

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pada tanggal 11 Maret 2020, Organisasi Kesehatan Dunia (WHO) telah menetapkan wabah virus corona baru (COVID-19) sebagai pandemi global dan mengoordinasikan upaya global untuk mengelola dampak dari COVID-19 (Cucinotta & Vanelli, 2020). Virus COVID-19 telah menyebar dengan sangat cepat ke seluruh dunia dan menginfeksi jutaan orang. Pada 2 Maret 2020, diumumkan kasus pertama COVID-19 di Indonesia oleh Presiden Joko Widodo. Sejak saat itu, kasus COVID-19 di Indonesia terus meningkat dan kehidupan masyarakat berubah secara drastis, dengan puncaknya pada bulan Juli 2021, di mana Indonesia mencatatkan kasus harian tertinggi, yaitu lebih dari 50.000 kasus (COVID-19, 2021). Presiden Jokowi telah mengeluarkan kebijakan Pembatasan Sosial Berskala Besar (PSBB) untuk membatasi aktivitas di tempat-tempat umum, seperti kantor, sekolah, dan kampus (Sahir et al., 2021). Pembatasan Sosial Berskala Besar (PSBB) di Indonesia diterapkan dalam beberapa gelombang, mulai dari PSBB awal di DKI Jakarta pada tahun 2020 hingga PSBB Mikro di seluruh Indonesia pada tahun 2021. Perubahan tersebut berdampak terhadap berbagai aspek kehidupan manusia, seperti cara bekerja, bersosialisasi, hingga dampak ekonomi. Pemerintah juga menetapkan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM) karena peningkatan jumlah kasus yang sangat signifikan. PPKM diterapkan pertama kali pada tanggal 11 Januari 2021 di tujuh provinsi di Pulau Jawa dan tanggal 2 Agustus 2021 di seluruh Indonesia dengan tujuan untuk mencegah penyebaran COVID-19 yang sedang meningkat pada saat itu. Berbagai macam kebijakan yang diterapkan diiringi kasus yang tak kunjung selesai mendorong masyarakat untuk menyuarakan pendapat mereka, baik positif maupun negatif, terhadap kebijakan tersebut.

Pada bulan Januari 2020, para ilmuwan dan ahli medis mulai melakukan pengembangan vaksin COVID-19 setelah virus SARS-CoV-2 diidentifikasi untuk pertama kalinya. Pengembangan vaksin COVID-19 dilakukan untuk membantu individu membentuk imunitas tubuh, sehingga diharapkan dapat mempercepat

tercapainya kekebalan kelompok (*herd immunity*) dengan tujuan akhirnya untuk menurunkan jumlah kasus yang terinfeksi (Nasir et al., 2021). Pada tanggal 11 Desember 2020, vaksin COVID-19 pertama kali disetujui untuk penggunaan di Amerika Serikat. Pengembangan vaksin umumnya membutuhkan waktu lama, hingga 15 tahun. Namun, dalam situasi darurat (pandemi), dapat dipersingkat menjadi 1,5 tahun atau lebih cepat. Meskipun penyingkatan waktu pengembangan vaksin dapat mempercepat akses masyarakat terhadap vaksin, hal ini juga berpotensi menimbulkan kekhawatiran dan tantangan, terutama terkait dengan penerimaan vaksin oleh publik dan kelompok anti-vaksin. Indonesia resmi memulai program vaksin COVID-19 pada bulan Januari 2021 yang tersedia untuk kelompok tertentu seperti tenaga kesehatan, kelompok rentan, dan petugas pelayanan publik. Sedangkan vaksin COVID-19 tersedia untuk masyarakat umum pada bulan Agustus 2021 (Niken et al., 2021). Berdasarkan hasil survei yang dilakukan oleh Kementerian Kesehatan, ITAGI, UNICEF, dan WHO menunjukkan bahwa, mayoritas masyarakat (75%) telah mengetahui tentang vaksin COVID-19 dan dua pertiganya (67%) bersedia menerima vaksin. Namun, tingkat kesediaan ini berbeda-beda di tiap provinsi. Faktor-faktor seperti status ekonomi, keyakinan agama, tingkat pendidikan, dan wilayah tempat tinggal memengaruhi tingkat penerimaan vaksin. Masyarakat dengan status ekonomi dan pendidikan yang lebih tinggi, serta memiliki akses informasi yang lebih banyak tentang vaksin, seperti mereka yang memiliki asuransi kesehatan, cenderung lebih bersedia menerima vaksin. Meskipun mayoritas bersedia divaksinasi, terdapat 10% yang ragu dan 7% yang menolak. Faktor keamanan, efektivitas, kehalalan, keterbatasan penyebaran informasi, dan kebenaran informasi menjadi pertimbangan utama bagi mereka yang menolak. Secara geografis, Papua memiliki tingkat penerimaan vaksin tertinggi (75%), diikuti oleh Jawa dan Kalimantan. Aceh memiliki tingkat penerimaan terendah (46%), bersama dengan beberapa daerah di Sumatera, Sulawesi, dan Maluku. Hasil survei tersebut menunjukkan bahwa edukasi dan informasi yang akurat tentang vaksin COVID-19 diperlukan untuk meningkatkan tingkat penerimaan di seluruh wilayah Indonesia (Rokom, 2021).

Media sosial memiliki peran yang besar dalam kehidupan manusia dan berjalannya program vaksinasi COVID-19. Media sosial merupakan aplikasi atau laman berbasis internet untuk berinteraksi dan berbagi informasi atau tanggapan antar pengguna dalam jangka waktu yang singkat (Harun & Ananda, 2021). Berdasarkan jurnal BMJ Global Health, media sosial memiliki peran ganda, dimana dapat menjadi sumber informasi yang cepat dan mudah diakses, tetapi juga dapat menjadi sarana penyebaran informasi yang salah dan ketakutan, yang dapat meningkatkan keraguan dan penolakan terhadap vaksin (Cascini et al., 2022). Selama pandemi, banyaknya orang yang berada di rumah, sehingga penggunaan media sosial, seperti Twitter meningkat. Twitter adalah media sosial untuk komunitas *online* di mana pengguna dapat menyampaikan pesan kapan saja dengan cepat. Pesan yang dikirim dari akun pengguna disebut dengan tweet. Tweet dapat berisi teks, tautan, foto, atau video. Twitter menjadi salah satu *platform* penting dalam penyebaran informasi terkait vaksinasi COVID-19 dengan penggunanya lebih dari 16 juta pada tahun 2021 dan selalu meningkat tiap tahunnya (Kemp, 2021). *Platform* ini memungkinkan individu untuk berbagi pengalaman pribadi dengan vaksin, membagi informasi dan opini yang dapat meningkatkan kepercayaan dan mendorong orang lain untuk divaksinasi. Twitter juga dapat menjadi *platform* penyebaran misinformasi tentang vaksin COVID-19 dan memanipulasi opini publik, baik dari akun pribadi maupun akun palsu (Cascini et al., 2022).

Analisis sentimen merupakan teknik untuk mengolah data tekstual untuk mengidentifikasi sentimen dan memahami emosi dalam sebuah teks (Sari & Wibowo, 2019). Analisis sentimen membantu dalam proses menilai sentimen dari masyarakat sehingga dapat memberikan masukan atau umpanbalik bagi penggunanya. Umpanbalik tersebut dapat menjadi sebuah faktor pendukung dalam pengambilan keputusan. Analisis sentimen dapat digunakan untuk memahami persepsi publik yang terdiri dari informasi dan opini pengguna Twitter terhadap vaksin COVID-19. Selain itu, analisis sentimen mampu melacak perubahan persepsi publik tentang vaksin COVID-19 secara berkelanjutan dan membantu mengidentifikasi sumber informasi yang menyesatkan atau tidak akurat terkait vaksin tersebut (Pristiyono et al., 2021). Analisis sentimen

mengelompokkan teks ke dalam kategori positif, negatif, atau netral, berdasarkan perasaan yang diungkapkan. Dengan menganalisis sentimen data tweet, hasil analisis dapat mengidentifikasi pola yang menunjukkan sentimen dari data tersebut (Sharma, 2023). Analisis sentimen juga dapat menjadi alat yang berharga untuk memahami persepsi publik terhadap vaksin COVID-19 dan dampak pandemi di tahun 2024. Dengan menggunakan analisis sentimen, pemerintah, organisasi kesehatan, platform media sosial, dan peneliti dapat mengambil langkah-langkah yang tepat untuk meningkatkan tingkat vaksinasi, memerangi misinformasi, dan memberikan layanan kesehatan yang dibutuhkan masyarakat.

Machine learning merupakan salah satu metode yang digunakan untuk melakukan klasifikasi sentimen. *Machine learning* adalah bagian dari *artificial intelligence* yang melatih komputer untuk belajar dari data tanpa diprogram secara langsung. Komputer akan belajar dari data yang diberikan untuk menemukan pola dan hubungan antar data atau kata, lalu digunakan untuk membuat keputusan (Indulkar & Patil, 2021). *Machine learning* ini menggunakan jenis pembelajaran *supervised learning*, dimana model yang dibangun dilatih menggunakan *dataset* berlabel. Label yang digunakan, yaitu positif, netral, dan negatif. Setelah dilatih, algoritma *machine learning* dapat digunakan untuk mengklasifikasi data baru (Min et al., 2023). Untuk meningkatkan kinerja model klasifikasi, digunakan Word2Vec. Word2Vec adalah model yang digunakan untuk merepresentasi kata-kata dalam bentuk vektor untuk memahami makna dan hubungan antara kata-kata tersebut. Word2Vec juga mudah digunakan dan menghasilkan vektor kata yang berkualitas tinggi (Li et al., 2019). Dengan Word2Vec, model dapat melakukan klasifikasi data baru yang lebih akurat.

Algoritma *machine learning* (klasifikasi) yang digunakan pada penelitian ini yaitu, *Support Vector Machine* (SVM), *Naive Bayes*, dan *Extreme Gradient Boost* (XGBoost). *Support Vector Machine* (SVM) adalah algoritma populer untuk menyelesaikan masalah klasifikasi dengan mengklasifikasi data ke dalam kelas yang berbeda dengan menemukan garis (*hyperplane*) yang memisahkan data dengan label berbeda (Alzubi et al., 2018). *Naive Bayes* adalah algoritma klasifikasi yang menghitung sekumpulan probabilitas dengan menghitung frekuensi dan kombinasi nilai dalam kumpulan data yang diberikan (Saritas &

Yasar, 2019). *Extreme Gradient Boost (XGBoost)*, merupakan algoritma klasifikasi berbasis *gradient boosting decision tree* yang digunakan untuk mengurangi kumpulan data untuk klasifikasi. Penelitian ini melakukan komperasi antara ketiga algoritma *machine learning* tersebut untuk memilih model algoritma yang terbaik, dan meningkatkan kinerja model algoritma tersebut. Aplikasi diharapkan dapat membantu dalam proses penyaringan informasi menjadi lebih cepat dan mudah karena dapat dilakukan secara kolektif dan masif. Sistem klasifikasi data dengan model terbaik dapat membantu mempercepat proses analisis sentimen dengan mengklasifikasikan tweet secara otomatis. Dibangun aplikasi untuk melakukan simulasi analisis sentimen menggunakan model algoritma terbaik.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah yang telah dijelaskan sebelumnya, maka masalah yang dirumuskan pada penelitian ini sebagai berikut :

1. Bagaimana menganalisis tweet untuk mengetahui sentimen positif dan negatif masyarakat terhadap program vaksinasi COVID-19 di Indonesia pada periode Maret 2021 hingga April 2022?
2. Bagaimana kinerja algoritma *Support Vector Machine*, Naive Bayes, dan *Extreme Gradient Boost* dalam melakukan analisis sentimen terhadap lebih dari 5.000 tweet terkait program vaksinasi COVID-19 di Indonesia, dengan menggunakan metrik akurasi, presisi, dan *recall*?
3. Algoritma mana yang terbaik untuk melakukan analisis sentimen terhadap tweet terkait program vaksinasi COVID-19 di Indonesia menggunakan metrik akurasi, presisi, dan *recall* sebagai metrik evaluasinya?
4. Bagaimana merancang dan mengembangkan aplikasi *web* untuk mengimplementasikan model analisis sentimen pada tweet terkait program vaksinasi COVID-19 di Indonesia?

1.3 Batasan Masalah

Dari rumusan masalah tersebut, maka dalam penelitian ini terdapat beberapa batasan masalah pada sebagai berikut :

1. Data yang digunakan adalah data tweet tanggapan masyarakat terhadap program vaksinasi COVID-19 oleh Kementerian Kesehatan pada tahun 2021 yang diambil dari Twitter.
2. Data yang digunakan adalah data yang berbahasa Indonesia yang dikumpulkan dari bulan Maret 2021 hingga April 2022 untuk proses *training* dan *testing* model.
3. Penulis menggunakan metode *machine learning*, yaitu *Support Vector Machine*, *Naive Bayes*, dan *Extreme Gradient Boost*.
4. Sistem dirancang dengan menggunakan bahasa pemrograman Python dan *framework* Flask.

1.4 Tujuan

Tujuan dari penelitian ini adalah sebagai berikut :

1. Menerapkan analisis sentimen untuk mengetahui emosi dan mengklasifikasi sentimen masyarakat terhadap program vaksinasi COVID-19.
2. Membangun model Word2Vec untuk data tweet sentimen masyarakat terhadap program vaksinasi COVID-19 dan menggunakannya untuk meningkatkan kinerja algoritma klasifikasi sentimen.
3. Membandingkan kinerja algoritma *machine learning* *Support Vector Machine* (SVM), *Extreme Gradient Boosting* (XGBoost), dan Naive Bayes untuk mengklasifikasi sentimen tweet terhadap program vaksinasi COVID-19.
4. Mengembangkan aplikasi *web* untuk melakukan simulasi klasifikasi sentimen tweet terhadap program vaksinasi COVID-19.

1.5 Manfaat

Adapun manfaat yang diharapkan didapat dari penelitian ini adalah :

1. Dapat mengklasifikasi tanggapan masyarakat Indonesia terhadap program vaksinasi COVID-19 yang diadakan oleh Kementerian Kesehatan dalam upaya mengurangi kasus COVID-19.
2. Dapat membantu pemerintah memahami persepsi dan memproses opini masyarakat dengan otomatis sehingga dapat dilakukan lebih cepat dan

dibandingkan secara manual yang memakan waktu lebih lama agar dapat membuat kebijakan selanjutnya.

1.6 Sistematika Penulisan

Untuk memberi gambaran yang jelas tentang penelitian ini, maka disusunlah sistematika penulisan yang berisi materi yang akan dibahas pada setiap bab. Sistematika dalam penulisan tugas akhir ini adalah sebagai berikut :

BAB I PENDAHULUAN

Pada bab pendahuluan menjelaskan mengenai latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, luaran penelitian dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini berisi penjelasan mengenai teori - teori dan penelitian terkait yang menjadi dasar penelitian ini seperti mengenai program vaksinasi COVID-19, *machine learning*, *text processing*, *Support Vector Machine*, *Naive Bayes*, *Extreme Gradient Boost*, *Word2Vec*, analisis sentimen, dan aplikasi.

BAB III METODOLOGI PENELITIAN

Pada bab ini diuraikan secara rinci mengenai tahapan penelitian, alat dan bahan, serta prosedur implementasi yang dilakukan.

BAB IV ANALISIS PERANCANGAN DAN PEMODELAN

Pada bab ini berisi tentang analisis model *machine learning* yang dibangun untuk mendapatkan kinerja terbaik dan melakukan analisis sentimen.

BAB V IMPLEMENTASI DAN PENGUJIAN

Pada bab ini membahas hasil dari penelitian yang dilakukan dan analisa terhadap implementasi dan pengujian hasil penelitian tersebut.

BAB VI PENUTUP

Pada bab ini merupakan penutup dari penelitian yang berisi hasil penelitian yang sudah dilakukan yang kemudian disajikan dalam bentuk kesimpulan dan saran.