

Analisis Sentimen Twitter pada #Indonesiaterserah

I Gusti Naufal Daffa Adyana¹, Rizki Indah Pratiwi²

^{1,2} *Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jakarta*

Jln. Rs. Fatmawati, Pondok Labu, Jakarta Selatan, DKI Jakarta, 12450, Indonesia

¹nofaldaffa@gmail.com, ²rizkiindah8899@gmail.com

Intisari – Menyebarnya wabah Covid-19 yang semakin meluas di Indonesia, menyebabkan pemerintah menerapkan beberapa kebijakan yang ditunjukan untuk memutus persebaran wabah tersebut. Mulai dari pemberlakuan Pembatasan Sosial Berskala Besar (PSBB), penetapan kebijakan *Work From Home* (WFH), belajar dirumah, ibadah dirumah, hingga *physical distancing*. Dimana penetapan kebijakan tersebut menyebabkan semua kegiatan masyarakat harus dilaksanakan dari rumah, serta adanya pembatasan kegiatan yang ada diluar rumah. Namun pada beberapa waktu lalu, sempat muncul topik yang ramai diperbincangkan oleh publik. Topik tersebut ditandai dengan #Indonesiaterserah pada *tweet* yang disampaikan oleh pengguna Twitter. Dimana topik tersebut muncul setelah dibagikan oleh para tenaga medis yang sudah “pasrah” dengan segala sikap masyarakat yang cenderung mengabaikan segala kebijakan yang telah ditetapkan, sehingga terus menyebabkan adanya peningkatan pada kasus persebaran Covid-19. Berdasarkan permasalahan tersebut, penulis bertujuan melaksanakan penelitian untuk menilai opini publik dengan melakukan analisis sentimen. Analisis sentimen akan dilakukan dengan memanfaatkan Twitter sebagai media sosial yang memberikan dampak besar dalam membangun opini, pandangan, dan sentimen publik. Dimana proses penarikan data yang digunakan untuk analisis sentiment akan didapatkan melalui *Twitter API* dengan menggunakan bahasa pemrograman R. Data yang digunakan yaitu berasal dari *tweet* yang mengandung #indonesiaterserah, dimana *tweet* yang diambil memiliki jumlah total 1000 *tweet*. Dari proses analisis sentimen yang dilakukan, kita dapat mengetahui respon publik mengenai topik dengan #Indonesiaterserah. Berdasarkan hasil penelitian yang telah dilakukan menunjukan bahwa sebagian besar opini yang disampaikan publik melalui Twitter cenderung menunjukan sentimen negatif. Hasil skor sentimen negatif memiliki nilai yang lebih tinggi dibandingkan dengan hasil sentimen positif dan netral. Dimana sentimen netral memiliki skor sebanyak 115, sedangkan untuk sentimen positif hanya memiliki skor 45 dan sentimen negatif memiliki skor tertinggi yaitu 119.

Kata Kunci – Covid-19, Work From Home, Tenaga Medis, Analisis Sentimen, Twitter API, Pemrograman R.

Abstract - The spread of the Covid-19 epidemic which is increasingly widespread in Indonesia, has caused the government to implement several policies aimed at deciding the spread of the outbreak. Starting from the implementation of Large-Scale Social Restrictions (PSBB), the establishment of Work From Home (WFH) policies, study at home, worship at home, to physical distancing. Where the stipulation of the policy causes all community activities to be carried out from home, and there are restrictions on activities that exist outside the home. But some time ago, a topic which had been widely discussed was publicly discussed. The topic is marked with #Indonesiaterserah on tweets delivered by Twitter users. Where the topic arose after being distributed by medical personnel who had "resigned" to all the attitudes of the community who tended to ignore all the policies that had been set, so that it continued to cause an increase in the case of the spread of Covid-19. Based on these problems, the authors aim to carry out research to assess public opinion by analyzing sentiments. Sentiment analysis will be carried out by utilizing Twitter as a social media that has a large impact in building public opinion, views and sentiments. Where the process of withdrawing data used for sentiment analysis will be obtained through Twitter API using the R programming language. The data used is derived from tweets containing #indonesiaterserah, where tweets taken have a total of 1000 tweets. From the sentiment analysis process, we can find out the public's response on the topic with #Indonesiaterserah. Based on the results of research that has been done shows that the majority of opinions expressed publicly via Twitter tend to show negative sentiment. The results of negative sentiment scores have a higher value than the results of positive and neutral sentiments. Where neutral sentiment has a score of 115, while for positive sentiment only has a score of 45 and negative sentiment has the highest score of 119.

Keywords - Covid-19, Work From Home, Medical Personnel, Sentiment Analysis, Twitter API, R. Programming.

I. PENDAHULUAN

Munculnya virus *corona* jenis baru yang disebut dengan Covid-19 pada akhir tahun 2019 lalu menyebabkan persebaran yang sangat luas ke berbagai negara. Indonesia termasuk kedalam salah satu negara yang terkena dampak dari persebaran virus tersebut. Dimana pada awal kemunculan kasus positif Covid-19 hingga saat ini, terhitung kurang lebih sudah 3 bulan lamanya persebaran Covid-19 terjadi. Belum ada tanda-tanda yang mengindikasikan bahwa persebaran virus tersebut akan segera berakhir. Hal tersebut memberikan dampak yang sangat besar diberbagai bidang kehidupan.

Berbagai lapisan masyarakat yang terdiri dari berbagai macam latar belakang profesi pun turut merasakan dampak dari persebaran Covid-19. Terutama para tenaga medis yang bersentuhan langsung untuk menangani pasien positif. Dalam kurun waktu 3 bulan terakhir, banyak kisah-kisah perihal perjuangan dan pengorbanan dari tenaga medis yang kita lihat dan dengar. Dengan terus meningkatnya kasus persebaran Covid-19, maka akan membuat tenaga medis bekerja lebih keras lagi untuk upaya perawatan pasien. Sehingga tidak sedikit tenaga medis yang gugur setelah dinyatakan terpapar Covid-19, ataupun hal lain yang masih berkaitan dengan tugas mereka dalam upaya melawan virus corona.

Akan tetapi akhir-akhir ini banyak beberapa masyarakat yang justru mengabaikan segala kebijakan yang telah ditetapkan oleh pemerintah dengan tidak mematuhi protokol kesehatan, tidak melakukan *social distancing*, selalu melakukan kegiatan yang tidak penting di luar rumah. Hal tersebut membuat tenaga medis “pasrah” dengan segala sikap masyarakat yang cenderung mengabaikan segala kebijakan yang telah ditetapkan, sehingga terus menyebabkan adanya peningkatan pada kasus persebaran Covid-19. Sehingga banyak dari tenaga medis yang membagikan keluh kesahnya melalui topik dengan menyertakan #Indonesiaterserah pada berbagai sosial media. Beberapa masyarakat yang merasa peduli pun turut serta mendukung tenaga medis, sekaligus menghimbau kepada masyarakat lainnya dengan menggunakan #Indonesiaterserah yang kemudian diunggah di berbagai sosial media yang mereka miliki.

Dengan banyaknya kemunculan topik tersebut, peneliti bertujuan untuk mengetahui bagaimana opini publik mengenai topik yang mengandung #indonesiaterserah. Dimana untuk mengetahui kecenderungan opini publik, maka peneliti akan melakukan proses analisis sentimen. Menurut Dang dkk (2010), analisis sentimen adalah metode untuk menganalisis sebagian data untuk mengetahui emosi manusia. Dimana

proses analisis sentimen akan menggunakan data yang diambil dari Twitter melalui *Twitter API* dengan menggunakan bahasa pemrograman R. Pemilihan media sosial Twitter pada penelitian ini didasari karena alasan bahwa Twitter menyediakan data yang bisa diakses secara bebas dengan menggunakan *Twitter API*. *Twitter API (Application Programming Interface)* merupakan sejumlah fungsi yang dapat digunakan pengembang perangkat lunak untuk mengolah data saat membangun perangkat lunak (Rustiana & Rahayu 2017).

II. METODE PENELITIAN

Data yang akan digunakan dalam melakukan proses analisis sentimen memiliki jumlah total 1000 *tweet*. Penelitian ini berfokus pada *tweet* yang memiliki topik dengan menyertakan #Indonesiaterserah. Dari proses analisis sentimen yang dilakukan, nantinya kita dapat mengetahui respon publik mengenai topik yang mengandung #Indonesiaterserah, selanjutnya data yang sudah diperoleh akan diproses menggunakan bahasa pemrograman R. Berikut merupakan tahapan yang digunakan dalam proses penelitian yang dilakukan yaitu :

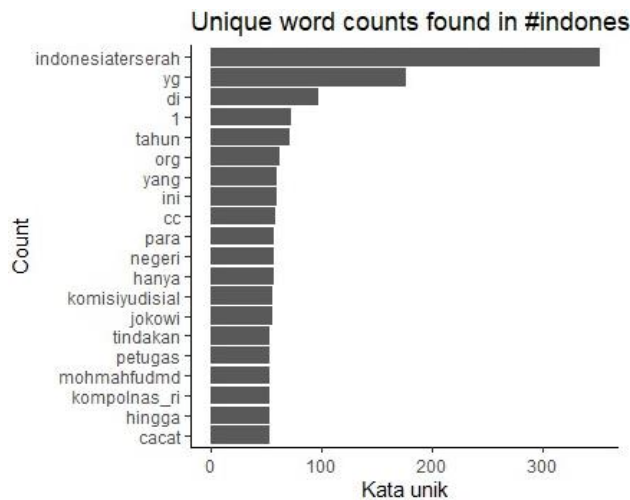
- a) *Input data*, dengan menyiapkan semua *library* yang berkaitan untuk menjalankan setiap proses didalam program yang telah dibuat. Selanjutnya mendapatkan *key* untuk *Twitter API* agar dapat mengakses dan mengambil data yang dibutuhkan dari Twitter.
- b) Membuat *corpus* untuk setiap set *tweet* yang telah didapat dari proses sebelumnya.
- c) Praproses data ke-1 dimana data teks yang telah didapatkan, akan dilakukan proses penghapusan elemen *http* yang ada didalamnya, *tokenisasi*, serta menghilangkan *stopwords* dari data teks tersebut.
- d) Visualisasi, yaitu dengan menampilkan visualisasi daftar kata teratas dari jumlah kata yang memiliki sentimen positif dan negatif.
- e) Praproses data ke-2, dimana akan dilakukan proses *data cleaning* pada *tweet* yang diperoleh. Beberapa proses yang dijalankan pada proses ini yaitu dengan menghapus data duplikat, menghapus data pada baris *index* ke-2 sampai dengan ke-6, mengimport kamus bahasa *slang*, mengganti (*replace*) internet *slang*, lalu melakukan *stemming*.
- f) Proses *tokenisasi* pada *tweet* yang telah melewati proses *data cleaning* sebelumnya.
- g) Data yang dihasilkan akan dibentuk ke dalam *wordcloud*, sehingga akan memudahkan dalam melakukan proses frekuensi analisis nantinya.

- h) Selanjutnya lakukan proses *import* kata positif dan negatif, proses ini bertujuan untuk mempermudah dalam proses analisis sentimen dari data yang sudah dihasilkan.
- i) Proses perhitungan nilai sentimen dari setiap *tweet*.
- j) Visualisasi histogram.

III. HASIL DAN PEMBAHASAN

Penelitian dilakukan dengan melakukan registrasi agar dapat memperoleh *key Twitter API*, hal tersebut dilakukan agar penguji dapat mengambil data *tweet* yang dibutuhkan. Total data yang diambil yaitu berjumlah 1000 *tweet*, serta *tweet* yang digunakan adalah *tweet* dengan menyertakan #Indonesiaterserah. Proses yang dilakukan dalam menganalisis sentimen mengenai #indonesiaterserah dirancang dengan menggunakan pemrograman R, dimana dalam eksekusi program yaitu dilakukan dengan menjalankan beberapa tahapan.

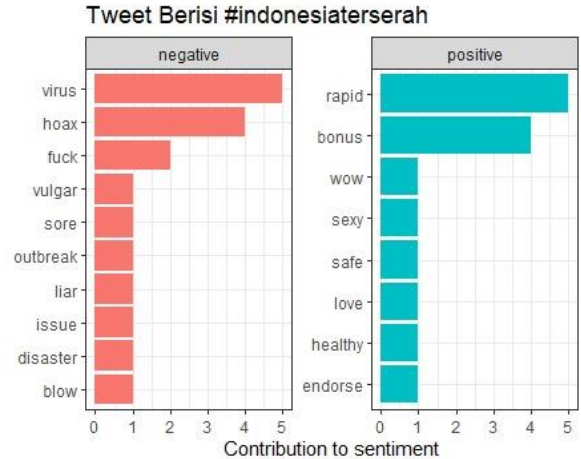
Beberapa tahapan yang dilalui diantaranya yaitu dengan membuat *corpus* untuk setiap *tweet* yang telah diambil, selanjutnya dilakukan praproses data pertama. Dimana pada praproses data pertama yaitu dengan melakukan penghapusan elemen *http* pada *tweet*, *tokenisasi*, serta menghilangkan *stopwords* yang ada. Praproses data pertama berpengaruh pada hasil visualisasi kata teratas pada *tweet* yang diamati. Berikut merupakan hasil visualisasi dari kata teratas yang didapatkan.



Gambar 1. Hasil Visualisasi daftar kata teratas pada *tweet* yang menyertakan #Indonesiaterserah.

Berdasarkan *output* diatas, menampilkan hasil visualisasi data yang ditampilkan berupa 20 daftar kata teratas pada *tweet* yang telah diambil dengan menyertakan #indonesiaterserah, dimana kata “indonesiaterserah” menempati posisi teratas.

Proses selanjutnya yang dilakukan yaitu dengan menganalisis sentimen dengan menggunakan *Bing*, dimana proses tersebut berhubungan untuk membandingkan emosi positif dan negatif. Hasil tersebut diperoleh setelah melakukan proses filterisasi serta memplot kata-kata yang telah melalui proses sebelumnya. Berikut merupakan hasil visualisasi dari proses yang dilakukan.



Gambar 2. Hasil visualisasi dari perbandingan emosi positif dan negatif

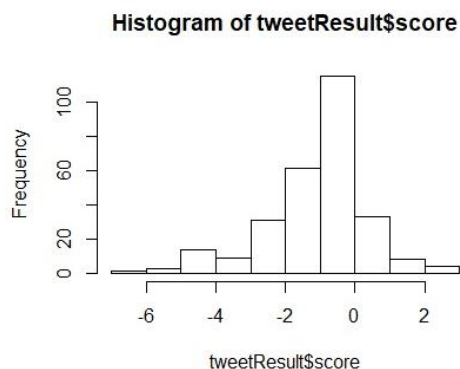
Setelah mendapatkan hasil dari 20 kata teratas serta hasil perbandingan emosi positif dan negatif dari *tweet* yang berisi #indonesiaterserah. Proses selanjutnya yang akan dilakukan yaitu menjalankan praproses data yang kedua. Proses yang dilakukan pada praproses kedua diantaranya yaitu melakukan *data cleaning* pada *tweet* yang telah diperoleh, beberapa proses yang dijalankan pada tahapan ini diantaranya yaitu dengan menghapus data duplikat, menghapus data pada baris index ke-2 sampai dengan ke- 6, mengimport kamus bahasa *slang*, mengganti (*replace*) internet *slang*, lalu melakukan *stemming*. Praproses data kedua berpengaruh terhadap hasil *wordcloud* yang berfungsi sebagai perhitungan analisis frekuensi untuk kata yang paling banyak muncul. Dimana semakin besar ukuran tulisan yang ditampilkan, menunjukkan bahwa kata tersebut memiliki frekuensi kemunculan kata yang besar (kata yang paling sering muncul). Berikut merupakan hasil visualisasi *wordcloud*.



Gambar 3. Hasil visualisasi wordcloud.

Berdasarkan hasil *wordcloud* yang ditampilkan, dapat dianalisis bahwa kata-kata seperti “orang”, “face”, “covid”, “normal” memiliki frekuensi yang tinggi dalam topik yang dibahas pada *tweet* #indonesiaterserah. Sedangkan kata lain seperti “nih”, “kalo”, “sih”, “makan”, “main” menjadi urutan frekuensi lebih rendah yang muncul dalam analisis, sedangkan sisa hasil kata lainnya yang berukuran lebih kecil seperti “test”, “jaga”, “cari”, “lucu”, “sehat”, “beli” menunjukkan frekuensi kemunculan yang lebih kecil dibandingkan kata-kata dengan ukuran tulisan yang lebih besar.

Setelah *wordcloud* terbentuk, maka proses selanjutnya yang akan dilakukan yaitu dengan menampilkan histogram. Dimana histogram akan menunjukkan hasil dari proses sentimen analisis yang telah dilakukan. Berikut merupakan hasil visualisasinya.



Gambar 4. Hasil visualisasi sentimen analisis.

Berdasarkan hasil histogram yang diperoleh, maka hasil akan dikelompokkan menjadi sentimen yang bersifat netral, positif, dan negatif. Berikut merupakan rincian hasil rincian dari setiap skor sentimen yang didapatkan.

```
> count(tweetResult$score)
```

	x	freq
1	-7	1
2	-5	3
3	-4	14
4	-3	9
5	-2	31
6	-1	61
7	0	115
8	1	33
9	2	8
10	3	4

Gambar 5. Hasil rincian dari setiap skor sentimen.

Berdasarkan hasil output diatas menunjukkan bahwa skor dari sentimen negatif memiliki nilai yang lebih tinggi dibandingkan dengan hasil sentimen positif dan netral. Dimana sentimen netral memiliki skor sebanyak 115, sedangkan untuk sentimen positif hanya memiliki skor 45 dan sentimen negatif memiliki skor tertinggi yaitu 119.

IV. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan mengenai analisis sentimen Twitter pada *tweet* yang menyertakan #Indonesiaterserah, dengan data yang digunakan berasal dari Twitter melalui *Twitter API* menunjukkan bahwa hasil sentimen yang ditunjukkan oleh publik cenderung negatif. Dimana proses analisis dilakukan dengan menggunakan pemrograman R, mulai dari proses pengambilan data *tweet* sebanyak 1000 *tweet*, hingga proses perhitungan serta proses visualisasi histogram dari hasil sentimen analisis.

Pada hasil akhir menunjukkan bahwa skor dari sentimen negatif memiliki nilai yang lebih tinggi dibandingkan dengan hasil sentimen positif dan netral. Dimana sentimen netral memiliki skor sebanyak 115, sedangkan untuk sentimen positif hanya memiliki skor 45 dan sentimen negatif memiliki skor tertinggi yaitu 119.

DAFTAR PUSTAKA

Pratama, Septian F., Riesa Andrean, & Aryo Nugroho. (2019). Analisis Sentimen Twitter Dabat Calon Presiden Indonesia Menggunakan Metode Fined-Grained Sentiment Analysis. *Journal of Information Technology and Computer Science*, 4, 39-43.

Hartanto. (2017). Text Mining dan Sentimen Analisis Twitter pada Gerakan LGBT. *Jurnal Psikologi Ilmiah*, 9(4), 18-25.