

BAB I PENDAHULUAN

1.1 Latar Belakang

Coronavirus Disease (Covid-19) merupakan salah satu topik yang masih hangat dibicarakan sejak kemunculannya di penghujung tahun 2019 hingga saat ini. Virus ini mula terdeteksi di kota Wuhan, Tiongkok [1]. Sejak dinyatakan sebagai pandemi Covid-19 pada 11 Maret 2020 oleh Organisasi Kesehatan Dunia (WHO), Covid-19 telah menyebar luas ke seluruh dunia. Berdasarkan data Organisasi Kesehatan Dunia (WHO) pada 21 September 2021, lebih dari 200 negara telah terkonfirmasi kasus Covid-19. Sehingga beberapa negara melakukan pembatasan sosial agar tidak menyebar luasnya Covid-19, salah satunya di Indonesia. Indonesia merupakan negara yang terkena dampaknya paparan Covid-19 yang masuk pada awal maret 2020. Adanya Covid-19 yang masuk ke Indonesia membuat beberapa daerah mulai melakukan pembatasan sosial, dampak dari penyebaran Covid-19 diberbagai negara menyebabkan sektor perekonomian menjadi turun drastis, Pendidikan, dan kehidupan sosial masyarakat [2]. Melihat pesatnya penyebaran Covid-19 di Indonesia dengan ini harus segera ditangani, salah satunya dengan mengembangkan vaksin. Pada kegiatan vaksinasi yang akan diberikan ke masyarakat, pemerintah juga terlibat penting dalam perencanaan itu. Tanggal 5 Oktober 2020 Presiden Jokowi Widodo meresmikan praturan Presiden (Perpres) [3]. Dengan adanya program pemerintah tentang vaksinasi, bertujuan untuk menekan jumlah kasus Covid-19 di Indonesia.

Analisis sentimen atau disebut juga dengan Sentimen Respon, yang bertujuan untuk menganalisis, mengolah, mengekstrak data tekstual berupa tanggapan terhadap suatu entitas pada suatu topik untuk memperoleh informasi. Teks dapat di klasifikasikan ke sebuah polaritas (positif, netral, dan negatif) [4].

Beberapa penelitian sebelumnya sistem analisis sentimen sudah banyak yang melakukan penelitian tersebut, namun masih belum mendapatkan hasil yang optimal. Seperti pada penelitian Tentang “Analisis

Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine” 2018 [5], menunjukkan metode Support Vector Machine mendapatkan hasil akurasi tertinggi sebesar 90% pada komposisi data latih, *precision* sebesar 94,44%, 85% recall sebesar dan f-measure sebesar 89,47% dengan komposisi data latih 50% dan data uji 50%. Sedangkan penelitian sebelumnya, dipaparkan oleh Brian Laurensz dan Eko Sedyono 2021 [6], menunjukkan rata-rata nilai terbaik di dapatkan Support Vector Machine di bandingkan metode lain. Hasil persentase akurasi sebesar 84,58%, recall 85,82%, dan presisi 82,14%. Oleh karena itu saya menerapkan metode “Support Vector Machine (SVM)” pada kasus ini dengan harapan dapat menghasilkan akurasi yang terbaik juga.

Oleh karena itu, penelitian ini mencoba merancang sistem dengan menggunakan Tf-Idf dan Svm untuk analisis sentimen terhadap vaksinasi Covid-19. Tujuan adalah dapat melakukan klasifikasi opini positif dan negatif dengan harapan mampu menghasilkan akurasi yang tinggi.

1.2 Perumusan Masalah

1.2.1 Permasalahan

Pada permasalahan yang saya lakukan ini dapat membantu pemerintah dalam menganalisis sentimen atau pendapat masyarakat terkait dengan vaksinasi Covid-19.

1.2.2 Metode Usulan

Berdasarkan latar belakang dan permasalahan yang telah di uraikan, pada penelitian ini metode usulan yang digunakan adalah metode TF-IDF dan SVM. Pada metode TF-IDF digunakan untuk pembobotan kata sedangkan metode SVM digunakan untuk mengklasifikasi sentimen tentang vaksinasi Covid-19.

1.2.3 Pertanyaan Penelitian

Bagaimana performa sistem yang dibangun menggunakan metode TF-IDF dan SVM untuk Sentimen analisis terhadap vaksinasi covid-19 ?

1.3 Tujuan & Manfaat

1.3.1 Tujuan Penelitian

Tujuan pada penelitian ini adalah melakukan sentimen analisis terhadap vaksinasi covid-19 dan mengetahui nilai persentase yang didapatkan dengan menerapkan metode TF-IDF sebagai pembobotan kata (Term Weighting) dan metode SVM sebagai klasifikasi berdasarkan *Accuracy*, *Precision*, *Recall*, *F1-score*.

1.3.2 Manfaat Penelitian

Manfaat yang dapat diberikan pada penelitian ini adalah :

Dapat menjadikan metode usulan untuk penelitian yang akan datang serta menjadi rujukan untuk mengembangkan sistem dengan akurasi yang lebih tinggi pada penelitian sebelumnya

1.4 Batasan-batasan

Berdasarkan permasalahan yang telah diperoleh di atas, maka penulis harus membatasi masalah agar tidak keluar dari topik yang dibahas. Adapun Batasan masalah sebagai berikut :

1. Data digunakan dari bulan Oktober-November 2020 yang berasal dari penelitian sebelumnya.
2. Opini yang digunakan Bahasa Indonesia
3. Keyword yang digunakan “Vaksin Corona” dan “Vaksin Covid” karena lebih dominan muncul atau sering digunakan di Twitter
4. Dataset yang digunakan adalah 2628 data dalam bentuk csv dan data dibagi menggunakan K-Fold Cross Validation dengan jumlah K=10 diantaranya data train sebanyak 2365 dan data test sebanyak 263 komentar.
5. Metode pembobotan kata menggunakan metode TF-IDF dan metode klasifikasi menggunakan metode SVM.
6. Data yang dihasilkan di bagi 2 kategori yaitu positif dan negatif
7. Pengukuran akurasi yang digunakan yaitu “*Confusion Matrix*” (*accuracy*, *precision*, *recall*, dan *f1-score*) dan “*K-Fold Cross Validation*”.
8. Sistem yang digunakan berbasis web.

1.5 Sistematika Proposal

Sistematika proposal ini terdiri dari :

Bab I Pendahuluan

Bab ini membahas tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah, dan sistem yang diusulkan.

Bab II Tinjauan Pustaka

Bab ini menjelaskan teori yang mendasari penelitian, dan dasar untuk mengatasinya pertanyaan.

Bab III Metode Usulan

Bab ini menjelaskan langkah-langkah penelitian yang akan dilakukan.

Bab IV Hasil Dan Penelitian

Bab ini menjelaskan hasil implementasi sistem berdasarkan hasil desain sebelumnya.

Bab V Penutup

Bab ini menjelaskan kesimpulan dari hasil dan pembahasan serta saran untuk pengembangan sistem selanjutnya.

Daftar Pustaka

Bab ini berisi referensi-referensi yang di gunakan.