



**ANALISIS RESPON MASYARAKAT TERHADAP
VAKSINASI COVID-19**

DIUSULKAN OLEH:

| | |
|-----------------------------------------|--------------------|
| Gloria Gabriel Regar | 20021106088 |
| Mentari Estefin Milania Pangkey | 20021106015 |
| Crysanti Andressa Angela Harjono | 20021106020 |

ABSTRAK

Dalam penelitian ini, kami mengambil data teks dari sosial media Twitter menggunakan metode *web scraping*. Berdasarkan kata kunci “VaksinasiCovid19” data yang dikumpulkan berjumlah 1500 *tweets*. Dari seluruh *tweets* yang terkumpul, kami menghapus beberapa *tweets* yang terduplikasi, sehingga jumlah *tweets* yang akan diproses lanjut berjumlah 1190 *tweets*. Data kemudian mengalami *preprocessing* yang melibatkan tahapan *Case Folding*, *Cleansing*, *Stemming*, *Stopword Removal*, *Tokenization*. Setelah itu data di analisis dalam 3 kategori sentimen. Berdasarkan hasil analisis, kami mendapati bahwa 1088 *tweets* atau setara 91.43% merupakan sentimen netral, 69 *tweet* atau 5.80% dikategorikan sebagai sentimen positif, sedangkan 33 *tweets* setara dengan 2.77% dikelompokkan sebagai sentimen negatif.

Kata kunci: Analisis Sentimen, Twitter, Vaksinasi, COVID-19

DAFTAR ISI

| | |
|--------------------------------|----|
| ABSTRAK | i |
| DAFTAR ISI | ii |
| 1. PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Tujuan | 1 |
| 1.3 Manfaat | 2 |
| 1.4 Batasan..... | 2 |
| 2. STUDI LITERATUR..... | 2 |
| 3. METODOLOGI | 3 |
| 3.1 Pengumpulan Data | 3 |
| 3.2 Metode Preprocessing | 3 |
| 3.3 Sentimen Analisis | 4 |
| 4. HASIL DAN PEMBAHASAN | 5 |
| 5. KESIMPULAN | 8 |
| DAFTAR PUSTAKA..... | 9 |

1. PENDAHULUAN

1.1 Latar Belakang

COVID-19 merupakan penyakit yang disebabkan oleh *SARS-CoV-2* dan menular yang menginfeksi saluran pernafasan pada manusia [1]. Pada Desember 2019, kasus pertama COVID-19 terjadi di Wuhan, Provinsi Hubei, China [1]. Hanya pada waktu beberapa bulan, penularan virus ini telah menyebar ke berbagai negara di dunia termasuk Indonesia.

Pemerintah Indonesia telah bekerja keras demi menekan penyebaran virus korona guna mengendalikan dampak negatifnya, salah satu upaya pemerintah yaitu dengan melakukan vaksinasi. Vaksinasi saat ini terus dilaksanakan oleh pemerintah lewat Program Vaksinasi Nasional, terdapat empat vaksin yang digunakan dalam program ini, yaitu *Sinovac*, *AstraZeneca*, *Pfizer*, dan *Novavax* [2]. Penggunaan keempat vaksin ini mendapat berbagai tanggapan dan opini dari masyarakat. Ada ungkapan pendapat yang baik dan konstruktif, tetapi ada juga kontradiksi dan penolakan.

Media sosial Twitter adalah tempat di mana orang dapat mengekspresikan pendapat mereka secara terbuka. Analisis sentimen didasarkan pada opini dan mengklasifikasikan emosi pribadi seperti kesedihan, kebahagiaan atau kepanikan dalam menghadapi vaksinasi COVID-19 ke dalam beberapa tingkatan yaitu positif, negatif, dan netral.

Menyadari pentingnya untuk mengetahui respon masyarakat Indonesia terhadap program vaksinasi COVID-19, maka kami melakukan penelitian singkat dengan melakukan penambangan data teks (Bahasa Indonesia) pada media sosial Twitter. Hasil dari penelitian ini diharapkan dapat membantu bangsa Indonesia untuk bangkit dari pandemi COVID-19 menuju Indonesia Maju.

1.2 Tujuan

Dalam penelitian ini kami bertujuan menggunakan metode *web scraping* untuk mengambil data dan melakukan analisis sentimen berdasarkan tren vaksinasi COVID-19 dari sosial media Twitter. Data yang di ambil akan diolah kemudian di analisis dan akan digolongkan ke dalam beberapa kategori sentimen yaitu: negatif, positif atau netral.

1.3 Manfaat

Manfaat yang dapat diperoleh dari penelitian ini adalah dapat mengetahui opini masyarakat tentang vaksinasi yang sementara dilakukan di Indonesia serta respon masyarakat setelah vaksinasi dilakukan dan sebelum dilakukan. Dengan melakukan analisis sentimen kita dapat mengkategorikan opini-opini dari masyarakat. Kemudian, opini tersebut dapat digolongkan ke dalam beberapa kategori yaitu: negatif, positif dan netral.

1.4 Batasan

Pada penelitian ini kami melakukan pengamatan kepada setiap respon masyarakat pengguna Twitter terhadap vaksinasi COVID-19. Penelitian ini dibatasi pada satu kata kunci yaitu “VaksinasiCovid19”. Penelitian ini juga dibatasi dengan memanfaatkan metode *web scraping* pada media sosial Twitter. Data yang dikumpulkan kemudian di analisis menggunakan metode *Natural Language Processing* setelah sebelumnya melewati proses *preprocessing* yang bertujuan membersihkan data teks.

2. STUDI LITERATUR

Web scraping merupakan metode pengumpulan data dari web dengan menggunakan teknologi yang ada [3]. *Web scraping* digunakan untuk membersihkan dan mengolah data yang tidak terstruktur dari suatu situs *web* menjadi terstruktur dengan memotong informasi dari halaman *web* berdasarkan skrip yang tersedia [4]. Teknik *web scraping* juga sering disebut sebagai pendekatan ekstraksi data yang melibatkan identifikasi dan pengkodean otomatis dari informasi yang ditemukan di halaman *web* [5].

Analisis Sentimen merupakan proses mengekstraksi, mengolah data teks yang belum terstruktur kemudian dapat diketahui sentimen yang terkandung didalamnya [6]. Sentimen analisis ini adalah proses mengklasifikasi teks ke dalam kategori positif atau negatif [7]. Jika sentimen lebih dari 0 maka *tweets* tersebut dikategorikan sebagai sentimen positif, sedangkan jika nilai sentimen kurang dari 0 maka *tweets* merupakan sentimen negatif, selain itu adalah sentimen netral. nilai tersebut di dapat dari penjumlahan nilai dari kata bersentimen [8].

Tahap *preprocessing* merupakan tahap untuk menyiapkan data mentah sebelum dilakukan proses selanjutnya, tujuan dari *preprocessing* ini dilakukan untuk menghilangkan data yang tidak cocok atau mengubah data agar lebih mudah diolah [9].

3. METODOLOGI

3.1 Pengumpulan Data

Kami mengumpulkan data yang tak terstruktur berbentuk *tweet* bahasa Indonesia yang diperoleh dari hasil *scraping* data yang disediakan Twitter melalui *Application Programming Interface (API)* dengan menggunakan kata kunci “VaksinasiCovid19”.

3.2 Metode Preprocessing

Dalam tahapan data mulai diolah dan diproses. Data akan diambil dari dataset dan kemudian diolah dengan membuang serta memisahkan data yang akan diambil dan data yang akan dibuang sehingga informasi yang akan kita ambil itu berguna. Tahapan ini mempermudah untuk mengambil data yang dibutuhkan. Adapun beberapa macam dalam data *preprocessing* yaitu:

1. *Case Folding*

Tahapan ini untuk membuat text yang awalnya berhuruf kapital akan diubah menjadi huruf kecil seperti kata “Bermanfaat” akan menjadi “bermanfaat” dan segala bentuk karakter akan dianggap sebagai delimiter.

2. *Cleansing*

Tahapan ini untuk mengeliminasi teks atau data yang tidak bernilai dan menyisakan data yang diperlukan.

3. *Stemming*

Pada tahapan ini bertujuan untuk mengubah kata-kata yang memiliki imbuhan menjadi kata awalnya (*root word*).

4. *Stopword Removal*

Tahapan ini bertujuan untuk membuang kata yang tidak penting atau menghapus kata – kata yang tidak memiliki informasi.

5. *Tokenization*

Tahapan dilakukan untuk tujuan membagi teks yang berbentuk kalimat serta paragraf menjadi beberapa bagian. Contohnya, "ayo semua jangan takut di vaksin" menghasilkan enam token, yaitu: "ayo", "semua", "jangan", "takut", "di", "vaksin".

Tabel 3.1 Simulasi Tahapan Proses Data

| Tahapan | Hasil |
|--------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <i>Tweets</i> | Badan pengawas obat dan makanan (BPOM) pada tgl 11 jan'21 telah memberikan persetujuan penggunaan vaksinn! https://t.co/TUgkjkQjFp |
| <i>Case Folding</i> | badan pengawas obat dan makanan bpom pada tgl 11 jan21 telah memberikan persetujuan penggunaan vaksinn httpstcotugkjkqjfp |
| <i>Cleansing</i> | badan pengawas obat dan makanan bpom pada tgl 11 jan21 telah memberikan persetujuan penggunaan vaksinn |
| <i>Stemming</i> | Badan pengawas obat dan makanan bpom pada tgl 11 jan21 telah memberi persetujuan penggunaan vaksin |
| <i>Stopword Removal & Tokenization</i> | ['badan' 'pengawas' 'obat' 'dan' 'makanan' 'pada' 'tgl' '11' 'jan21' 'memberi' 'persetujuan' 'penggunaan' 'vaksin'] |

3.3 Sentimen Analisis

Setelah membersihkan dan memproses data menjadi terstruktur, kemudian data-data tersebut ditentukan sentimen analisisnya dimana kata dalam *tweet* akan diklasifikasikan menjadi 3 kategori yaitu: positif, negatif, dan netral. sehingga pemberian skor pada kalimat digunakan untuk bisa membedakan kategori dari *tweet* yang ada, disini kami mengklasifikasikan skor >0 untuk sentimen positif, <0 untuk sentimen negatif, selain itu merupakan sentimen netral.

| 0 | text | polarity | subjectivity | sentiment | neg | neu | pos | compound | text_len | text_word_count | punct | tokenized | nonstop | stemmed |
|---|---------------------------------------------------------------|----------|--------------|-----------|-------|-------|-------|----------|----------|-----------------|---------------------------------------------------|---------------------------------------------------------|---------------------------------------------------------|------------------------------------------------------|
| 0 | Partisi bilik vaksinasi dengan berbagai ukuran... | 0.0 | 0.0 | neutral | 0.000 | 1.000 | 0.000 | 0.0000 | 118 | 16 | partisi bilik vaksinasi dengan berbagai ukuran... | [partisi, bilik, vaksinasi, dengan, berbagai, ...] | [partisi, bilik, vaksinasi, dengan, berbagai, ...] | [partisi, bilik, vaksinasi, dengan, berbagai, ...] |
| 1 | Update data vaksinasi COVID-19 di Kabupaten Ro... | 0.0 | 0.0 | neutral | 0.000 | 1.000 | 0.000 | 0.0000 | 114 | 18 | update data vaksinasi covid di kabupaten ro... | [update, data, vaksinasi, covid, di, kabupaten, ...] | [update, data, vaksinasi, covid, di, kabupaten, ...] | [updat, data, vaksinasi, covid, di, kabupaten, ...] |
| 2 | RT @TelekungSK: "Nak... Mak teringin nak cucuk vaksin vak..." | 0.0 | 0.0 | neutral | 0.000 | 1.000 | 0.000 | 0.0000 | 59 | 7 | nak mak teringin nak cucuk vaksin vak... | [, nak, mak, teringin, nak, cucuk, vaksin, vak, ...] | [, nak, mak, teringin, nak, cucuk, vaksin, vak, ...] | [, nak, mak, teringin, nak, cucuk, vaksin, vak, ...] |
| 3 | Come on vaccinated vaccinatedVaksinasiCovi... | 0.0 | 0.0 | neutral | 0.000 | 1.000 | 0.000 | 0.0000 | 58 | 6 | come on vaccinated vaccinated vaksinasicovid... | [come, on, vaccinated, vaccinated, vaksinasicovid, ...] | [come, on, vaccinated, vaccinated, vaksinasicovid, ...] | [come, vaccin, vaccinated, vaksinasicovid, vaksin,] |
| 4 | Top News Koran Rakyat MerdekaWaring Politis... | 0.5 | 0.5 | negative | 0.132 | 0.769 | 0.099 | -0.1531 | 115 | 16 | top news koran rakyat merdeka warning politis... | [top, news, koran, rakyat, merdeka, warning, p, ...] | [top, news, koran, rakyat, merdeka, warning, p, ...] | [top, news, koran, rakyat, merdeka, warn, poli, ...] |

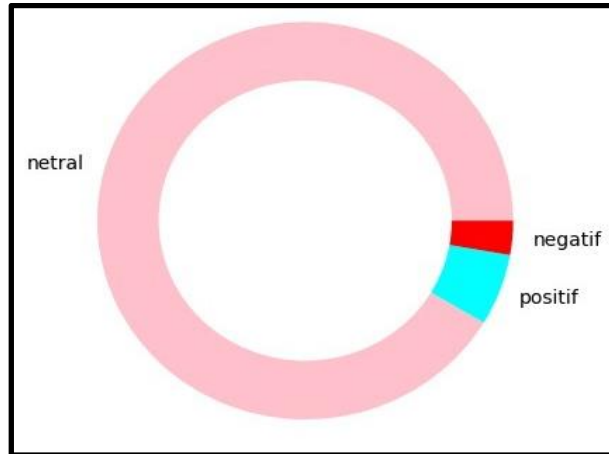
Gambar 3.1 Praproses Tweet

4. HASIL DAN PEMBAHASAN

Dari hasil pengujian kami dapat melaporkan bahwa sebanyak 1190 *tweets* yang direview dengan total 3196 kata. Dari total tersebut setidaknya ada 10 kata yang paling sering muncul dalam tweet yaitu: “vaksin” sebanyak 492 kata, “vaksinasicovid” sebanyak 472 kata, “vaksinasi” sebanyak 414, “covid” sebanyak 299 kata, “di” sebanyak 274 kata, kata “yang” sebanyak 202 kata, kata “dan” sebanyak 185 kata, kata “untuk” sebanyak 132 kata, kata “xm” sebanyak 117 kata, dan kata “dengan” sebanyak 114 kata. Kami juga menghitung beberapa kata yang muncul secara bersamaan. Untuk dua kata yang muncul secara bersamaan adalah kata: “covid” dan “19” sebanyak 215 kali kemudian “vaksinasi” dan “covid” sebanyak 86 kali, kata “juli dan “2021” sebanyak 61, kata “gerai” dan “vaksin” sebanyak 57 kali, yang terakhir adalah kata “vaksin” dan “presisi” sebanyak 56 kali. Sedangkan empat kata yang secara bersamaan muncul dalam twit adalah: “vaksinasi-covid-19” sebanyak 81 kali, “gerai-vaksin-presisi” sebanyak 56 kali, “partisi-bilik-vaksinasi” sebanyak 39 kali, “bilik-vaksinasi-dengan” sebanyak 39 kali, dan “vaksinasi-dengan-berbagai” sebanyak 39 kali. Untuk lebih jelas dapat dilihat pada tabel 4.1.

Tabel 4.1 Distribusi Kata yang Sering Digunakan

| Kata | Jumlah | Kombinasi 2 kata | Jumlah | Kombinasi 3 kata | Jumlah |
|--------|--------|------------------|--------|--------------------|--------|
| vaksin | 492 | covid-19 | 215 | vaksinasi-covid-19 | 81 |



Gambar 4.5 *Pie Chart Persentase Tweet Berdasarkan Jenis Sentimen*

Nilai sentimen pada setiap *tweets* di dapat pada perhitungan jumlah kata netral, positif, dan negatif.

5. KESIMPULAN

Dari hasil penelitian, kami dapat simpulkan bahwa masyarakat lebih banyak memberikan respon yang bersentimen netral yaitu 91.43% dibanding dengan positif 5.80% dan negatif 2.77%. Hal tersebut berarti masyarakat lebih banyak memberikan respon yang netral tentang vaksinasi COVID-19. Sentimen netral berarti *tweets* masyarakat yang di dapat tidak hanya terdiri dari pendapat yang menyatakan pro atau kontra tentang vaksinasi tetapi juga terdapat respon lain seperti harapan, pendapat umum, serta pengetahuan.

DAFTAR PUSTAKA

- [1] WHO, "Coronavirus disease (COVID-19)," WHO, 12 Oktober 2020. [Online]. Available: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19>.
- [2] Rokom, "Vaksin COVID-19 Merek Sinovac, AstraZeneca, Pfizer, dan Novavax Tidak Dapat Dipergunakan untuk Vaksinasi Gotong Royong," Kementerian Kesehatan RI, 15 Juni 2021. [Online]. Available: <https://sehatnegeriku.kemkes.go.id/baca/berita-utama/20210615/1837903/vaksin-covid-19-merek-sinovac-astrazeneca-pfizer-dan-novavax-tidak-dapat-dipergunakan-untuk-vaksinasi-gotong-royong/>.
- [3] R. N. Landers, R. Brusso, K. Cavanaugh and A. B. Collmus, "A Primer on Theory-Driven Web Scraping: Automatic Extraction of Big Data from the Internet for Use in Psychological Research," vol. 21, pp. 475-492, 2016.
- [4] S. A. Gode, A. V. Saurkar and K. G. Pathare, "An Overview on Web Scraping Techniques and Tools," International Journal on Future Revolution in Computer Science & Communication Engineering, vol. 4, no. 4, pp. 2454-4248, 2018.
- [5] R. Cooley, B. Mobasher and J. Srivastava, "Web mining: Information and Pattern Discovery on the World Wide Web," In Proceedings of the Ninth IEEE International Conference on Tools with Artificial Intelligence, pp. 558-567, Desember 1997.
- [6] P. Arsi and R. Waluyo, "Analisis Sentimen Wacana Pemindahan Ibu Kota Indonesia Menggunakan Algoritma Support Vector Machine (SVM)," Jurnal Teknologi Informasi dan Ilmu Komputer, vol. 8, no. 1, pp. 147-156, 2021.
- [7] E. Haddi, X. Liu and Y. Shi, "The Role of Text Pre-processing in Sentiment Analysis," Procedia Computer Science, vol. 17, pp. 26-32, 2013.
- [8] B. Ohana and B. Tierney, "Sentiment Classification of Reviews Using SentiWordNet," 9th. IT & T Conference, p. 3–10, 2009.
- [9] S. Mujilawati, "Pre-processing Text Mining pada Data Twitter," Seminar Nasional Teknologi Informasi dan Komunika, pp. 49-56, 2016.
- [10] F. F. Rachman and S. Pramana, "Analisis Sentimen Pro dan Kontra Masyarakat Indonesia tentang Vaksin COVID-19 pada Media Sosial Twitter," Indonesian of Health Information Management Journal, vol. 8, no. 2, pp. 100 - 109, 2020.