

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan dunia teknologi informasi dan komunikasi yang pesat tidak terlepas dari penyedia layanan web yang menyediakan informasi yang beragam. Penyebaran data berupa data teks dapat dijadikan sumber yang sangat potensial untuk digali lebih dalam dan salah satu contohnya adalah data text yang diambil dari *twitter*. Twitter merupakan layanan jejaring sosial dan *microblog* yang memungkinkan penggunanya untuk mengirim dan membaca pesan berbasis text hingga 140 karakter yang dikenal dengan istilah *tweet*. *Twitter* didirikan oleh Jack Dorsey pada tahun 2006 dan menjadi salah satu dari sepuluh situs yang paling dikunjungi di internet. Pada kuartal I pada tahun 2017 pengguna twitter bertambah hingga 6% atau 9 juta sehingga secara keseluruhan pengguna twitter mencapai sekitar 328 juta jiwa. *Twitter* juga jejaring social yang memiliki peran besar dalam mempengaruhi aspek dari segi social, komunikasi, psikologi, marketing dan juga politik. Setiap *tweet* yang di posting pengguna beraneka ragam sesuai dengan keinginan pengguna. *Tweet* itu bisa berupa pendapat, saran, ataupun kritikan tentang topik-topik tertentu. Keanekaragaman postingan tersebut serta banyaknya penggunaan bahasa yang tidak baku pada *tweet* menjadi alasan diperlukan Sentimen analisis.

Sentiment Analisis atau *opinion mining* adalah jenis *natural language* pengolahan kata untuk melacak *mood* atau *ekspresi* masyarakat terhadap suatu topik atau produk tertentu (Vinodhini & Chandrasekaran, 2012). Analisis sentiment juga bertujuan untuk menganalisis, memahami, mengolah, dan mengekstrak data tekstual (Liu, 2010) Analisis sentimen digunakan untuk melihat kecenderungan suatu sentimen, apakah sentiment tersebut cenderung beropini positif atau negatif. Sebelum melakukan analisa sentimen dilakukan preprocessing teks dengan metode *text mining* yang meliputi data cleansing, case folding, tokenizing, stopwords, dan stemming.

Terdapat banyak metode klasifikasi sentiment yang digunakan untuk analisis sentimen namun metode yang banyak digunakan pada klasifikasi teks adalah metode *Naïve Bayes Classifier* (NBC). Metode NBC banyak digunakan karena metode ini sangat efektif pada klasifikasi data karena dengan algoritma yang sederhana tetapi mampu memberikan tingkat akurasi yang tinggi.

Para peneliti yang telah mencoba menerapkan sentiment analisis dalam berbagai kasus. Fajar Ratnawati melakukan penelitian Implementasi Algoritma Naive Bayes Terhadap Analisis Sentimen Opini Film Pada Twitter (Ratnawati, 2018). Faishol Nurhada, Sari widya Sihwi, Afrizal Doewes melakukan penelitian Analisis Sentimen Masyarakat terhadap Calon Presiden Indonesia 2014 berdasarkan Opini dari Twitter Menggunakan Metode Naive Bayes Classifier (Nurhada, et al., 2013). Robet Habibi, Djoko Budiyanto Setyohadi, Ernawati melakukan penelitian Analisis Seentimen Pada Twitter Mahasiswa Menggunakan Metode Backpropagation (Habibi, et al., 2016). Dwatra Apriandi, M. Irwan Afandi, Eka Dyar Wahyuni melakukan penelitian Analisis Sentimen pelanggan WIFI.ID Pada Twitter Dengan Support Vector Machine (Ariandi, et al., 2016). Nur Khotimah melakukan penelitian Analisis Sentimen Terhadap Review E-Commerce Dengan Metode Stochastic Gradient Descent (Khotimah, 2019).

Berdasarkan latar belakang yang telah dikembangkan diatas maka dilakukan penelitian yang berjudul **“SENTIMENT ANALYSIS TWEET VAKSINASI COVID-19 MENGGUNAKAN METODE NAÏVE BAYES CLASSIFIER”**.

1.1 Rumusan Masalah

Berdasarkan latar belakang maka rumusan masalah pada tugas akhir ini adalah sebagai berikut :

1. Bagaimana data Twitter digunakan sebagai sentimen analisis pada suatu topik pembahasan yang sedang terjadi.
2. Bagaimana implementasi Text mining dan sentiment analisis data menggunakan metode *Naïve Bayes Classifier*

1. Bagaimana nilai akurasi ketetapan sentiment menggunakan metode *Naïve Bayes Classifier*

1.1 Batasan Masalah

Batasan masalah pada penelitian ini dapat ditentukan berdasarkan rumusan masalah yang telah dijelaskan sebelumnya yaitu sebagai berikut :

1. Dokumen teks uji dalam penelitian ini berupa tweet dari Twitter.
2. Tweet yang digunakan yaitu kalimat *tweet* dalam Bahasa Indonesia.
3. Metode yang digunakan pada *Sentiment Analisis* yaitu metode *Naïve Bayes Classifier*.
4. Penelitian tidak memperhatikan latar belakang dari pemilik akun twitter.
5. Data yang digunakan berupa data tweet hasil pencarian *keyword* “vaksin covid-19” dan didapatkan sebanyak 1631 *tweet*
6. Bahasa pemrograman yang digunakan pada proses *Sentiment Analisis* yaitu R programming menggunakan aplikasi RStudio

1.2 Tujuan Penelitian

Berdasarkan rumusan masalah tersebut, maka penelitian ini dibuat dengan tujuan sebagai berikut.

1. Mendapatkan hasil ketetapan klasifikasi sentiment pengguna twitter terhadap tweet vaksin covid-19
2. Mendapatkan hasil sentiment pada topik vaksin covid-19 pada twitter.
3. Mendapatkan kata-kata yang sering muncul dan memvisualisasikan kedalam *word cloud* dan *barplot* jumlah tweet positif dan negatif pada tweet vaksin covid-19

1.3 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut :

1. Menambah pengetahuan penulis mengenai sentiment analisis dan implementasi dari metode *Naïve Bayes Classifier* untuk klasifikasi sentimen pada suatu kasus.
2. Memberikan informasi kepada publik bagaimana sentiment tweet masyarakat terhadap suatu tokoh publik, produk, atau suatu permasalahan.

1.1 Sistematika Penulisan

Sistematika penulisan yang digunakan dalam laporan ini adalah sebagai berikut:

a. BAB I PENDAHULUAN

Bab ini menjelaskan tentang latar belakang, rumusan masalah, batasan masalah, tujuan, manfaat penelitian, dan sistematika penulisan.

b. BAB II TINJAUAN PUSTAKA

Bab ini menjelaskan tentang kajian-kajian literatur yang digunakan untuk menunjang penelitian tugas akhir ini yang terdiri dari *Text Mining*, *Naïve Bayes Classifier*, *Sentiment Analisis*, dan R Programming.

c. BAB III METODOLOGI PENELITIAN

Pada bab ini membahas langkah-langkah yang dilaksanakan dalam proses penelitian, yaitu identifikasi masalah, studi literatur, pengumpulan data, olah data, proses klasifikasi judul buku dengan menggunakan metode *Sentiment Analyst* dan *Naïve Bayes Classifier*, pengujian, implementasi dan kesimpulan dan saran.

d. BAB IV ANALISIS

Bab ini menjelaskan tentang implementasi yang dilakukan pada penelitian ini.

e. BAB V PENUTUP

Bab ini berisi tentang kesimpulan yang didapatkan pada penelitian ini dan saran untuk pengembangan penelitian selanjutnya.