

BAB I

PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi internet yang semakin pesat terutama media sosial di Indonesia, memberikan dampak langsung terhadap penggunaan media sosial. Media sosial seperti Facebook, Instagram, dan Twitter merupakan beberapa media sosial yang populer di masyarakat saat ini. Pengguna internet di Indonesia sesuai survei Asosiasi Penyelenggara Jasa Internet Indonesia (APJII) mencapai 171,17 juta orang pada tahun 2018[1]. Terdapat tiga alasan orang Indonesia dalam menggunakan internet. Tiga alasan tersebut untuk sumber informasi harian, mengakses sarana sosial/komunikasi, dan mengikuti perkembangan zaman[2].

Twitter merupakan salah satu media sosial yang seringkali digunakan oleh masyarakat untuk mengungkapkan sesuatu informasi, dengan menggunakan teks pendek yang disebut tweet. Informasi yang didapatkan melalui media ini beragam seperti pendapat, komentar, pertanyaan, berita, kritikan tentang topik tertentu[2]. Salah satu yang dibicarakan saat ini di twitter oleh para pengguna yaitu tentang kebijakan COVID-19 yang diterapkan Indonesia.

COVID-19 merupakan infeksi penyakit pernapasan yang disebabkan oleh virus severe acute respiratory syndrome coronavirus 2(SARS-CoV-2) yang pertama kali dideteksi di kota wuhan Desember 2019. Virus corona dapat menyebabkan gangguan sistem pernapasan, mulai dari efek samping yang ringan seperti influenza, hingga infeksi paru-paru [3]. Masyarakat memanfaatkan twitter untuk mengungkapkan pandangan mereka terhadap kebijakan COVID-19. Keberagaman tweet oleh pengguna twitter ini memiliki karakteristik yang tidak terstruktur dan banyak noise sehingga dibutuhkan teks mining.

Teks mining adalah cara untuk memisahkan pola (data dan informasi yang berguna) dari berbagai informasi tidak terstruktur yang nantinya akan mendapatkan pola, dan ekstraksi informasi potensial dari informasi teks. Masukan untuk Text Mining merupakan informasi yang tidak terstruktur, seperti PDF, word, dokumen, kutipan teks dan lain-lain sementara itu masukan untuk Data Mining adalah data yang terstruktur. Salah satu alasan untuk menggunakan teks mining adalah analisis

sentimen[4].

Analisis sentimen disebut juga *opinion analysis* bertujuan untuk menganalisa, mengolah pendapat atau kecenderungan opini terhadap suatu topik atau masalah yang terjadi. Analisis sentimen juga berfokus pada pengolahan opini yang memiliki nilai sentimen positif maupun negatif[2]. Banyaknya penggunaan bahasa yang tidak baku pada tweet menjadi alasan dilakukan analisis sentimen. Di Amerika ada sekitar 20 hingga 30 perusahaan yang memperhatikan pada layanan analisis sentimen. Karena besarnya pengaruh dan manfaat dari analisis sentimen menyebabkan penelitian mengenai analisis sentimen berkembang cepat[5]. *Preprocessing* merupakan salah satu tahapan yang penting dalam beberapa penelitian tentang analisis sentimen. Adapun tahapan *preprocessing* antara lain *case folding*, *cleansing*, *tokenizing*, *filtering* dan *stemming*.

Penelitian pada bagian analisis sentimen ini sudah banyak dilakukan salah satunya adalah yang dilakukan Fajar Ratnawati pada tahun 2018 perihal klasifikasi opini film yang informasi analisisnya diperoleh dari komentar yang disusun pada Twitter berdasar sentimen netral, sentimen positif dan sentimen negatif pada tingkat kalimat. Analisa sentimen yang dilakukan menggunakan metode *Naïve Bayes* dengan nilai akurasi yang diperoleh sebesar 90%[6]. Pada tahun 2017, Abdan Syakuro melakukan penelitian berjudul “Analisis Sentimen Masyarakat Terhadap E-Commerce Pada Media Sosial Menggunakan Metode *Naïve Bayes Classifier*”. Penelitian ini menggunakan data *tweet* dari *e-commerce* akun Tokopedia, Bukalapak, dan Lazada dengan seleksi fitur *information gain*. Tingkat akurasi yang didapat dalam penelitian ini sebesar 88,8%[8]. Metode *Naïve Bayes* dipilih dikarenakan salah satu metode pengklasifikasian teks yang memiliki karakteristik sederhana, mudah diaplikasikan, dan memiliki akurasi yang baik.

Berdasarkan uraian latar belakang diatas, penelitian ini mengimplementasikan metode *Naïve Bayes* untuk mengklasifikasikan teks. Teks yang dijadikan data didapatkan dari twitter dengan menggunakan keyword kebijakan covid. Diharapkan dengan adanya penelitian ini dapat mengetahui bagaimana respon masyarakat terhadap kebijakan Covid-19 yang diterapkan pada saat ini.

1.2. Perumusan Masalah

1.2.1. Permasalahan (Research Problem)

Mengetahui tingkat akurasi penerapan metode *Naïve Bayes* dalam mengklasifikasi tweet tentang kebijakan covid-19 yang diterapkan di Indonesia.

1.2.2. Metode Usulan (Proposed Method)

Digunakan metode *Naïve Bayes* untuk mengukur nilai akurasi klasifikasi respon masyarakat terhadap kebijakan *Covid-19* yang diterapkan di Indonesia.

1.2.3. Pertanyaan Penelitian (Research Question)

Berapa nilai akurasi yang didapatkan dari penerapan metode *Naïve Bayes* untuk mengklasifikasi tweet masyarakat terhadap kebijakan covid-19 ?

1.3. Tujuan dan Manfaat Penelitian

1.3.1. Tujuan Penelitian

1. Mengetahui tingkat akurasi yang diperoleh dari penerapan metode *Naïve Bayes*.
2. Menerapkan metode *Naïve Bayes* untuk mengetahui respon masyarakat terhadap kebijakan *Covid-19* di Indonesia.

1.3.2. Manfaat Penelitian

Diharapkan dapat memberikan manfaat atau kontribusi dalam pengembangan ilmu text mining khususnya yang berkaitan langsung dengan metode *Naïve Bayes*.

1.4. Batasan Masalah

Adapun batasan masalah pada penelitian ini adalah:

1. Data yang digunakan berasal dari twitter dengan cara crawling dan manual search.
2. *Tweet* yang akan digunakan merupakan tweet yang menggunakan bahasa Indonesia saja.
3. Data *tweet* diambil mulai bulan September – Desember 2020.
4. Data yang akan digunakan menggunakan kata kunci *kebijakan covid*.
5. Klasifikasi hanya menggunakan dua kategori, positif dan negatif untuk

untuk setiap kelasnya 250 dengan total data 500.

6. Data *tweet* divalidasi dengan 1 orang ahli bahasa.
7. Implementasi *filtering* dan *stemming* menggunakan library Sastrawi
8. Implementasi sistem hanya digunakan untuk research.

1.5. Sistematika Proposal

Sistematika yang digunakan dalam penulisan laporan skripsi adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini akan menjelaskan dasar-dasar penulisan tugas akhir, yaitu tentang Latar Belakang, Rumusan Masalah, tujuan dan manfaat, Batasan Masalah, dan sistematika penulisan.

BAB II KAJIAN PUSTAKA

Pada bab ini akan menjelaskan mengenai landasan teori dan penelitian sebelumnya. Landasan teori berisi tentang konsep-konsep dan teori-teori yang mendukung dalam penyusunan penelitian ini.

BAB III METODE USULAN

Pada bab ini akan menjelaskan tentang metode yang menjadi usulan (algoritma), arsitektur sistem, tahapan penelitian dan skenario pengujian sistem.