

BAB I PENDAHULUAN

1.1 Latar Belakang

Corona virus atau yang sering disebut dengan COVID-19 merupakan salah satu penyakit menular yang disebabkan oleh *corona virus* yang ditemukan pada akhir Desember tahun 2019 di kota Wuhan, Tiongkok. Sebagian besar orang yang terinfeksi COVID-19 akan mengalami gangguan pernapasan ringan hingga gejala yang berat. Dalam rangka memutus penularan COVID-19, Pemerintah Indonesia telah melakukan vaksinasi tahap 1 dan tahap 2. Kemudian saat ini pemerintah sedang melaksanakan program vaksinasi tahap ketiga atau *booster* yang bertujuan untuk lebih meningkatkan kekebalan tubuh dan sistem imun.

Vaksinasi *booster* ini menjadi perbincangan hangat di kalangan masyarakat baik secara langsung maupun melalui media sosial, salah satunya melalui *twitter*. Media sosial tersebut merupakan *platform* situs jejaring sosial yang berkembang secara pesat. Pengguna *twitter* dapat saling berinteraksi melalui komputer ataupun *smartphone* dari mana pun dan kapan pun. Setelah diluncurkan pada Juli 2006, jumlah pengguna *twitter* meningkat sangat pesat. Pada Januari 2021, diperkirakan jumlah pengguna *twitter* yang terdaftar sekitar 187 juta pengguna aktif (Lestari & Saepudin, 2021).

Pemanfaatan data yang bersumber dari media sosial merupakan suatu terobosan baru yang dapat dijadikan sebagai alternatif sumber data yang dapat menggantikan *survey* tradisional. Pada satu sisi, pengumpulan data respon dan opini masyarakat secara langsung/*real time* menggunakan *survey* tradisional dinilai sulit untuk dilakukan. Sedangkan pada sisi lain, pengumpulan data melalui media sosial dinilai dapat memberikan efisiensi dalam segala hal. *Efisiensi* tersebut mencakup biaya yang minimal dalam memperoleh data, data diperoleh secara *real time*, dan menghasilkan data yang dapat menggambarkan opini masyarakat yang sebenarnya.

Terkait vaksinasi *booster*, banyak persepsi, pandangan dan opini dari masyarakat. Sebagian diantaranya menyetujui penggunaan vaksin *booster* sebagai solusi penanganan wabah COVID-19, dan sebagian lainnya menolak dengan alasan takut akan adanya efek samping serta berbagai alasan lainnya. Dari beragamnya persepsi, pandangan dan opini masyarakat tersebut kemudian penulis mencoba melakukan analisis sentimen untuk mendapatkan simpulan yang dapat menjadi informasi penting bagi pemerintah dalam menentukan langkah penanganan COVID-19 berikutnya (Fathur Rachman & Pramana, 2021).

Analisis sentimen merupakan metode yang dapat digunakan untuk menentukan apakah tulisan seseorang termasuk dalam katageori positif atau negatif. Metode ini juga dikenal sebagai *opinion mining*, yang bertujuan untuk menganalisis pendapat atau sikap masyarakat terhadap suatu masalah tertentu. Terdapat banyak algoritma yang dapat digunakan untuk melakukan analisis sentimen, salah satunya *Naïve Bayes* (Alhajji et al., 2020).

Algoritma *Naïve Bayes* merupakan penggolongan menggunakan statistik sederhana berdasarkan *teorema bayes* yang mengasumsikan keberadaan atau ketiadaan dari suatu kelas dengan fitur lainnya. Beberapa keuntungan algoritma *Naïve Bayes* diantaranya adalah sangat cocok untuk digunakan pada klasifikasi data teks, dapat digunakan pada *dataset* yang besar, serta bekerja sangat baik pada *multi-class prediction* (Laurensz & Eko Sedyono, 2021).

Penelitian ini diarahkan untuk mencari simpulan dari opini masyarakat terkait pelaksanaan vaksinasi *booster* di Indonesia berdasarkan data yang tersebar di *twitter* dengan memanfaatkan algoritma *Naïve Bayes*. Dengan demikian penelitian ini diberi judul **“Analisis Sentimen Vaksin Booster Covid-19 Menggunakan Algoritma Naïve Bayes Pada Media Sosial Twitter”**.

1.2 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah “apakah algoritma *Naïve Bayes* efektif digunakan untuk analisis sentimen opini pengguna *twitter* terkait vaksin *booster* COVID-19 ?”

1.3 Batasan Masalah

Penelitian ini dibatasi pada hal-hal berikut.

- a. Data yang dianalisis adalah opini masyarakat terkait vaksin *booster* di Indonesia yang disampaikan melalui media sosial *twitter*.
- b. Penelitian ini menggunakan *tools rapid miner* dengan menerapkan algoritma *Naïve Bayes*.
- c. Evaluasi model dilakukan menggunakan *confusion matrix* yang terdiri dari: *accuracy*, *precision*, dan *recall*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengevaluasi tingkat *accuracy*, *precision*, dan *recall* dari algoritma *Naïve Bayes* dalam melakukan analisis sentimen terhadap opini masyarakat yang ada di *twitter* terkait masalah vaksin *booster*.

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah sebagai berikut.

- a. Memberikan pengetahuan kepada peneliti maupun akademis tentang kemampuan algoritma *Naïve Bayes* dalam bidang analisis sentimen, khususnya untuk *dataset* yang diperoleh dari media sosial *twitter*.
- b. Memberikan simpulan tentang opini masyarakat terkait pelaksanaan vaksin *booster* di Indonesia yang dapat digunakan pemerintah sebagai bahan evaluasi dan perencanaan tindak lanjut berikutnya.

1.6 Sistematika Penulisan

Penulisan proposal ini terdiri dari 3 (tiga) bab dengan sistematika penulisan sebagai berikut.

a. Bab I Pendahuluan.

Pada bab ini diuraikan latar belakang Covid-19, vaksin booster, analisis sentimen, rumusan masalah, batasan masalah, tujuan dan manfaat penelitian dan sistematika penulisan.

b. Bab II Tinjauan Pustaka

Pada bab ini berisi tentang analisis sentimen, covid-19, vaksin, text mining, klasifikasi, algoritma Naïve bayes, twitter, twitter api, rapidminer, crawling data, confusion matrix dan penelitian terkait.

c. Bab III Metodologi Penelitian

Bab ini menjelaskan tentang metodologi penelitian, pengumpulan data, analisa data, pembersihan data, pelebelan data, pembagian data dan implementasi algoritma naïve bayes.