakan fungsi Oauth (spesifikasi 1.0)
3. **Plyr**: Mengelola problem besar dengan membagi menjadi kepingan yang dapat ditangani, diolah dan dikumpulkan kembali menjadi suatu solusi
4. **Stringr**: Bertugas menangani semua fungsi String di R. Paket ini mempermudah penanganan data karakter dengan panjang nol serta NA (not applicable).
5. **Ggplot2**: Paket untuk implementasi grafik di R dalam format grafik dasar maupun lattice serta mendukung multi sumber data \.
6. **Rcolorbrewer**: Digunakan untuk menggambar peta bersama sesuai variable palet
7. **Devtools**: Paket yang digunakan untuk mempermudah pekerjaan developer dengan menyederhanakan beberapa task yang umum digunakan.

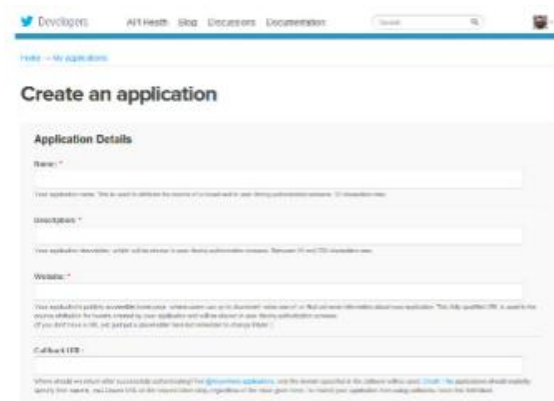
METODOLOGI

Analisa Sentimen Twitter dilakukan dengan langkah-langkah sebagai berikut:

1. Membuat aplikasi Twitter
2. Eksekusi kode API Twitter di R Studio
3. Mengumpulkan arsip data Twitter
4. Klasifikasi data dengan perintah R
5. Jalankan perintah R untuk mengolah Tweets
6. Persiapkan plotter R untuk menampilkan hasil dalam bentuk grafik

Membuat Aplikasi Twitter

Fungsi pencarian API Twitter digunakan untuk mengakses tweet terbaru dalam periode yang sudah ditentukan, yaitu 7 (tujuh) hari sebelum dan sesudah tanggal 2 Desember 2018. Hasil pencarian tweet tersebut akan digunakan sebagai data yang akan diolah pada riset ini. Untuk memperoleh API Twitter, pengguna harus berhubungan langsung dengan Twitter dan menjelaskan rencana penggunaan API yang diminta. Pemanfaatan API harus menggunakan aplikasi Twitter, dan untuk membuat aplikasi Twitter dapat dilihat pada Gambar 2 dibawah ini.



Gambar 3. Identifikasi Aplikasi Twitter

Eksekusi API Twitter di R

Kode interface program aplikasi Twitter search harus dijalankan melalui R Studio. Agar dapat mengumpulkan Tweet yang dikehendaki, harus terbentuk lebih dahulu interface dengan Twitter melalui website Twitter. Setelah tweet yang diinginkan terkumpul kemudian disimpan kedalam file CSV (Comma Separated Values). Agar proses pengumpulan tweet dapat berjalan dengan baik, diperlukan instalasi beberapa paket sebelum pencarian dilakukan.

```
# Web scrapping project in Twitter
library(stringr)
library(RCurl)
library(tm)
library(RColorBrewer)
library(wordcloud)
library(SnowballC)
library(RXKCD)
library(twitter)
library(ROAuth)
library(plyr)
library(ggplot2)
library(RSentiment)
```

```
library(stopwords)
```

Diikuti dengan perintah untuk mengaktifkan API key dan token yang diberikan oleh Twitter dengan contoh sbb:

```
# API Keys
api_key <- "your API keys"
api_secret <- "your API Secret"
access_token_secret <- "your token secret"
access_token <- "your token"

setup_twitter_oauth(api_key,api_secret,access_token,access_token_secret)
```

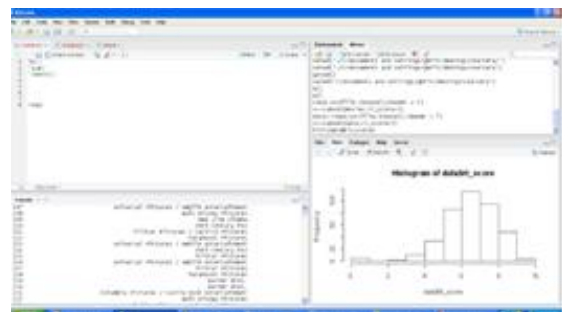
Selanjutnya mencari Tweet berbahasa Indonesia dengan topik "Reuni 212", dibatasi sebanyak 1000 tweet pertama dari periode tanggal 25 Nopember 2018 sampai dengan tanggal 9 Desember 2018.

```
tweets <- searchTwitter("reuni 212", n=1000, lang="id", since='2018-11-25', until='2018-12-09')
```

Perintah-perintah R ditulis di area script R Studio dan dapat dieksekusi per baris untuk menerima, mengolah dan melihat hasil pengolahan data. Dataset dan hasil pengolahannya dapat dilihat di area kanan atas. Grafik hasil pengolahannya dapat dilihat di area kanan bawah.

Import Data via Twitter API

Setelah terjadi *property handshake* dengan Twitter, *Tweet* yang sesuai dengan keyword pencarian dan rentang waktu yang ditentukan akan diterima oleh R Studio.



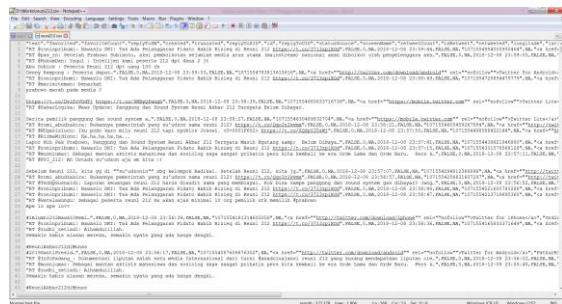
Gambar 4. Tampilan R-Studio

Fungsi *searchTwitter* akan digunakan untuk memperpanjang fase akhir mengunduh tweet dari *timeline*. *Tweet* yang telah terkumpul dapat disimpan kedalam file CSV di folder lokal atau langsung

diproses dengan perintah selanjutnya. Dibawah ini contoh Tweet yang telah disimpan kedalam file CSV:

Standarisasi Data

Tweet yang terkumpul biasanya masih becampur dengan karakter-karakter khusus dan URL. Oleh karena itu, perlu dijalankan “data cleansing” dengan memanggil beberapa beberapa fungsi perintah untuk mengubah kumpulan tweet menjadi informasi yang dapat dibaca.



Gambar 5. File CSV berisi Tweet

Proses ini disebut sebagai standarisasi data, yaitu proses penghapusan simbol-simbol tambahan dan link alamat yang tidak memiliki arti pada tweet agar beban klasifikasi bisa lebih ringan.

```
clean.text = function(x)
{
  # remove RT
  x = gsub(" (RT|via) ((?:\\b\\W*@[\\w+)+)", "", x)
  # tolower
  x = tolower(x)
  # remove at
  x = gsub("@\\w+", "", x)
  # remove U
  x = gsub("^\\s*<U\\+|\\w+>|-", "", x)
  # q <- "<U+00A6> 1000-66329"
  x = gsub("<U+00A6> 1000-66329", "", x)
  # x <- q
  # remove punctuation
  x = gsub("[[:punct:]]", "", x)
  # remove numbers
  x = gsub("[[:digit:]]", "", x)
  # remove links http
  # x = gsub("
  ?(f|ht)tp(s?)://(.*)[.][a-z]+",
  "", x)
}
```

```
x =
gsub("(f|ht)tp(s?)://(.*)[.][a-z]{2,6}", "", x)
x = gsub("http\\w+", "", x)
# remove tabs
x = gsub("[ \\t]{2,}", "", x)
# remove blank spaces at the beginning
x = gsub("^ ", "", x)
# remove blank spaces at the end
x = gsub(" $", "", x)
# remove â€¦
x = gsub("â€¦", "", x)
# remove â€œ
x = gsub("sâ€œ", "", x)
return(x)
}
```

Klasifikasi Data

Proses analisis sentimen adalah sinkronisasi kata-kata tweet dengan daftar kata positif dan daftar kata negatif yang telah diunduh dan disimpan di D:/Works. Setiap kata yang memenuhi salah satu dari klasifikasi tersebut dikumpulkan dalam masing-masing variable teks penampungan untuk diproses pada tahapan selanjutnya. Selain kebutuhan file referensi, analisa sentimen juga membutuhkan dua paket tambahan kepastakaan **plyr** dan **stringR** terinstal di environment R Studio yang berguna untuk memanipulasi kumpulan kata tersebut.

Perintah-perintah dibawah ini digunakan untuk membandingkan kata yang diperoleh dengan file kata positif dan negatif dan hasilnya disimpan di variable memory.

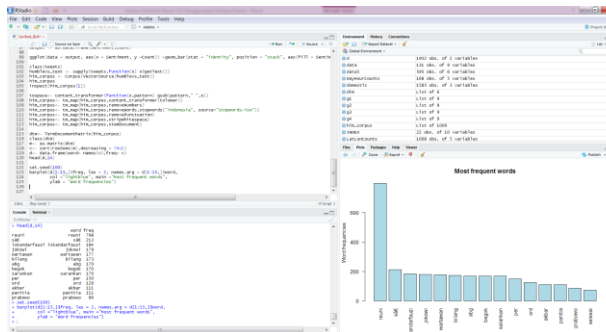
```
setwd("D:/Works")
pos <- scan("positif.txt", what="character", comment.char=";")
neg <- scan("negatif.txt", what="character", comment.char=";")
source("sentiment.R")
```

Menghitung Skor

Analisa sentimen dilakukan dengan menghitung banyaknya kata yang bersifat positif atau negatif pada setiap tweet dan kemudian dihitung skor untuk tiap tweet. Skor klasifikasi positif ditambah bila kata-kata di tweet terdapat dalam kepastakaan kata positif, demikian juga dengan skor klasifikasi negatif bertambah bila ditemukan didalam daftar kata negatif. Semakin sering bertemu dengan kata-kata yang cocok dengan daftar positif akan

menempatkan semua tweet di sumbu X dan skala skor yang sesuai pada sumbu Y.

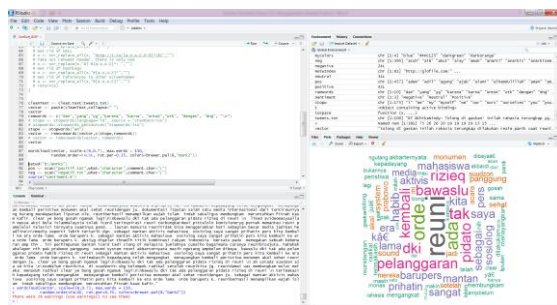
```
set.seed(100)
barplot(d[1:10,]$freq, las = 2,
names.arg = d[1:10,]$word,
      col="lightblue", main
="Most frequent words",
      ylab = "Word
frequencies")
```



Gambar 11 Histogram Pemunculan Kata

Untuk menampilkan beberapa kata dengan pemunculan terbanyak, bisa menggunakan fungsi Wordcloud sbb:

```
wordcloud(vector, scale=c(6,0.7),
max.words =150,random.order
=FALSE, Rot.per=0.35,colors=
brewer.pal(8,"Dark2"))
```



Gambar 12 Wordcloud pemunculan kata

PENUTUP
Simpulan dan Saran

Berkembang pesatnya jejaring sosial akan menghasilkan data dalam jumlah besar dan membutuhkan lebih dari sekedar alat data mining biasa untuk menganalisa datanya. Dengan digunakannya R Studio untuk mengakses environment Twitter dan analisa data peristiwa

Reuni 212 tahun 2018, tampak bahwa opini publik yang terrekam di dalam tweet masing-masing dapat diklasifikasikan kedalam kelompok Positif, Netral dan Negatif beserta visualisasinya. Oleh karena itu, R dan R Studio ini patut direkomendasikan sebagai alat analisa Big Data untuk kasus-kasus lainnya setara dengan bahasa pemrogram Python atau yang sejenisnya.

DAFTAR PUSTAKA

Pak A, Paroubek P. *Twitter as a corpus for sentiment analysis and opinion mining*. Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10); 2010 May 19-21.

Ding X, Liu B, Yu PS. *A holistic lexicon-based approach to opinion mining*. Proceedings of First ACM International Conference on Web Search and Data Mining WSDM; 2008.

Paradis, Emanuel, R for beginners, https://cran.r-project.org/doc/contrib/Paradis-rdebuts_en.pdf [diakses 1 Agustus 2018]

Faisal, M. Reza, *Seri Belajar Pemrograman: Pengenalan Bahasa Pemrograman R*, e-book, (2016)

Implementasi Big Data untuk pencarian pattern data gudang pada PT. Bank Mandiri (Persero)TBK. https://www.researchgate.net/publication/13847608_Implementasi_Big_Data_untuk_pencarian_pattern_data_gudang_pada_PT_Bank_Mandiri_Persero_TBK [Diakses 5 Agustus 2018].

Liu, B.. *Sentiment Analysis and Opinion Mining. Synthesis Lectures On Human Language Technologies*, (2012), 1-167.

Patel, T., Undavia, J., & Patela, A. (2015). *Sentiment Analysis of Parents Feedback for Educational Institutes*. International Journal of Innovative and Emerging Research in Engineering , (2015), 2(3).

Pang, Lee, *Opinion Mining and Sentiment Analysis, Foundations and Trends in information*, 2008, 1-135