

**LAPORAN PROYEK AKHIR
PRAKTIKUM DATA SCIENCE
ANALISIS SENTIMEN COVID-19 PADA TWEETS**



AZQIA ADISTYA ABDILLAH	123190064
WAFFIQ AZIZ	123190070

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

1. PENDAHULUAN

Sejak awal tahun 2020, ramai menjadi perbincangan masyarakat mengenai adanya sebuah virus yang memiliki siklus penyebaran yang sangat cepat. Virus tersebut sering disebut dengan CoronaVirus, kemudian penyakit yang disebabkan oleh virus tersebut dinamakan CoronaVirus disease 2019(COVID-19). Dampak dari terjadinya pandemic ini membuat banyak kekacauan di berbagai belahan dunia, tidak terkecuali Indonesia, pada awal bulan Maret 2020, Presiden Joko Widodo menyampaikan secara langsung temuan kasus COVID-19 di Indonesia.

Banyak penutupan sekolah, fasilitas umum, pembatasan moda transportasi, pelayanan masyarakat, dan juga berbagai macam hal lainnya. Di Indonesia sendiri, pemerintah mengeluarkan kebijakan Pembatasan Sosial Berskala Besar (PSBB).

Setiap individu memiliki pendapat dan opini yang berbeda-beda mengenai COVID-19. Pendapat ini sangat penting dan juga merupakan salah satu hal yang mempengaruhi perilaku utama manusia. Sentimen analisis (sentiment analysis) atau yang juga sering disebut sebagai opinion mining merupakan bidang studi yang menganalisa opini masyarakat, sentimen, evaluasi, penilaian, sikap, dan emosi terhadap sebuah produk, pelayanan, organisasi dan perhimpunan, seorang tokoh, dan isu atau masalah serta peristiwa yang terjadi pada masyarakat itu sendiri. Sentimen analisis banyak digunakan di berbagai bidang dalam kehidupan sehari-hari. Seperti misalnya dalam bidang bisnis, sentimen analisis digunakan untuk memprediksi harga saham, menganalisis kebutuhan pasar, hingga penyusunan strategi pemasaran berdasarkan sentimen masyarakat. Sentimen analisis juga digunakan untuk mengetahui opini publik terkait isu, kejadian, dan juga peristiwa yang terjadi di masyarakat.

Dalam proyek kali ini kami melakukan sentimen analisis pada tweets mengenai Corona Virus. Banyak peristiwa, opini, pendapat masyarakat yang banyak disampaikan di twitter, sehingga dengan memanfaatkan tweets dari twitter mengenai corona virus tersebut kita dapat menganalisis baik itu komentar positif maupun komentar negative.

2. METODE

2.1. Pengumpulan Data

Tahap yang pertama adalah kita melakukan scrapping data dari tweets mengenai Corona Virus. Dalam proyek ini kami menggunakan dataset yang telah tersedia di internet. Untuk link dataset sebagai berikut :

<https://github.com/rajeshmore1/Capstone-Project-2/blob/main/Copy%20of%20Coronavirus%20Tweets.csv>

2.2. Pengenalan Data

Pada tahap ini kita mencari jumlah tiap sentiment, kemudian diikuti dengan mencari sentiment yang bernilai extremely positive serta tweets dalam satu hari, hal itu dilakukan untuk mengenali data apa saja yang ada didalam dataset tersebut.

2.3. Data Cleansing

Data cleansing adalah proses analisa kualitas dari suatu data dengan cara mengubah, mengoreksi, atau menghapus data-data yang salah, tidak lengkap, tidak akurat, atau memiliki format yang salah dalam basis data guna menghasilkan data berkualitas tinggi.

Dalam tahap ini dataset yang sudah didapatkan melalui tahap scrapping data pada corona virus tweets akan dilakukan sortir beberapa data, antara lain :

- Mengubah format kolom TweetAt menjadi date
- Menghilangkan kata yang didalamnya terdapat '@' dan 'https'
- Data Preprocessing menghilangkan nomor, stop words(konjungsi)
- Stem data untuk menghilangkan data yang mirip
- Merubah format data menjadi Lowercase
- Menghapus baris yang kosong (na)

2.4. Text Analysis

Kita mencari berapa banyak suatu kata muncul pada dataset yang kita miliki. Hal ini bertujuan supaya kita tahu kata apa saja yang sering digunakan oleh pengguna twitter yang membahas mengenai corona virus.

2.5. Sentimen analysis

Untuk proses sentimen kita hanya melakukan pengelompokan kata mana yang masuk kedalam sentiment positif, negative, extremely positive, extremely negative, dan neutral.

2.6. Menampilkan hasil klasifikasi

Shiny adalah paket R yang memungkinkan pengguna membuat aplikasi web interaktif. Alat ini membuat aplikasi web setara HTML dari kode Shiny. Kami mengintegrasikan kode HTML dan CSS asli dengan fungsi R Shiny untuk membuat aplikasi rapi. Shiny menggabungkan kekuatan komputasi R dengan interaktivitas web modern. Shiny membuat aplikasi web yang diterapkan di web menggunakan server Anda atau layanan hosting R Shiny.

Pada proyek kali ini kami membuat suatu GUI dengan menggunakan Shiny dengan beberapa menu, dengan menu tersebut pengguna dapat melihat hasil dari sentiment analisis mengenai corona virus.

3. HASIL DAN PEMBAHASAN

3.1. Data Covid-19 tweets

	UserName	ScreenName	Location	TweetAt	OriginalTweet	Sentiment
1	3799	48751	London	16-03-2020	@MeNyrbie @Phil_Gahan @Chrisitv https://t.co/Fz9FAn2Pa...	Neutral
2	3800	48752	UK	16-03-2020	advice Talk to your neighbours family to exchange phone n...	Positive
3	3801	48753	Vagabonds	16-03-2020	Coronavirus Australia: Woolworths to give elderly, disabled ...	Positive
4	3802	48754		16-03-2020	My food stock is not the only one which is empty... PLEASE, ...	Positive
5	3803	48755		16-03-2020	Me, ready to go at supermarket during the #COVID19 outbr...	Extremely Negative
6	3804	48756	ÄoeT: 36.319708,-82.363649	16-03-2020	As news of the region's first confirmed COVID-19 case ca...	Positive
7	3805	48757	35.926541,-78.753267	16-03-2020	Cashier at grocery store was sharing his insights on #Covid...	Positive
8	3806	48758	Austria	16-03-2020	Was at the supermarket today. Didn't buy toilet paper. #Reb...	Neutral
9	3807	48759	Atlanta, GA USA	16-03-2020	Due to COVID-19 our retail store and classroom in Atlanta ...	Positive
10	3808	48760	BHAVNAGAR,GUJRAT	16-03-2020	For corona prevention,we should stop to buy things with th...	Negative
11	3809	48761	Makati, Manila	16-03-2020	All month there hasn't been crowding in the supermarkets o...	Neutral
12	3810	48762	Pitt Meadows, BC, Canada	16-03-2020	Due to the Covid-19 situation, we have increased demand f...	Extremely Positive
13	3811	48763	Horningssea	16-03-2020	#horningssea is a caring community. Let's ALL look after th...	Extremely Positive
14	3812	48764	Chicago, IL	16-03-2020	Me: I don't need to stock up on food, I'll just have Amazon ...	Positive
15	3813	48765		16-03-2020	ADARA Releases COVID-19 Resource Center for Travel Brand...	Positive

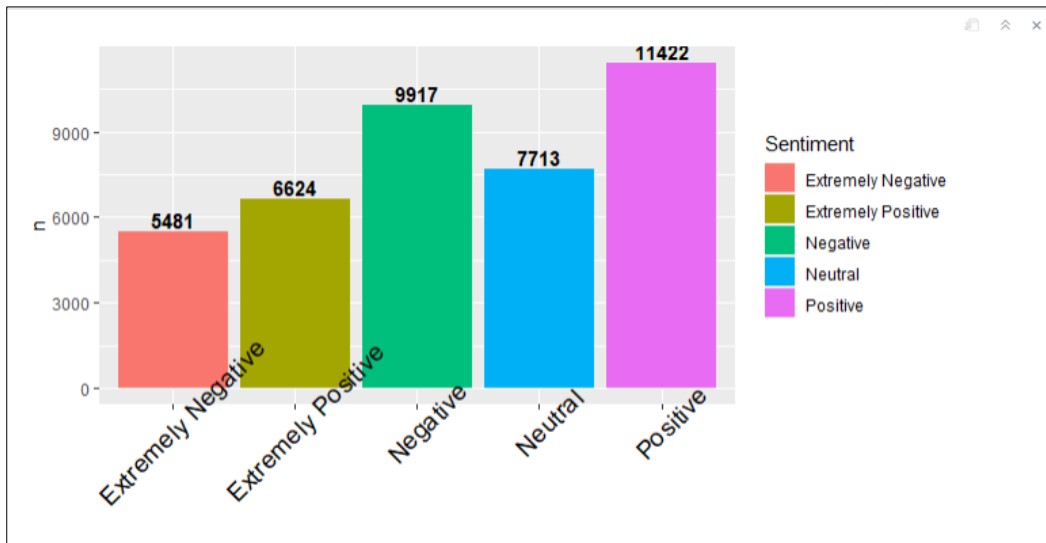
Gambar 3.1 Dataset Covid-19 tweets

Data diatas merupakan data tweets dari Twitter yang berisi tweets mengenai Covid-19. Didalamnya terdapat beberapa kolom, antara lain:

- Username
- ScreenName
- Location
- TweetAt

- OriginalTweet
- Sentiment

3.2. Melihat jumlah tiap sentiment



Gambar 3.2 Plot jumlah tiap sentimen

Pada proses tersebut digunakan untuk melakukan pengelompokan data, berupa jumlah masing-masing tiap sentimen yang akan dianalisis.

3.3. Mengambil sentimen yang bernilai extremely positive

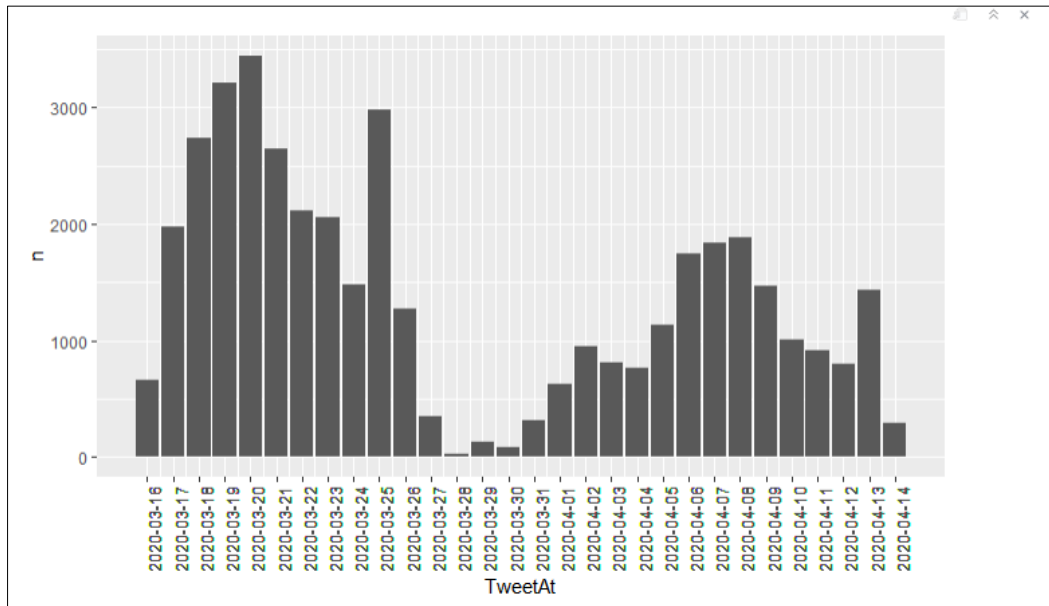
User	ScreenName	Location	TweetAt	OriginalTweet	Sentiment	Tweets
1	3810	48762 Pitt Meadows, BC, Canada	2020-03-16	Due to the Covid-19 situation, we have increased demand f...	Extremely Positive	Due to the Covid-19 we have increased demand for all food...
2	3811	48763 Horningssea	2020-03-16	#horningssea is a caring community. Let's ALL look after th...	Extremely Positive	is a caring Let's ALL look after the less capable in our villa...
3	3817	48769 North America	2020-03-16	Amazon Glitch Stymies Whole Foods, Fresh Grocery Deliveri...	Extremely Positive	Amazon Glitch Stymies Whole Fresh Grocery Deliveries A'As...
4	3821	48773	2020-03-16	We have AMAZING CHEAP DEALS! FOR THE #COVID2019 g...	Extremely Positive	We have AMAZING CHEAP FOR THE going on to help And ...
5	3822	48774	2020-03-16	We have AMAZING CHEAP DEALS! FOR THE #COVID2019 g...	Extremely Positive	We have AMAZING CHEAP FOR THE going on to help And ...
6	3824	48776 London	2020-03-16	UK #consumer poll indicates the majority expect #covid19's...	Extremely Positive	UK poll indicates the majority expect impact to last 4-12 mo...
7	3830	48782	2020-03-16	'Hole' Foods... images from the nicest grocery store in one ...	Extremely Positive	from the nicest grocery store in one of the richest neighbor...
8	3832	48784 Virginia, USA	2020-03-16	Coronavirus fun fact: if you cough at the grocery store, you ...	Extremely Positive	Coronavirus fun if you cough at the grocery you get the wh...
9	3839	48791 Fort Worth, Texas	2020-03-16	We're here to provide a safe shopping experience for our cu...	Extremely Positive	We're here to provide a safe shopping experience for our cu...
10	3845	48797 Drogheda	2020-03-16	Please Share. Know someone who's 65 Living on their own ...	Extremely Positive	Please Share Know someone who's 65 Living on their own s...
11	3847	48799 Arkansas, USA	2020-03-16	Never thought I'd say this, but, 2019. Will you come back? ...	Extremely Positive	Never thought I'd say Will you come
12	3854	48806 West Midlands, England	2020-03-16	Morning everyone have a great and safe day. ??? #coronavir...	Extremely Positive	Morning everyone have a great and safe
13	3857	48809	2020-03-16	With the outbreak of Covid-19 in entire world, the retail sho...	Extremely Positive	With the outbreak of Covid-19 In entire the retail shops in ...
14	3862	48814 Dallas, TX	2020-03-16	So, are we feeling like it's ethical to still do stuff like order d...	Extremely Positive	are we feeling like it's ethical to still do stuff like order deliv...
15	3874	48826 MA & New England by way of CA	2020-03-16	Lost wages either due to illness from 19 or to the virus eco...	Extremely Positive	Lost wages either due to illness from 19 or to the virus econ...
16	3888	48840	2020-03-16	The fact that canned food, toxic chemicals and store bought...	Extremely Positive	The fact that canned toxic chemicals and store bought hand...
17	3889	48841 Houston, TX	2020-03-16	Just called mum and dad in UK (over 70). They are great but...	Extremely Positive	Just called mum and dad in UK They are great but I offered ...
18	3892	48844 Leicester, England	2020-03-16	Pretty sure within a week or two, supermarket supply chains...	Extremely Positive	Pretty sure within a week or supermarket supply chains will ...
19	3894	48846 FREE goal setting guide ??	2020-03-16	Worried about COVID-19? I more worried about people pa...	Extremely Positive	Worried about I more worried about people Having a You're...
20	3901	48853	2020-03-16	THANK YOU To all the grocery store employees for working ...	Extremely Positive	THANK YOU To all the grocery store employees for working ...
21	3904	48856	2020-03-16	You know it's the end of the world when @BootsUK Surgical...	Extremely Positive	You know it's the end of the world when Surgical Spirit is ou...
22	3920	48872	2020-03-16	I love this ?? all grocery stores should have a set hour (prefe...	Extremely Positive	I love this all grocery stores should have a set hour in the fo...
23	3923	48875 Wakefield, England	2020-03-16	Currently self isolating as I'm symptomatic. Thinking how t...	Extremely Positive	Currently self isolating as I'm Thinking how to support eid...
24	3929	48881 United States	2020-03-16	FYI It's called the not the Chinese Virus Notice nasty mentio...	Extremely Positive	FYI It's called the not the Chinese Virus Notice nasty mentio...
25	3932	48884 Michigan	2020-03-16	I HAPPY M O N D A Y Friends ! - I know everyone is uneasy abo...	Extremely Positive	HAPPY M O N D A Y Friends I know everyone is uneasy abo...
26	3937	48889 Dunedin City, New Zealand	2020-03-16	Spark's Manila call centre is closing to help prevent the spre...	Extremely Positive	Spark's Manila call centre is closing to help prevent the spre...

Gambar 3.3 Tweets dengan sentiment extremely positive

Setelah melihat jumlah dari masing-masing sentiment maka selanjutnya yakni mengambil sentiment dari data yang memiliki nilai extremely positive.

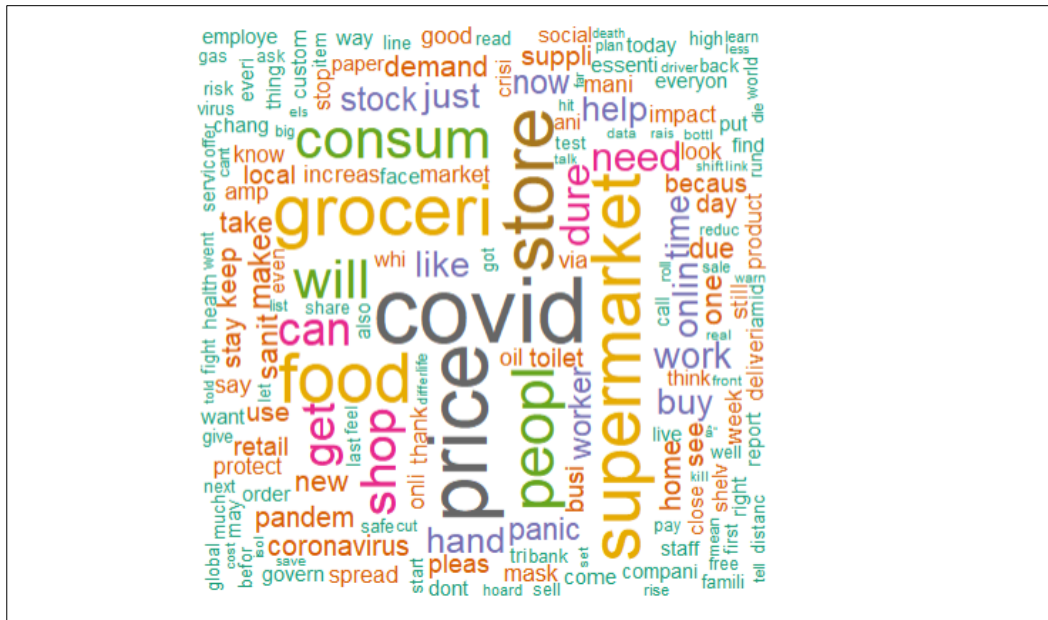
3.4. Mencari jumlah tweets tiap hari

Dari data kemudian diolah agar dapat menampilkan plot yang menampilkan jumlah tweets setiap hari, seperti dalam gambar dibawah pada sumbu x terdapat tanggal dimana tweets tersebut diunggah, kemudian pada sumbu y terdapat jumlah tweetsnya.



Gambar 3.4 Plot jumlah tweets tiap hari

3.5. Word Cloud



Gambar 3.5 Word Cloud

Gambar diatas merupakan gambaran kata yang paling banyak muncul, semakin besar ukuran font maka semakin banyak pula orang yang menulis kata tersebut di twitter.

3.6. Tampilan GUI Shiny

UserName	ScreenName	Location	TweetAt	OriginalTweet	Sentiment	Tweets
3799	48751	London	2020-03-16	@MeNyrble @Phil_Gahan @Chrisitv https://t.co/IFz9FAn2Pa and https://t.co/xX6ghGFzCC and https://t.co/l2NlzdXNo8	Neutral	and and
3800	48752	UK	2020-03-16	advice Talk to your neighbours family to exchange phone numbers create contact list with phone numbers of neighbours schools employer chemist GP set up online shopping accounts if poss adequate supplies of regular meds but not over order	Positive	advice Talk to your neighbours family to exchange phone numbers create contact list with phone numbers of neighbours schools employer

Gambar 3.6 Tampilan GUI dengan Shiny

Dalam gui tersebut terdapat navigation bar yang akan memudahkan pengguna untuk memilih menu yang akan dipilih. User dapat memilih data yang akan dilihat, apakah jumlah sentiment, tweets harian, lokasi tweets, jumlah kata dan world cloud .

4. KESIMPULAN

Proyek ini melakukan analisis sentimen terhadap Covid-19 tweets, karena dapat menentukan sentimen skala yang tepat untuk review yang negatif maupun yang positif. dari data yang ada. Kemudian harapannya dengan adanya sentiment analisis tersebut dapat membantu dalam melakukan analisa-analisa selanjutnya.