

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Pada tahun 2019 akhir menjadi awal mula munculnya virus korona di Kota Wuhan, China. Wabah virus yang disebut COVID-19 ini telah melanda Negara Indonesia hampir 2 tahun lamanya yang menyebabkan banyaknya jumlah kasus kematian di Indonesia. COVID-19 adalah penyakit yang menyerang saluran pernafasan dan mengakibatkan korbannya mengalami sesak nafas. Penyebaran virus yang sangat pesat dengan melalui kontak langsung antara orang satu ke orang lainnya ataupun melalui batuk dan bersin.

Dilihat dari mudahnya penyebaran COVID-19 yang sangat cepat, pemerintah Indonesia mewaspadai potensi lonjakan baru dengan cara mengadakan program vaksinasi secara gratis untuk menghentikan tingkat penyebaran virus. Vaksin yang dibagikan ke masyarakat diharapkan dapat melindungi atau mengurangi kemungkinan terkenanya COVID-19. Adapun aspek yang perlu dipertimbangkan oleh pemerintah dalam program kegiatan vaksinasi tersebut yaitu, kelayakan vaksin, efek samping setelah pemakaian, hingga tahapan dan prosedur yang benar.

Dengan adanya program vaksinasi dari pemerintah untuk masyarakat Indonesia, banyak kalangan yang memberikan respon dan opininya di berbagai media sosial. Jumlah pengguna media sosial di Indonesia bisa mencapai 150 juta lebih. Media sosial kini menjadi suatu hal yang wajib dimiliki oleh setiap orang karena banyaknya informasi yang bisa didapatkan. Salah satu media sosial yang

banyak digunakan oleh masyarakat Indonesia adalah twitter. Masyarakat dapat membagikan hal yang sedang dirasakan melalui postingan dan mendapat informasi dalam hal apapun pada twitter. Dengan hal ini dapat dilakukan analisis sentimen dari masyarakat terhadap vaksin dengan memanfaatkan data tweets pada media sosial twitter dan melakukan klasifikasi pendapat dan opini kedalam 3 kelas yaitu positif, netral, dan negatif. Analisis sentimen ini juga dapat dijadikan referensi atau bahan evaluasi untuk meningkatkan suatu layanan atau produk.

Pada penelitian ini akan menganalisis bagaimana respon masyarakat Indonesia terhadap vaksin COVID-19 dengan menggunakan data yang bersumber dari media sosial twitter dan melakukan klasifikasi menggunakan algoritma *Naïve Bayes*, *Decision Tree C4.5*, dan metode pendekatan *CRISP-DM* (Cross-Industry Standard Process for Data Mining). Algoritma tersebut digunakan karena dapat menghasilkan akurasi yang baik dalam penelitian.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah disampaikan, maka perlu dirumuskan suatu masalah sebagai berikut :

1. Apakah metode pendekatan *CRISP-DM* dapat digunakan sebagai tahapan analisis sentimen terhadap vaksin pada twitter?
2. Apakah tingkat akurasi algoritma *Naïve Bayes* lebih baik dibandingkan *Decision Tree* dalam pengklasifikasian sentimen terhadap vaksin pada twitter?
3. Apakah jumlah dataset dapat mempengaruhi tingkat akurasi algoritma dalam pengklasifikasian sentimen terhadap vaksin pada twitter?

1.3 Batasan Masalah

Batasan-batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Jumlah data tweet yang akan diambil sebanyak 3.000 tweet.
2. Pengambilan data tweet menggunakan tools Twint.
3. Data tweet yang diambil dan dianalisis hanya tweet yang berbahasa Indonesia.
4. Klasifikasi sentimen terhadap vaksin pada tweet dibagi menjadi 3 kelas, yaitu positif, netral, dan negatif.
5. Metode yang digunakan untuk klasifikasi data tweet adalah algoritma *Naïve Bayes*, *Decision Tree* dan pendekatan *CRISP-DM*.
6. Pengolahan data analisis sentimen pada tweet menggunakan bahasa pemrograman python.
7. Hasil penelitian ini adalah mengukur tingkat akurasi dari algoritma *Naïve Bayes* dan *Decision Tree*.
8. *Confusion Matrix* digunakan untuk mengevaluasi performa klasifikasi dari algoritma *Naïve Bayes* dan *Decision Tree*.

1.4 Maksud dan Tujuan Penelitian

Adapun tujuan yang akan dicapai dari penelitian ini adalah sebagai berikut:

1. Untuk mengukur tingkat akurasi dari algoritma *Naïve Bayes* dan *Decision Tree* dalam pengklasifikasian.
2. Untuk mengetahui sentimen masyarakat terhadap vaksin menggunakan algoritma *Naive Bayes* dan *Decision Tree* pada media sosial twitter.

1.5 Manfaat Penelitian

Berdasarkan dengan latar belakang masalah dan tujuan dari penelitian, penulis mengharapkan penelitian ini dapat memberikan beberapa manfaat sebagai berikut:

1. Memberikan informasi sentimen masyarakat terhadap vaksin pada twitter yang dibagi menjadi positif, netral, dan negatif.
2. Penulis mendapatkan pengetahuan lebih dalam menganalisa sentimen masyarakat terhadap vaksin menggunakan algoritma *Naive Bayes* dan *Decision Tree*.
3. Hasil penelitian dapat menjadi bahan referensi untuk peneliti lain dengan menggunakan topik yang sama.

1.6 Metode Penelitian

Dalam menyusun penelitian ini, metode yang digunakan adalah sebagai berikut :

1.6.1 Metode Analisis

Metode pendekatan yang digunakan untuk menganalisis sentimen dalam penelitian ini yaitu CRISP-DM, yang meliputi :

a. Business Understanding

Pada tahap ini menentukan bagaimana mendapatkan data, menentukan tujuan dan batasan dalam rumusan masalah data mining.

b. Data Understanding

Tahapan kedua yaitu memeriksa atau mengidentifikasi data yang sudah diambil.

c. Data Preparation

Tahapan ini meliputi, membersihkan data dan mengurangi jumlah data dengan cara menormalisasi jika diperlukan.

d. Modelling

Tahapan keempat membuat model prediktif dan deskriptif dengan menerapkan teknik dan algoritma data mining terhadap dataset menggunakan bantuan tools.

e. Evaluation

Pada tahap ini melakukan evaluasi terhadap model yang diterapkan agar model sesuai dengan tujuan yang ingin dicapai.

1.6.2 Metode Klasifikasi

Pada penelitian ini metode yang digunakan untuk klasifikasi yaitu algoritma *Naïve Bayes*. Algoritma *Naïve bayes* termasuk ke dalam *supervised learning* dan keuntungan dari *Naïve Bayes* yaitu hanya membutuhkan sedikit *data training* untuk memperkirakan parameter yang diperlukan untuk klasifikasi.

Sedangkan algoritma *Decision Tree* dalam penelitian ini digunakan untuk melakukan klasifikasi dan prediksi sentimen masyarakat terhadap vaksin apakah termasuk ke dalam kelas positif, netral, atau negatif.

1.6.3 Metode Evaluasi

Evaluasi dilakukan terhadap kualitas dan efektivitas suatu model yang diterapkan untuk mencapai tujuan awal. Metode yang digunakan untuk mengevaluasi yaitu *Confusion Matrix*. Metode *Confusion Matrix* dilakukan untuk menghitung nilai *accuracy*, *precision*, dan *recall* sebagai hasil akhir.

1.7 Sistematika Penulisan

Dalam menyusun penelitian ini, sistematika penulisan yang digunakan oleh penulis adalah sebagai berikut:

BAB I PENDAHULUAN

Dalam bagian pendahuluan berisi latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metode penelitian, sistematika penulisan.

BAB II LANDASAN TEORI

Pada landasan teori berisi tinjauan pustaka, dasar-dasar teori yang digunakan dan dasar teori kutipan.

BAB III METODE PENELITIAN

Pada bab 3 berisi hipotesis, alur penelitian, peralatan dan bahan serta rancangan system.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini berisi implementasi tahapan yang dilakukan penulis dalam menganalisis sentimen dan membahas hasil dari penelitian yang telah diperoleh.

BAB V PENUTUP

Dalam bagian penutup berisi kesimpulan dan saran yang didapat penulis dari hasil penelitian yang berdasarkan rumusan masalah.