

Analisis Sentimen untuk Melihat Respon Masyarakat Terhadap Vaksin Pfizer

Dita Nurmadewi¹, Christophora Putri Gusti², Ergi Cahya Hernanto³, Farrell Ananda⁴,
Mery Andani Hutagalung⁵, Nadyatul Hikmah⁶

^{1,2,3,4,5,6}Program Studi Sistem Informasi, Universitas Bakrie, Jakarta Selatan, Indonesia

e-mail: dita.nurmadewi@bakrie.ac.id¹, 1202002034@student.bakrie.ac.id²,

1202002028@student.bakrie.ac.id³, 1202002007@student.bakrie.ac.id⁴,

12020029@student.bakrie.ac.id⁵, 1202002022@student.bakrie.ac.id⁶

* Penulis Korespondensi: E-mail: dita.nurmadewi@bakrie.ac.id

Abstrak: *Covid-19* merupakan penyakit yang penyebarannya sangat luas dan cepat. Dalam waktu singkat, penyakit ini menyebar ke seluruh dunia, termasuk Indonesia. Tentunya peristiwa ini menyerang masyarakat dari berbagai faktor, khususnya di bidang kesehatan. Dalam rangka melakukan percepatan penanganan penyebaran *Covid-19*, pemerintah Indonesia mewajibkan warga negaranya melakukan vaksinasi *Covid-19* salah satunya adalah vaksin *Pfizer*. Vaksin *Pfizer* adalah salah satu vaksin terbaik diantara vaksin lainnya dan kemungkinan munculnya efek samping sangat kecil. Namun, penggunaan vaksin ini tetap menimbulkan reaksi dan opini yang berbeda dari masyarakat umum. Oleh karena itu, penelitian ini bertujuan untuk mendapatkan sentimen masyarakat terhadap vaksin *Pfizer* melalui data di *Twitter* yang diambil melalui *netlytic*. Analisis sentimen merupakan kegiatan untuk mencari opini masyarakat tentang sebuah objek yang ingin diketahui. Untuk mendukung penelitian mengenai hal tersebut, peneliti menggunakan algoritma *naive bayes* untuk mengidentifikasi sentimen masyarakat terhadap vaksin *Pfizer*. Algoritma ini dipilih karena efisiensi dan kemudahan dalam penggunaannya serta keakuratannya dalam proses klasifikasi data. Hasil dari penelitian ini adalah analisis sentimen menggunakan algoritma *naive bayes* berhasil mengklasifikasikan respon masyarakat terhadap vaksin *Pfizer*. Hasil klasifikasi menunjukkan sentimen positif lebih mendominasi dari sentimen negatif dan netral. Hasil klasifikasi menunjukkan sentimen positif lebih mendominasi dari sentimen negatif dan netral dengan jumlah sentimen positif sebesar 67% dari data yang berhasil di analisis.

Kata Kunci: Analisis Sentimen; Naive Bayes; Pfizer; Vaksin

Abstract: *Covid-19* is a disease with rapidly widespread transmission. In a short period of time, the disease has spread to all parts of the world, including Indonesia. Of course, this event affects people from various factors, especially in health. In order to speed up handling the spread of *Covid-19*, the Indonesian government requires its citizens to get vaccinated, one of which is the *Pfizer* vaccine. The *Pfizer* vaccine is one of the best among other vaccines and the possibility of side effects is very small. However, the use of this vaccine still triggers different reactions and opinions from the general public. Therefore, this study aims to obtain the sentiment of the public towards the *Pfizer* vaccine through data on *Twitter* collected through *Netlytic*. Sentiment analysis is an activity to find out public opinions about an object that needs to be known. To support research in this matter, the researcher used a *naive bayes* algorithm to identify public sentiment towards the *Pfizer* vaccine. This algorithm was chosen because of its efficiency and ease of use and its accuracy in data classification processes. The result of this study is that sentiment analysis using the *naive bayes* algorithm successfully classifies the response of the public towards the *Pfizer* vaccine. The classification results show that positive sentiment dominates over negative and neutral sentiment. The classification results show that positive sentiment dominates over negative and neutral sentiment with a positive sentiment of 67% from the data that was successfully analyzed.

Keywords: *Naive Bayes; Pfizer; Sentiment Analysis; Vaccines*

PENDAHULUAN

Saat ini di seluruh negara sedang digemparkan karena adanya wabah virus *corona (covid-19)*. Virus ini pertama kali ditemukan di Wuhan, China pada bulan November hingga Desember tahun 2019. *Covid-19* merupakan salah satu jenis virus yang berbahaya karena prevelensinya yang tinggi dan tingkat penyebarannya yang cepat dari satu orang ke banyak orang. Tercatat menurut data dari *World Health Organization (WHO)* sudah ada lebih dari 497 juta kasus *covid-19* tersebar di seluruh dunia yang statusnya terkonfirmasi dan memiliki angka kematian di atas 6 juta jiwa [1]. Indonesia juga

menjadi salah satu negara yang beberapa masyarakatnya terjangkit wabah *covid-19*. Kasus *covid-19* di Indonesia pertama kali ditemukan pada 2 Maret 2020 di Jawa Barat, dimana dua warga Depok terkonfirmasi positif terjangkit wabah *covid-19* [2]. Sejak terkonfirmasi pertama kali virus ini masuk di Indonesia, semakin banyak masyarakat yang terinfeksi dan harus dirawat di rumah sakit.

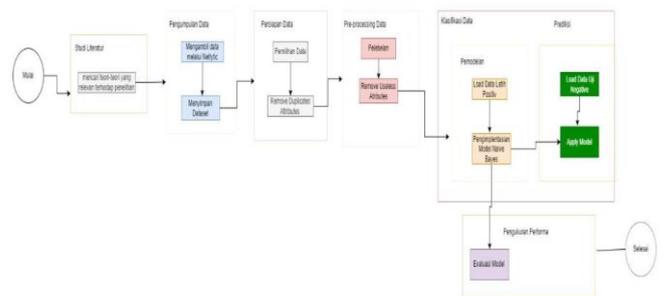
Pemerintah Indonesia berupaya sebaik mungkin untuk menanggulangi penyebaran *covid-19* dengan mengeluarkan kebijakan Pembatasan Sosial Berskala Besar (PSBB). Kebijakan PSBB meliputi pembatasan

pada kegiatan pendidikan yang awalnya kegiatan belajar mengajar dilakukan di kelas atau secara *offline* dirubah menjadi *online*, kegiatan kerja juga dilakukan dengan *work from home*, ibadah dianjurkan untuk dilaksanakan di rumah, kemudian di tiadakannya agenda kegiatan berkumpul, dan masih banyak pembatasan lainnya. Kebijakan PSBB ini diatur dalam Peraturan Pemerintah No. 21 yang dikeluarkan pada tahun 2020 tentang Pembatasan Sosial Berskala Besar dalam rangka percepatan penanganan *covid-19* [3]. Upaya pemerintah Indonesia tidak hanya terkait kebijakan PSBB, tetapi juga melalui kegiatan vaksinasi terhadap masyarakat. Bersumber dari sejarah bagaimana vaksinasi telah menyelamatkan masyarakat dari beberapa penyakit dan kecacatan dengan bantuan vaksinasi. Kegiatan vaksinasi sebagai alternatif yang dianggap efektif dalam mencegah penularan penyakit berbahaya [4]. Kegiatan vaksinasi juga bertujuan untuk menekan tingkat penyebaran wabah *covid-19* dan menurunkan angka positif kematian di Indonesia, sehingga masyarakat yang sudah melakukan vaksinasi diharapkan tubuhnya akan kebal terhadap virus *covid-19*. Ada beberapa jenis vaksin yang di tawarkan yaitu *Sinovac, Pfizer, Moderna, Astra Zeneca, Biofarma* dan *Sinopharm* [5]. *Pfizer* menjadi salah satu vaksin yang dapat dipilih oleh masyarakat dimana menurut penelitian terdahulu vaksin ini memiliki tingkat efektivitas dan keamanan yang bagus dengan nilai 94,6%. *Pfizer* juga tidak menimbulkan efek samping yang serius sehingga aman digunakan oleh masyarakat [6].

Informasi mengenai vaksinasi ini juga sudah tersebar ke beberapa sosial media, salah satunya di *twitter*. Media sosial *Twitter* menjadi alternatif untuk berkomunikasi dan bertukar informasi mengenai vaksin *Pfizer*. Tidak dipungkiri adanya pro dan kontra dari masyarakat yang membicarakan vaksin *Pfizer* di *Twitter* serta adanya potensi mis-informasi yang dapat terjadi juga. Oleh karena itu, penting untuk melihat bagaimana pendapat masyarakat mengenai vaksin *Pfizer* sebagai salah satu upaya untuk mencegah penyebaran wabah *covid-19* menggunakan algoritma *naive bayes* dengan *tools rapidminer*. Algoritma *naive bayes* dipilih karena keakuratannya dalam memberikan nilai probabilitas yang tepat. Hasil penelitian ini nantinya dapat menjadi informasi bagi pemerintah untuk melihat bagaimana reaksi dari masyarakat dan dampak yang ditimbulkan terkait adanya vaksin *Pfizer*.

METODE PENELITIAN

Metode yang digunakan pada penelitian ini digambarkan dan dijelaskan sebagai berikut :



Gambar 1. Alur Penelitian

Studi Literatur

Studi literatur digunakan untuk mencari teori-teori yang relevan terhadap penelitian sejenis yang pernah dilakukan sebelumnya terkait hasil penemuannya. Ini menjadi rujukan guna mendapat solusi dari permasalahan yang peneliti dihadapi. Studi literatur berasal dari beberapa jurnal, buku dan *website* terkait analisis sentimen terhadap vaksin *Pfizer*.

Penentuan dan Pengumpulan Sumber Data

Sumber data yang digunakan pada penelitian ini berasal dari data *Twitter*. Sumber data yang digunakan pada penelitian ini berasal dari data *Twitter* yang diambil melalui *website Netlytic* (netlytic.org). Data berupa *tweet* yang di *publish* di *Twitter*. Data yang digunakan adalah data pada tahun 2022 dengan jumlah *user_id* 460 pengguna.

Persiapan Data

Data di ambil dengan metode *crawling* atau pengambilan data untuk mengindeks informasi di halaman menggunakan URL (*Uniform Resource Locator*). *Tools* yang digunakan adalah aplikasi *Rapidminer*. *Rapidminer* merupakan perangkat lunak mandiri yang digunakan untuk analisis data dan mesin penambangan data [7]. *Rapidminer* dapat membantu dalam melakukan analisa secara otomatis [8]. Data hasil *crawling* dibersihkan dan dilakukan proses *filter* berdasarkan kebutuhan penelitian ini. Selanjutnya, data yang sudah sesuai dengan kebutuhan penelitian ini menjadi *dataset* yang nantinya akan diolah.

Pre-Processing Data

Pada tahap ini dilakukan pelabelan data untuk menentukan kelas sentimen *positive* dan *negative* dari data yang telah di *filter* atau di bersihkan. Pada tahap ini juga atribut yang tidak digunakan dihilangkan. Seperti atribut tanggal, atribut nama, dsb.

Penggunaan Algoritma Naive Bayes

Teknik analisis yang digunakan untuk mengidentifikasi sentimen masyarakat terkait vaksin *Pfizer* yaitu menggunakan algoritma *naive bayes*. *Naive bayes* merupakan salah satu metode *machine learning* yang menggunakan perhitungan dengan probabilitas [9]. Nilai probabilitas didasarkan pada setiap atribut pada masing-masing kelas [10]. Algoritma ini mudah digunakan untuk perhitungan yang sederhana dan tidak panjang [11]. Cara kerja dari algoritma ini di dasarkan

pada teorema bayes, dimana bentuk umum dari perhitungannya sebagai berikut :

$$P(Y|X) = \frac{P(X|Y) * P(Y)}{P(X)} \dots\dots\dots(1)$$

Keterangan :

X = Kelas data yang belum diketahui

Y = Hipotesa data X adalah kelas spesifik

P(Y|X) = Kemungkinan Hipotesa H berdasarkan keadaan X

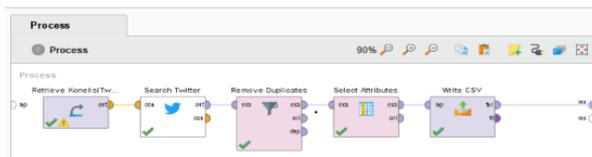
P(X|Y) = Kemungkinan X berdasarkan keadaan tersebut

P(X) = Kemungkinan dari X

Rumus ini digunakan untuk menghitung peluang suatu data masukan X yang termasuk ke dalam kelas target Y [12].

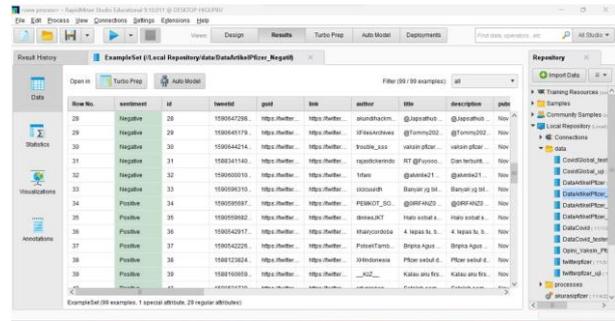
HASIL DAN PEMBAHASAN

Hasil dari penelitian dijelaskan sesuai dengan metode penelitian yang telah dijelaskan sebelumnya. Pada tahapan pengumpulan data atau pembuatan dataset dengan metode *crawling* atau pengambilan data menggunakan beberapa *query* atau kata kunci pencarian yaitu “vaksin Pfizer” menggunakan bahasa Indonesia menggunakan *netlytic*. Data yang berhasil di *crawling* sebanyak 460 data dan dilanjutkan analisis data menggunakan *software Rapidminer*. Selanjutnya data dibersihkan dari duplikasi data dan dilakukan pemilihan atribut yang sesuai menggunakan operator yang ada di *rapidminer*. Operator yang digunakan yaitu *retrieve file* untuk *import file* dengan format *.raw* data yang sudah berhasil didapatkan, *remove duplicates* untuk menghapus baris yang berulang, *select attributes* yang digunakan untuk memilih satu atau sekumpulan data yang memiliki informasi tertentu, *write csv* untuk *import* data dengan format *csv* dan *subprocess* untuk menguji suatu *text* memiliki nilai *true* atau *false*. Data yang berhasil dibersihkan dari adanya duplikasi dan data kosong, serta yang memenuhi kriteria, ditemukan sebanyak 99 data. Penelitian ini mengambil atribut teks untuk diolah. Prosesnya digambarkan seperti dibawah ini:



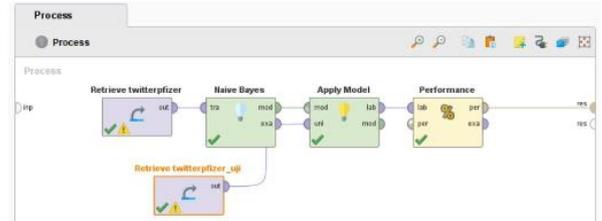
Gambar 2. Proses Pengambilan Data

Data selanjutnya disimpan ke dalam format teks atau *.csv*. Kemudian dilanjutkan ke proses berikutnya yaitu pemberian label. Pelabelan dilakukan untuk membagi data menjadi dua kelas sentimen yang bernilai *positif* dan *negatif* yang di dasarkan pada reaksi atau informasi yang di publish oleh pengguna *Twitter*. Proses digambarkan seperti di bawah ini :



Gambar 3. Pelabelan Data

Proses berikutnya yaitu pemrosesan data yang dilakukan menggunakan algoritma *naive bayes*. Algoritma ini dipilih karena efisiensi dan kemudahan dalam penggunaannya serta keakuratannya dalam proses klasifikasi data. Algoritma digunakan untuk klasifikasi data. Dari sini akan diketahui nilai akurasi, presisi dan *recall*. Akurasi untuk melihat rasio dari prediksi benar yang memiliki nilai *positive* dan *negative* dari keseluruhan data, presisi untuk melihat rasio prediksi benar *positive* yang dibandingkan dengan keseluruhan hasil yang diprediksi bernilai *positive* dan *recall* untuk melihat sensitifitas data dengan melihat rasio prediksi benar positif yang dibandingkan dengan keseluruhan data yang benar positif. Prosesnya digambarkan seperti dibawah ini :



Gambar 4. Pengimplementasian Algoritma Naive Bayes

Tahap ini juga menggunakan beberapa operator pendukung, seperti operator *apply model* dan *performance test*. Operator *apply model* untuk mendapatkan prediksi data yang digunakan dan operator *performance test* digunakan untuk mengevaluasi hasil akhir. Hasil akhir dari penelitian ini yaitu :

Performance	accuracy: 98.99%	True Positive	True Negative	Class Precision
AUC (symmetric)		66	0	100.00%
AUC (asymmetric)		66	1	97.96%
Class Recall	98.48%	100.00%		

Gambar 5. Hasil Akhir

Hasil akhir yang didapatkan dengan menguji *dataset* menggunakan algoritma *naive bayes* yaitu nilai akurasi sebesar 98%, dengan *precision* untuk *positive class* bernilai benar sebesar 100% dan nilai *recall* yang diperoleh untuk *positive class* bernilai benar sebesar 98.48%. Dari hasil di atas juga dapat dilihat bahwa total sentimen positif sebanyak 66 *tweet* atau sekitar 67%,

dengan total sentimen negatif sebanyak 33 atau sekitar 33%.

KESIMPULAN DAN SARAN

Berdasarkan analisis dan pengujian yang dilakukan, maka kesimpulan yang dapat diambil adalah hasil dari implementasi algoritma *naive bayes*, pada penelitian ini data menunjukkan bahwa tingkat keakurasian untuk sentimen masyarakat terhadap vaksin *Pfizer* yaitu 98%. Hasil akurasi pada penelitian dinilai sangat baik. Analisis sentimen dengan algoritma *naive bayes* berhasil mengklasifikasikan sentimen masyarakat terhadap vaksin *Pfizer*. Hasil klasifikasi menunjukkan sentimen positif lebih mendominasi dari sentimen negatif.

Dapat disimpulkan bahwa banyak masyarakat yang mendukung adanya vaksin *Pfizer*. Selain itu dari data juga didapatkan info bahwa pengguna vaksin *Pfizer* merasakan efek samping yang kecil atau bahkan tidak merasakan efek samping apapun dibandingkan dengan vaksin lainnya. Namun, disisi lain beberapa orang berpendapat negatif terhadap vaksin *pfizer* dikarenakan mereka merasakan efek sampingnya seperti tangannya menjadi sakit, mudah tertular batuk atau pilek, atau mudah lelah, dan lain-lain.

Saran untuk peneliti berikutnya yaitu dapat dilakukan penelitian lanjutan mengenai analisis sentimen terhadap jenis vaksin lain dengan menggunakan algoritma klasifikasi selain *naive bayes*.

DAFTAR PUSTAKA

- [1] «WHO Covid-19,» 7 November 2022. [En línea]. Available: <https://covid19.who.int>.
- [2] A. Susilo, C. M. Rumende, C. W. Pitoyo y W. D. Santoso, «Coronavirus Disease 2019: Tinjauan Literatur Terkini,» *Jurnal Penyakit Dalam Indonesia*, vol. 7, pp. 45-67, 2020.
- [3] «JDIH BPK RI,» 31 Maret 2020. [En línea]. Available: <https://peraturan.bpk.go.id/Home/Details/135059/pp-no-21-tahun-2020>. [Último acceso: 7 November 2022].
- [4] L. A. Octafia, «Vaksin Covid-19 Perdebatan, Persepsi dan Pilihan,» *Jurnal EMIK*, vol. 4, n° 2, pp. 160-174, 2021.
- [5] Satuan Tugas Penanganan Covid-19, Pengendalian Covid-19 dengan 3M, 3T, Vaksinasi, Disiplin, Kompak dan Konsisten, Jakarta, 2021.
- [6] S. A. Nugroho y I. N. Hidayat, «Efektivitas Dan Keamanan Vaksin Covid-19 : Studi Refrensi,» *Jurnal Keperawatan Profesional (JKP)*, vol. 9, n° 2, 2021.
- [7] V. R. Prasetyo, H. Lazuardi, A. A. Mulyono y C. Lauw, «Penerapan Aplikasi RapidMiner Untuk Prediksi Nilai Tukar Rupiah Terhadap US Dollar Dengan Metode Regresi Linier,» *Jurnal Nasional Teknologi dan Sistem Informasi*, vol. 7, 2021.
- [8] D. A. C., D. A. Baskoro, L. Ambarwati y I. W. S. Wicaksana, Belajar Data Mining dengan RapidMiner, Jakarta, 2013.
- [9] H. Susana, N. Suarna, Fathurrohman y Kaslani, «Penerapan Model Klasifikasi Metode Naive Bayes Terhadap Penggunaan Akses Internet,» *Jurnal Sistem Informasi dan Teknologi Informasi*, vol. 4, n° 1, pp. 1-8, 2022.
- [10] R. Y. Hayuningtyas, «Penerapan Algoritma Naive Bayes untuk Rekomendasi Pakaian Wanita,» *Jurnal Informatika*, vol. 6, n° 1, pp. 18-22, 2019.
- [11] H. F. Putro, R. T. Vulandari y W. L. Y. Saptomo, «Penerapan Metode Naive Bayes Untuk Klasifikasi Pelanggan,» *Jurnal Teknologi Informasi dan Komunikasi Sinar Nusantara*, vol. 8, n° 2, 2020.
- [12] R. Rachman y R. N. Handayani, «Klasifikasi Algoritma Naive Bayes Dalam Memprediksi Tingkat Kelancaran Pembayaran Sewa Teras UMKM,» *Jurnal Informatika*, vol. 8, n° 2, pp. 111-122, 2021.