Analisis Pro Kontra Vaksin Covid 19 Menggunakan Sentiment Analysis Sumber Media Sosial Twitter

I Wayan Desta Gafatia & Novri Hadinata

Fakultas Teknik Ilmu Komputer, Universitas Bina Dharma

Email: Wayandesta1235@gmail.com

Abstract

The development of information technology today has experienced very rapid growth. One of the developments in information technology, namely social media such as Twitter, Facebook, and Youtube, are some of the most popular communication media in today's society. Twitter is often used to express emotions about something, either praising or criticizing in the form of emotion. Human emotions can be categorized into five basic emotions, namely love, joy, sadness, anger, and fear. Twitter users' emotional tweets can be known as opinion or sentiment analysis (opinion analysis or sentiment analysis). Sentiment analysis is also carried out to see opinions or tendencies towards a problem or policy, whether they tend to have negative or positive opinions. The COVID-19 vaccine has become one of the discussions with a fairly high intensity on social media. Vaccine-related tweets have increased as government policies evolve. The responses of netizens also varied, ranging from clinical trials of vaccines, free vaccines, vaccine effectiveness, halal vaccines, to the implementation of vaccinations. This research produces a system that can analyze tweet sentiment related to the covid 19 vaccine in Indonesia where the tweet is obtained using the Twitter API. This system uses the Multinomial Naive Bayes method for the classification process.

Keywords: Covid 19 vaccine, Sentiment analysis, Twitter.

1. Pendahuluan

Perkembangan teknologi informasi saat ini telah mengalami pertumbuhan yang sangat pesat [Yustiani and Yunanto, 2017], Salah satu perkembangan teknologi informasi yakni Media sosial seperti Twitter, Facebook, dan Youtube merupakan beberapa media perangkat komunikasi terpopuler di masyarakat saat ini. Twitter seringkali digunakan untuk mengungkapkan emosi mengenai sesuatu hal, baik memuji ataupun mencela dalam bentuk emosi. Emosi manusia dapat dikategorikan menjadi lima emosi dasar yaitu cinta, senang, sedih, marah, dan takut. Emosi tweet para pengguna twitter dapat dikenali dengan analisis opini atau sentimen (opinion analysis atau sentiment analysis [Rizal, 2017]).

Analisis sentimen juga dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah maupun kebijakan, apakah cenderung beropini negatif atau positif terhadap suatu tokoh tertentu [Rizal, 2017]. Oleh karena itu dibutuhkan sebuah sistem yang dapat menganalisis sentimen, terutama tweet yang berbahasa Indonesia. Vaksin covid 19 menjadi salah satu pembahasan dengan intensi cukup tinggi di media sosial. Cuitan terkait vaksin meningkat seiring perkembangan kebijakan dari pemerintah. Respons warganet pun beragam, mulai dari soal uji klinis vaksin, vaksin gratis, efektivitas vaksin, kehalalan vaksin, hingga pelaksanaan vaksinasi. Pro & kontra ini banyak di bicarakan oleh masyarakat Indonesia melalui media sosial Twitter.

Penelitian ini menghasilkan sebuah sistem yang dapat menganalisis sentimen tweet yang berhubungan dengan vaksin covid 19 di Indonesia dimana tweet tersebut didapat menggunakan Twitter API. Sistem ini menggunakan metode Multinomial Naive Bayes untuk proses klasifikasinya. Oleh karena itu, dilakukan peneliti dengan judul

Analisis Pro Kontra Vaksin Covid 19 Menggunakan Sentiment Analysis Sumber Media Sosial Twitter.

2. Tinjauan Literatur

Analisis sentimen atau analisis sentimen dalam bahasa Indonesia adalah teknik atau cara Anda menentukan bagaimana perasaan dengan teks dinyatakan dan bagaimana perasaan ini dapat dikategorikan sebagai katup sentimen positif dan negatif. **[Error! Reference source not found.]** Analisis sentimen atau rexifompinasi mengacu pada area luas pemrosesan bahasa alami, komputer linguistik dan penurunan teks bahwa tujuan menganalisis pendapat, suasana hati, evaluasi, sikap, peringkat dan emosi seseorang atau penulis Salah satu tema puas atau dengan topik yang puas adalah penulis, produk, layanan, organisasi, individu atau kegiatan tertentu. Tugas analisis sentimen adalah untuk mengelompokkan teks dalam kalimat atau dokumen, dan kemudian pendapat yang dianalisis dalam kalimat atau dokumen menentukan apakah mereka positif, negatif atau netral **[Error! Reference source not found.]**.

Analisis sentimen atau komentar mengacu pada berbagai macam pemrosesan bahasa alami, komputer linguistik dan ekstraksi teks, pendapat, sentimen, ulasan, saksi, ulasan, dan emosi seseorang atau juru bicara atau penulis dalam kaitannya dengan produk topik, suatu produk, sebuah produk, produk, layanan, wastafel, organisasi, individu atau kegiatan tertentu (Liu, 2011).

Tugas dasar dalam analisis sentimen adalah untuk mengelompokkan dalam satu kalimat atau dokumen dalam satu kalimat atau dokumen, kemudian menentukan saran yang telah disampaikan ke operasi atau dokumen, baik positif, negatif atau netral (dehaff, m ., 2010) adalah. Analisis sentimen juga dapat mengekspresikan perasaan emosional emosional, bahagia atau marah.

Kami dapat mencari produk, merek atau orang setelah berkonsultasi dan menentukan apakah mereka positif atau negatif di Internet (Saraswati, 2011). Ini memungkinkan kita untuk menemukan informasi tentang:

- a. Pengenalan Api (Rants Bad)
- b. Persepsi produk baru.
- c. Persepsi merek.
- d. Manajemen reputasi.

Ekspresi atau sentimen mengacu pada fokus topik tertentu, instruksi pada suatu topik dapat menjadi makna yang berbeda dengan pernyataan yang sama pada berbagai topik. Dalam beberapa penelitian, terutama dalam penilaian produk, elemen suatu produk akan dipromosikan, yang dibahas sebelum proses pembukaan ditentukan (Barber, 2010).

Analisis sentimental atau penambangan pendapat adalah deteksi keberadaan benda atau orang [beineke et al. 2014]. Analisis sentimen dapat digunakan untuk mendapatkan proporsi positif perasaan dan perasaan negatif seseorang, perusahaan, institusi, produk atau dalam kondisi tertentu. Nilai analisis sentimen dapat dibagi menjadi 3, yaitu perasaan positif, perasaan negatif dan perasaan netral atau diperbarui, sehingga ia dapat menemukan siapa atau kelompok itu adalah sumber perasaan positif atau perasaan negatif.

Analisis sentimental harus emosi, sikap, opini, ulasan yang disampaikan oleh seseorang pembicara atau penulis produk atau tokoh masyarakat. Alasan-alasan ini menyebabkan studi yang berbeda, terutama dalam ulasan produk yang mendahului elemen produk yang dibahas sebelum membahas proses tindakan.

Penelitian yang dilakukan oleh Saputra dan kawan – kawan (2015) yang berjudul “ANALISIS SENTIMEN DATA PRESIDEN JOKOWI DENGAN PREPROCESSING NORMALISASI DAN STEMMING MENGGUNAKAN METODE NAIVE BAYES DAN SVM ” penelitian ini Boolean searching dengan operator “AND” sangat efektif dan efisien dalam pencarian data, selain itu cukup berpengaruh terhadap akurasi analisis sentimen, karena menghasilkan tidak semua data mengenai Jokowi, melainkan yang mengandung kata yang dimasukkan. Berdasarkan penelitian yang dilakukan, akurasi yang dihasilkan ketika data dilakukan stemming, terdapat peningkatan rata-rata sebesar 0,85% untuk metode Naive Bayes dan 0,85% untuk metode SVM.[Saputra et al., 2015]

Penelitian yang dilakukan oleh Sudiantoro (2018) dengan judul penelitian “ANALISIS SENTIMEN TWITTER MENGGUNAKAN TEXT MINING DENGAN ALGORITMA NAÏVE BAYES CLASSIFIER” Hasil penelitian ini menunjukkan bahwa: Algoritma Naïve Bayes Classifier sangat efektif untuk digunakan sebagai proses klasifikasi tweet yang dibutuhkan dalam sistem analisis sentiment ini dimana nilai yang didapatkan dalam pengujian sampai 84%. Metode Naïve Bayes Classifier dapat digunakan untuk melakukan klasifikasi tweets dengan cukup baik pada sistem analisis sentiment [Sudiantoro and Zuliarso, 2018]. Penelitian yang dilakukan oleh AS Eddoyono (2014) dengan judul penelitian “Kawan, sekaligus lawan bagi redaksi berita” Hasil penelitian ini menunjukkan bahwa Twitter masih berpeluang menjadi sumber informasi baru sekaligus media sosialisasi dan interaksi. Twitter juga berpeluang menjadi pesaing bagi media massa mainstream. Media massa memiliki beragam strategi untuk memanfaatkan secara optimal keberadaan Twitter, sekaligus mengimbangnya. Salah satu upaya yang dilakukan adalah terus berinovasi memanfaatkan fitur-fitur yang ditawarkan Twitter [**Error! Reference source not found.**].

Penelitian yang dilakukan oleh AS Eddoyono (2014) dengan judul penelitian “Kawan, sekaligus lawan bagi redaksi berita” Hasil penelitian ini menunjukkan bahwa Twitter masih berpeluang menjadi sumber informasi baru sekaligus media sosialisasi dan interaksi. Twitter juga berpeluang menjadi pesaing bagi media massa mainstream. Media massa memiliki beragam strategi untuk memanfaatkan secara optimal keberadaan Twitter, sekaligus mengimbangnya. Salah satu upaya yang dilakukan adalah terus berinovasi memanfaatkan fitur-fitur yang ditawarkan Twitter [**Error! Reference source not found.**].

Penelitian yang dilakukan oleh Nurul Gayatri Indah Reza (2018) dengan judul “Penerapan Algoritma DbSCAN Untuk Pencarian Trend Topik Pilkada Pekanbaru 2017 Pada Twitter” penelitian ini melakukan beberapa percobaan dengan beberapa parameter Eps dan MinPts yang berbeda terhadap 2.184 data teks yang sudah melalui tahapan *cleaning*, penghapusan duplikasi, serta pra-proses seperti *stemming* dan *stopwords* [**Error! Reference source not found.**].

3. Metode Penelitian

Metode penelitian ini menggunakan Teknik Sentiment Analysis Sumber Media Sosial Twitter. Skema penelitian yang dilakukan penulis yang pertama adalah mendaftarkan akun ke twitter developer yang beralamat di situs <https://developer.twitter.com/> yang telah disediakan oleh pihak Twitter, setelah mendaftarkan twitter developer kemudian penulis bisa mendapatkan akses keys inilah yang akan digunakan untuk aplikasi meneliti tweets dari sebuah akun twitter. Langkah selanjutnya adalah memnentukan klasifikasi naive bayes sesuai yang telah ditetapkan oleh penulis berdasarkan topik berdasarkan kata kunci atau hastag vaksinisasi yang merupakan topik trending tahun 2021 yakni covid 19. Setelah menentukan kata kunci berupa hastag kemudian penulis mendapatkan hasil sentimen analysis yang berupa pro kontra meliputi positive, netral maupun negative dan kemudian didokumentasi untuk keperluan peneitian.

3.1 Tahapan Penelitian

Tahapan penelitian yang dilakukan oleh peneliti ialah berdasarkan metode deksriptif untuk mendapatkan *trends* yang terjadi di ruang lingkup masyarakat dalam bentuk kata, melalui *Aplication Programming Interface* (API) adalah suatu library yang berbentuk JavaScript[3] yang telah di sediakan oleh twitter metode tersebut merupakan metode penelitian yang berlandaskan yang berlandaskan paspositivisme yaitu memandang realita, fenomena yang diklarifikasikan tetap dan terukur.[4]

Gambar 3.1 Skema Penelitian

3.2 Naïve Bayes

Naïve Bayes merupakan salah satu metode statistik probabilitas berdasarkan penerapan Teorema Bayes dengan asumsi independen (naif) yang kuat, untuk memprediksi kelas dari suatu dokumen berdasarkan probabilitasnya

Bentuk umum atau persamaan dari teorema Bayes adalah :

$$P(H|X) = \frac{P(X|H) \cdot P(H)}{P(X)}$$

Keterangan : X : Data dengan class yang belum diketahui H : Hipotesa data X merupakan suatu cass spesifik P(H|X) : Probabilitas hipotesis H berdasar kondisi X (posteriori probability) P(H) : Probabilitas hipotesis H (prior probability) P(X|H) : Probabilitas X berdasarkan kondisi pada hipotesis H P(X) : Probabilitas X Untuk menjelaskan metode Naive Bayes, perlu diketahui bahwa proses klasifikasi memerlukan sejumlah petunjuk

untuk menentukan kelas apa yang cocok bagi sample yang dianalisis tersebut. Karena itu, metode Naïve Bayes di atas disesuaikan sebagai berikut: $(C|F1...Fn) = (C)P(F1...Fn|C)P(F1...Fn)$ Di mana Variabel C merepresentasikan kelas, sementara variabel “F1 ... Fn” merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (Posterior) adalah peluang munculnya kelas C (sebelum masuknya sampel tersebut, seringkali disebut prior)/

Naïve Bayes classifier berasumsi bahwa kehadiran fitur tertentu dari suatu kelas tidak berhubungan dengan kehadiran fitur lainnya (independen). Naïve bayes merupakan salah satu pendekatan yang sangat sederhana pada teks klasifikasi [6]

Polarity						
Spositive	Positive	WPositive	Netral	WNegative	Negative	SNegative
>0 <=0.3	>0.3 <=0.6	>0.6 <-1	0	>-0.3 <=0	>=0.6 <=0.3	>-1 <=0.6

Dari gambar diatas peneliti melakukan klasifikasi berdasarkan polarity labelisasi data yakni netral adalah jika polarirty > 0 dan polarity <= 0.3 maka dapat disimpulkan sentimennya adalah weekness positive, jika polarirty > 0.3 dan polarity <= 0.6 maka dapat disimpulkan sentimennya adalah positive, jika polarirty > 0.6 dan polarity <= 1 maka dapat disimpulkan sentimannya adalah strenght positive, jika polarirty > -0.3 dan polarity <= 0 maka dapat disimpulkan sentimenya adalah weakly negative, jika polarirty > -0.6 dan polarity <= 0.4 maka dapat disimpulkan sentimenya adalah Negative, jika polarirty > -1 dan polarity <= 0.6 maka dapat disimpulkan sentimenya adalah Strongly Negative.

3.3 Crawling Data

Crawling data pada Twitter merupakan tahap proses dalam penarikan data dari serverTwitter dengan memanfaatkan API (*Application Programming Interface*).

Gambar 3.2 Alur Crawling Data

Berikut adalah penjelasan dari gambar 3.1 yakni:

1. Penulis mendaftarkan akun di Twitter Developer dengan alamat.
2. Setelah mendaftarkan akun Twitter Developer maka akan mendapatkan akses keys dan token yang berfungsi sebagai proses autentikasi untuk mengakses steam Twitter yang telah disediakan oleh API search.
3. Penulis melakukan eksekusi untuk menarik data dengan memanfaatkan kode pmrograman Python yang telah di buat oleh penulis.
4. Menyimpan hasil penarikan data tweets kedalam bentuk Json.

3.4 Analyzing Data

Pada tahap *Analyzing data* penulis akan mengetahui informasi berdasarkan data yang telah disimpan dengan menggunakan bantuan *tools* Jupyter.

Gambar 3.3 Bentuk Data Preprocessing.

Dokumen yang di input akan dihilangkan proses menghilangkan duplikat data dan proses *Tokenization* adalah proses memisahkan data di dalam kalimat menjadi potongan kata tunggal atau termmed word. *Stopwords* adalah kata umum yang tidak memiliki makna dan biasanya muncul dalam jumlah yang besar. Mengingat tingginya frekuensi kemunculan *stopwords* dalam sebuah dokumen, membuat tahap menghilangkan *stopwords* sangat penting, karena akan membuat topik tidak dapat diinterpretasi dengan baik. Dalam proses *Stemming* digunakan untuk mengganti bentuk dari suatu kata menjadi kata dasar dari kata tersebut. Proses *stemming* akan menghilangkan semua imbuhan, yang terdiri dari awalan, sisipan, akhiran dan gabungan dari awalan dan akhiran pada kata turunan, Dikarenakan data teks perlu dibentuk menjadi kata dasar agar tidak terdapat kata yang sama namun berbeda karna adanya imbuhan.

Setelah mendapatkan consumer key dan consumer secret selanjutnya adalah mengimport

library yang diperlukan yakni sys, tweepy, csv, re, texblob, matplotlib

4. Hasil dan Pembahasan

4.1 Hasil Twitter Developer Akses token

Penulis menggunakan bahasa pemrograman python untuk menarik data pada Twitter dengan menggunakan API (*Application Programming Interface*) yang telah disediakan oleh Twitter Developer dengan memanfaatkan akses keys dan token, dengan terdaptarnya sebagai pengguna Twitter Developer kemudian melakukan pengajuan untuk meminta hak akses dengan mengisi beberapa ketentuan pembuatan yang berlaku oleh Twitter, setelah disetujui oleh pihak Twitter, akan mendapatkan kode akses keys dan token seperti gambar berikut ini :

Gambar 4.1 Keys & Token Twitter Developer

4.2 Hasil Crawling Data

Terlihat data csv hasil tarikan data berupa hastag sebanyak 100 tweets telah berhasil didapatkan, data tersebut hanyalah data yakni suatu , data memiliki bentuk yang kompleks seperti kata, huruf, angka, grafik tabel, gambar, serta lambang belum memiliki arti maupun sebuah informasi yang jelas, dari data inilah didapatkan untuk melakukan sentiment analysis yang berupa pro dan kontra mengenai vaksinasi tersebut dan klasifikasi naive bayes yang berupa satu metode statistik probabilitas berdasarkan penerapan Teorema Bayes dengan asumsi independen (naif) yang kuat, untuk memprediksi kelas dari suatu dokumen berdasarkan probabilitasnya.

4.3 Hasil Sentimen Analysis

Menjalankan perintah python dengan membuka melalui CLI atau Command Prompt (CMD) sebagai berikut:

Gambar 4.2. eksekusi python

Setelah berhasil menjalankan perintah python maka akan tampil inputan yang harus diisi

yakni memasukan kata kunci yakni berupa hastag yang diperoleh di twitter dan berapa banyak data yang akan diambil, disini penulis mengambil 100 data tweets sebagai sumber analisa yang dilakukan dalam tahap analisis.

Data yang diambil sebanyak 100 tweets kemudian diolah dalam algoritma python dan klasifikasi naive bayes terlihat hasil dari analisis yakni respown sentimen analisis netizen yang merespon positif sebanyak 8 %, positive sebanyak 21 %, weakly positive, 4 % strogly positive, 61 % metral, 1% negativ, 4% weakly negative, 1% strongly negative. Berikut hasil yang didapat:

Gambar 4. 3 Hasil sentimen analisis

Dapat dilihat dari hasil visualisasi melalui algoritma bahasa pemrograman python data sebanyak 100 hastag vaksinasi orang-orang di Indonesia dapat dilihat sebuah informasi grafik pie yang dimana informasi tersebut menghasilkan sebuah informasi yakni rata-rata masyarakat Indonesia memiliki persepsi netral terhadap vaksin tersebut sebanyak 61 %

5. Kesimpulan

Dari penelitian yang dilakukan, dapat disimpulkan bahwa :

- Penulis berhasil mendaftarkan akun twitter developer di akun resmi twitter dan berhasil juga mendapatkan token API key yang telah didapatkan secara legal
- Penulis berhasil mendapatkan crawling data sebanyak 100 tweets twitter mengenai kata kunci ataupun hastag vaksinasi
- Data sebanyak 100 merupakan data csv yang berhasil di analisa dan diklasifikasikan dalam parameter polarity yang telah ditetapkan dalam algoritma python. Berdasarkan data yang telah dianalisa melalui bantuan bahasa pemrograman python dapat dilihat bahwa klasifikasi naive bayes terlihat hasil dari analisis yakni respown sentimen analisis netizen yang merespon positif sebanyak 8 %, positive sebanyak 21 %, weakly positive, 4 % strogly positive, 61 % netral, 1% negativ, 4% weakly negative, 1% strongly negative.

Berdasarkan hasil visualisasi data informasi tersebut dapat disimpulkan vaksinasi tersebut

memiliki respon netral

Referensi

Antoni, D., & Akbar, M. (2019). E-supply chain management value concept for the palm oil industry. *Jurnal Sistem Informasi*, 15(2), 15-29.

Antoni, D., Fikari, D., & Akbar, M. (2018). The readiness of palm oil industry in enterprise resource planning. *Telkomnika*, 16(6), 2692-2702.

Antoni, D., Herdiansyah, M. I., Akbar, M., & Sumitro, A. (2021). Pengembangan Infrastruktur Jaringan Untuk Meningkatkan Pelayanan Publik di Kota Palembang. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 5(4), 1652-1659.

Antoni, D., Jie, F., & Abareshi, A. (2020). Critical factors in information technology capability for enhancing firm's environmental performance: case of Indonesian ICT sector. *International Journal of Agile Systems and Management*, 13(2), 159-181.

Bustami. 2014. Penerapan Algoritma Naïve Bayes untuk Mengklasifikasi Data nasabah Asuransi. *Jurnal informatika* Vol. 8, no. 1. Aceh Indonesia

Ebi, M., Yesi Novaria, K., and Ria, A. Eskplorasi trending topik twitter menggunakan text mining. *SENTIKOM 2017*.

Fauzi, F., Dencik, A. B., & Asiati, D. I. (2019). Metodologi Penelitian untuk manajemen dan akuntansi. *Jakarta: Salemba Empat*.

Masykur, F. (2014). Implementasi sistem informasi geografis menggunakan google maps api dalam pemetaan asal mahasiswa. *Simetris: Jurnal Teknik Mesin, Elektro dan Ilmu Komputer*, 5(2):181–186.

Rizal, (2017). Analisis sentimen tentang opini pilkada DKI 2017 tentang dokumen Twitter berbahasa indonesia menggunakan naïve bayes dan pembobotan emoji. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* vol. 1, no. 12, desember 2017, pages 1718-1724

Sembodo, J. E., Setiawan, E. B., and Baizal, Z. A. (2016). Data crawling otomatis pada twitter. In *Indonesian Symposium on Computing (Indo-SC)*, pages 11–16.

Simanullang, J. W., Adiwijaya, A., and Al Faraby, S. (2017). Klasifikasi sentimen pada movie review dengan metode multinomial nave bayes. *Proceedings of Engineering*, 4(2).

Yustiani and Yunanto, 2017. Peran market Place alternatif Bisnis di era Teknologi Informasi. *Jurnal ilmiah komputer dan informatika (KOMPUTA)* vol. 6, no. 2, Oktober 2017

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>)