# 

# **LAPORAN PROYEK AKHIR**

**PRAKTIKUM DATA SCIENCE**

**ANALISIS SENTIMEN MASYARAKAT TERHADAP**

**KULIAH ONLINE**



|  |  |
| --- | --- |
| FARAH AISYAH | 123200009 |
| NICHOLAS INDRANTO HARTONO | 123200038 |
| JUAN BONTHOR HUTAPEA | 123200117 |

**PROGRAM STUDI INFORMATIKA**

**JURUSAN TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK INDUSTRI**

**UNIVERSITAS PEMBANGUNAN NASIONAL**

**“VETERAN” YOGYAKARTA**

**2022**

**HALAMAN PENGESAHAN LAPORAN PROYEK AKHIR**



Disusun oleh :

FARAH AISYAH123200009

NICHOLASINDRANTOHARTONO123200038

JUAN BONTHOR HUTAPEA 123200117

Telah Diperiksa dan Disetujui oleh Asisten Praktikum Data Science

Pada Tanggal : ................................

.

|  |  |  |
| --- | --- | --- |
| **Asisten Praktikum**  **Ega Erinovian Megananda Putra**  123190113 |  | **Asisten Praktikum**  **Luckman Nathan Syarif Aljustin**  123190151 |

# **PENDAHULUAN**

Dampak pandemi Covid-19 pada tahun 2020 ini di Indonesia mulai merambah dunia pendidikan, pemerintah pusat hingga daerah memberikan kebijakan untuk meliburkan seluruh lembaga pendidikan. Hal ini dilakukan sebagai upaya mencegah meluasnya penularan virus corona. Diharapkan dengan seluruh lembaga pendidikan tidak melaksanakan aktivitas seperti biasanya, hal ini dapat meminimalisir menyebarnya penyakit covid 19 ini

Kebijakan yang diambil oleh banyak negara termasuk Indonesia dengan meliburkan seluruh aktivitas pendidikan, membuat pemerintah dan lembaga terkait harus menghadirkan alternatif proses pendidikan bagi peserta didik maupun mahasiswa yang tidak bisa melaksanakan proses pendidikan pada lembaga pendidikan. Salah satu kebijakan di bidang pendidikan yang diambil oleh pemerintah terkait kasus COVID 19 yaitu : kegiatan pembelajaran baik sekolah maupun perkulaiahn secara daring (*online*).

Dalam pendidikan, kesulitan muncul, seperti ketidaksiapan teknologi, media yang digunakan, atau aspek psikologis siswa. Sementara itu, masa depan akan terjadi ketimpangan pendidikan antara penduduk dan daerah di Indonesia dalam jangka Panjang. Dari pemaparan tersebut terdapat berbagai tanggapan dari masyarakat terhadap kuliah online. Maka dari itu penyusun membuat program untuk mengetahui tanggapan dari masyarakat terhadap kuliah online tersebut dengan menggunakan bahasa pemrograman R dan metode naïve bayes.

# **METODE**

## Data Processing

Data processing merupakan proses dikumpulkannya sebuah data serta melakukan proses cleaning data agar menjadi sebuah infomasi yang bermanfaat serta mudah digunakan untuk proses modeling data yang akan dilanjutkan pada tahap selanjutnya. Untuk data processing ini sendiri terdapat beberapa proses yaitu:

1. Impor Data

Pada tahap ini dilakukan mengupload data set kedalam program, hal ini dilakukan bertujuan agar data yang diimpor dapat diproses pada tahap selanjutnya.

1. Data Cleaning

Pada proses ini dilakukan sebuah proses pembersihan serta proses mengubah beberapa data yang terdapat pada dataset. Pada proses ini cukup penting dilakukan untuk memudah dilakukannya data modeling, selain itu data cleaning ini juga berfungsi untuk mendapatkan data yang sederhana dan lebih efisien ketika ingin disajikan. Data dibersihkan dari karakter-karakter yang dapat menjadi sumber miskalkulasi, seperti tanda-tanda baca hingga url atau link web dengan menggunakan bantuan library tm (text-mining) dengan menggunakan fungsi tm\_map(). Data juga dibersihkan dari kata-kata yang tidak diperlukan seperti kata-kata subject, dan katakata lain yang relatif tidak memiliki sentimen atau emosi. Kata-kata tersebut ditampung ke dalam file ‘stopwordsen.txt’.

1. Export Data

Setelah melakukan data cleaning, data akan di export agar bisa diproses ditahap *Data Visualization*. Data tersebut akan diubah ke dalam bentuk data frame dan akan disimpan ke dalam file baru yang memiliki ekstensi csv.

## 2.2. Data Visualization

Setelah dilakukan data processing, dilakukan visualisai data berupa scatterplot dan wordcloud. Visualisasi data ini bisa dilakukan menggunakan bantuan library yang telah di import pada awal program di Rstudio. Untuk library yang digunakan untuk mengubah data menjadi ke bentuk visual yang bisa dilihat antara lain:

1) Ggplot2, *library ggplot2* ini berfungsi untuk dilakukannya proses visualisasi data, agar data dapat berubah menjadi bentuk yang lebih sederhana dan mudah untuk dipahami.

2) Plotly, *library plotly* ini digunakan agar data dapat divisualisasi ke dalam bentuk chart agar data dapat dilihat lebih mudah.

3) Wordcloud, library word cloud digunakan agar data dapat divisualisasi ke dalam bentuk word cloud. Sehingga frekuensi data dapat dilihat lebih mudah dianalisis, khususnya data-data yang sering muncul.

## 2.3 Deployement

Deployment dilakukan dengan bantuan library shiny. Shiny adalah sebuah paket dalam R yang mengizinkan penggunanya membangun web apps yang interaktif. Shiny menggabungkan antara kekuatan komputasi statistika R dan interaksinya dengan web modern. User Interface (UI) merupakan fungsi yang mendefinisikan tampilan web dari aplikasi.

Data-data yang ditampilkan merupakan data yang telah melewati tahap data visualization. Data-data tersebut berupa dataset yang telah melalui tahap data cleansing, scatterplot dan wordcloud. Masing-masing data tersebut akan dibungkus ke dalam ui yang ditampilkan dengan shiny.

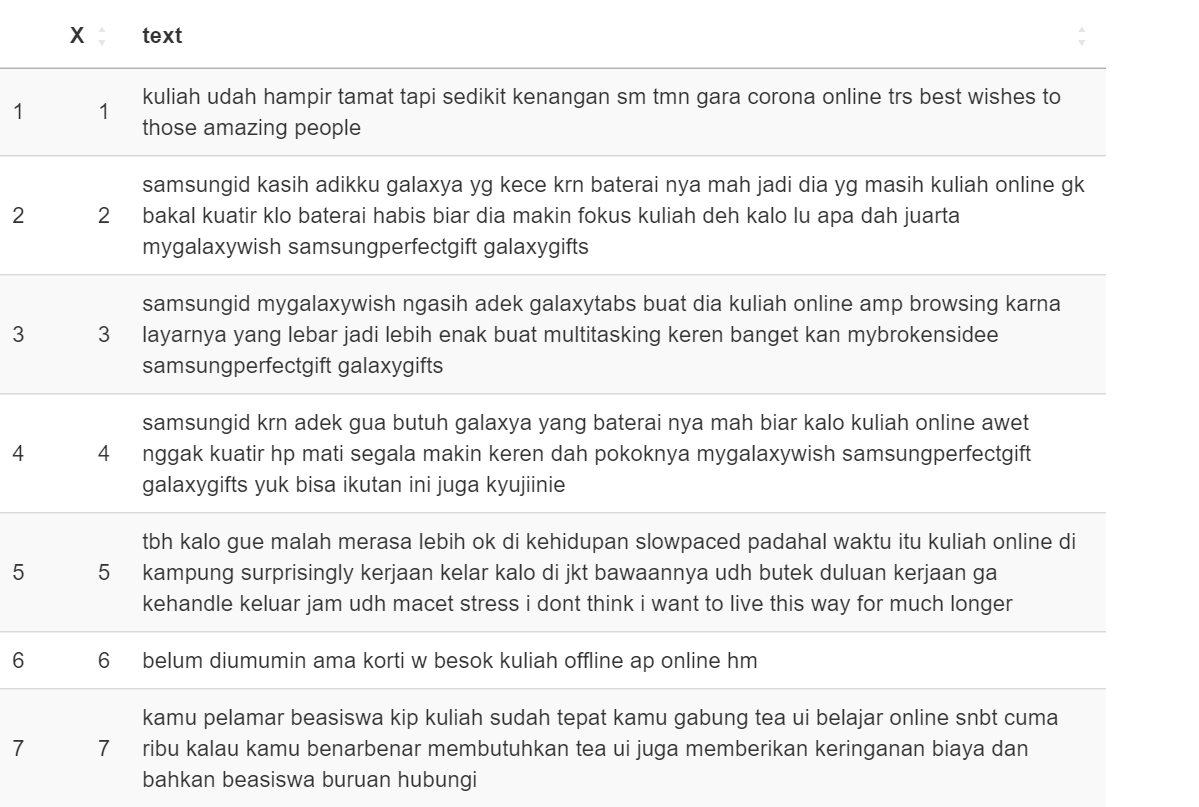
# **HASIL DAN PEMBAHASAN**

## Data Processing

## 

## Gambar 3.1 Dataset

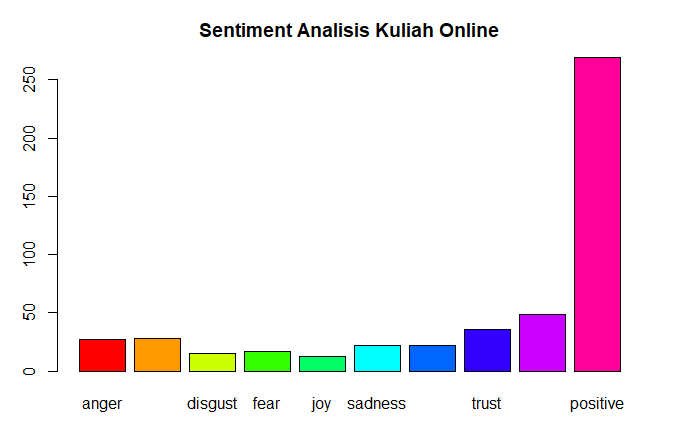
Data yang diambil pada dataset hanya pada kolom review title kemudian dilanjutkan dengan proses data processing. Pada bagian data processing, dataset yang telah diimport akan diubah ke dalam bentuk dataframe.



**Gambar 3.2 Dataset setelah melalui proses data processing**

Pada bagian data processing, dataset yang telah diimport akan diubah ke dalam bentuk dataframe. Kemudian dilakukan proses data cleaning untuk menghilangkan karakter-karakter seperti tanda baca, hingga karakter url atau web link yang dapat menjadi sumber miskalkulasi.

## 3.2. Data Visualization



**Gambar 3.3 Diagram Batang Sentiment Masyarakat**

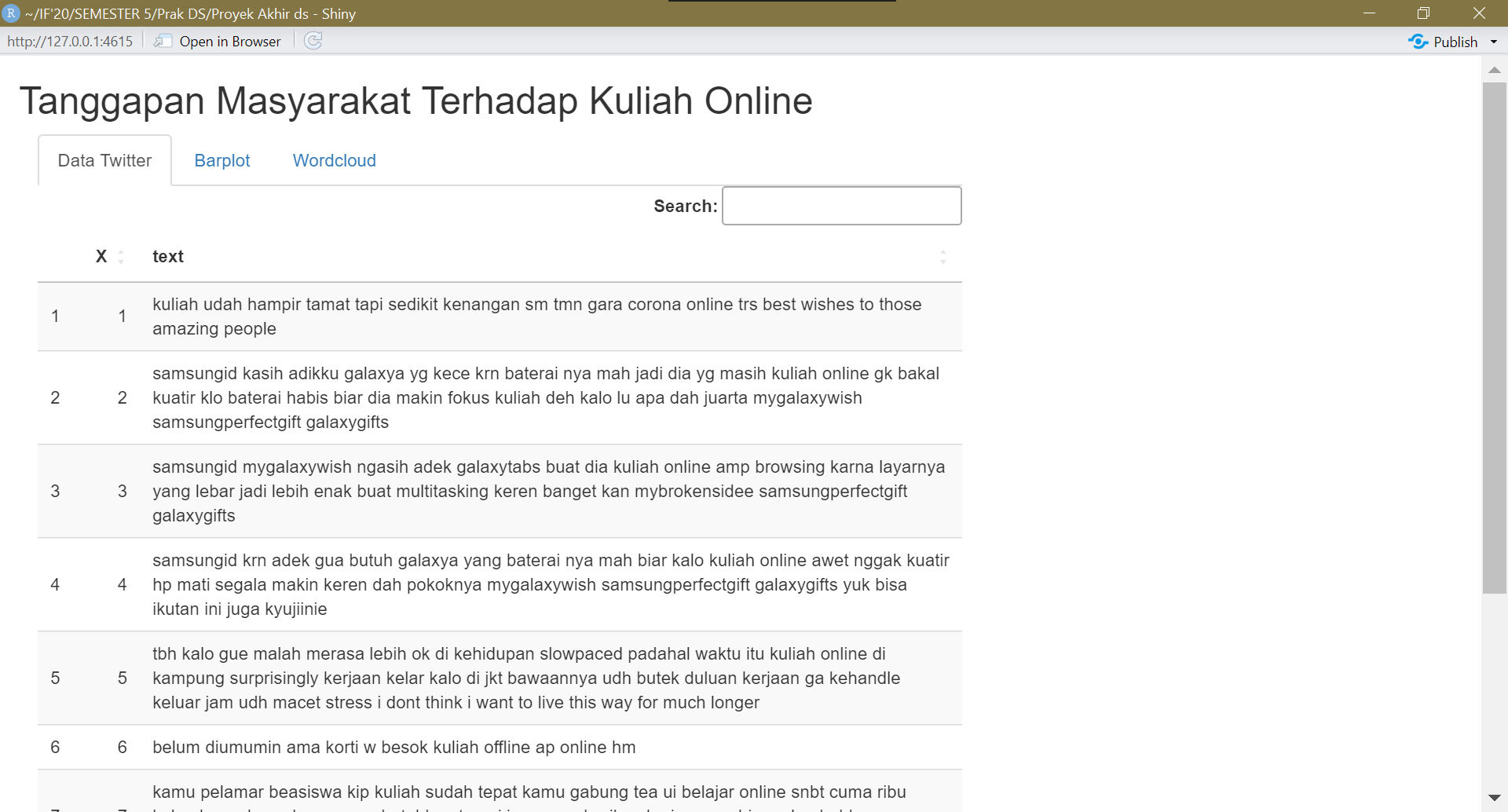
Data divisualisasikan menjadi scatterplot. Sumbu X merupakan sentimen dari kata yang berada di dalam dataset dan Sumbu Y merupakan banyaknya kata yang masuk ke dalam kategori suatu sentimen.



**Gambar 3.4 Wordcloud**

Interpretasikan dari visualisasi wordcloud, kata-kata yang sering muncul dari dataset yaitu “online”, “kuliah”, “dana”, dan lainnya.

## 3.3 Deployement



**Gambar 3.5 Tampilan GUI Shiny**

Menampilkan dataset tanggapan masyarakat terhadap kuliah online yang telah melalui tahap data processing, scatterplot atau diagram batang sentimen, dan wordcloud.

# **KESIMPULAN**

Dari hasil analisis sentiment masyarakat terhadap kuliah online yang telah dibuat, rata-rata masyarakat memberikan respon positif yang besar. Dan pada barplot dapat dilihat bahwa sentiment masyarakat terbanyak ialah positif yang jumlah datanya lebih dari 250.