

BAB I

PENDAHULUAN

1.1 Latar Belakang

Di zaman modern saat ini, Indonesia menjadi salah satu negara dengan pengakses internet tertinggi di dunia. Menurut data dari *website We are Social*, dari Januari 2019 sampai Januari 2020, penetrasi pengguna internet di dunia telah meningkat sebanyak 9,2% dan pengguna sosial media meningkat sebesar 9,8% sedangkan di Indonesia, pengakses internet telah meningkat sebanyak 17% dalam kurun waktu tersebut. Hal ini berarti bahwa ada sekitar 25,3 juta pengguna internet baru selama setahun. Total pengguna internet di Indonesia pada Januari 2020 sebesar 64% dari total penduduk yaitu sekitar 174 juta orang.

Durasi rata-rata dalam sehari yang dikonsumsi oleh 64% dari penduduk Indonesia yang mengakses internet yaitu 7 jam 59 menit. Dari waktu tersebut, 4 jam 46 menit diakses menggunakan *mobile* dan waktu yang digunakan untuk sosial media yaitu 3 jam 26 menit, membuat Indonesia berada pada peringkat ke 5 di dunia dalam waktu penggunaan sosial media.

Twitter merupakan salah satu platform sosial media yang berada di peringkat 5 dalam penggunaannya di Indonesia. 80% dari pengguna mengakses *twitter* menggunakan perangkat *mobile*. Didirikan pada tahun 2006, pengguna *twitter* terus meningkat setiap tahunnya. Pada oktober 2020, Ada total 1,3 miliar akun *twitter* serta *twitter* mempunyai pengguna aktif setiap bulan yang telah mencapai 340 juta. *Twitter* menjadi salah satu platform dalam menyampaikan berbagai opini masyarakat.

Salah satu *query* yang sempat trending di *twitter* yaitu #Covid-19. Wabah yang muncul pertama kali di Wuhan, China. Wabah ini telah menyebar hampir ke seluruh bagian dunia sampai 30 Desember 2020, *World Health Organization* (WHO) telah mencatat 80,773,033 total *case* secara global dan sebanyak 1,783,619 jiwa telah meninggal dunia terkena Covid-19.

Menurut Komite Penanganan Covid-19 dan Pemulihan Ekonomi Nasional, jumlah kasus Covid-19 di Indonesia telah mencapai 818,386 jiwa dengan total kematian yang telah meninggal sebanyak 23,947 jiwa. Ada 2 daerah yang memiliki kasus tertinggi yaitu DKI Jakarta dan Surabaya. Menurut *Wiku*, penyebaran ini terjadi cukup cepat karena mobilitas sangat tinggi sehingga mengakibatkan penyebaran Covid-19 lebih cepat sehingga mengetahui bagaimana analisis sentiment pada 2 daerah yang memiliki kasus positif Covid-19 tertinggi.

Analisis sentimen merupakan salah satu cara dalam mengetahui opini dimasyarakat dengan cara memanfaatkan komentar masyarakat. Dalam proses klasifikasi pada penelitian ini, digunakan metode klasifikasi *Naïve Bayes*. Metode *Naïve Bayes* dipilih karena memiliki tingkat akurasi yang tinggi dibandingkan dengan algoritma lainya.(Nugraha, 2020)

1.2 Rumusan Masalah

Berdasarkan latar belakang yang ada, maka rumusan masalah yaitu bagaimana tingkat akurasi dalam klasifikasi sentimen menggunakan fitur *geolocation* dari data *twitter* yang mengandung opini masyarakat tentang Covid-19 dengan metode *Naïve Bayes* ?

1.3 Ruang Lingkup

Ruang lingkup dalam penelitian ini sebagai berikut :

1. Data yang digunakan berasal dari media sosial *twitter* dengan tweet yang berisi #Covid-19.
2. Data menggunakan bahasa Indonesia.
3. Jumlah data *tweet* sebanyak 1000 *tweet* per wilayah yang diambil berdasarkan 40 kilometer dari titik *geolocation* yang di input. Pada 12 Desember 2020.
4. Ada 2 wilayah yaitu DKI Jakarta dan Surabaya.
5. Metode klasifikasi yang digunakan yaitu *Naïve Bayes*.

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah melihat akurasi dari klasifikasi sentimen positif dan sentimen negatif menggunakan metode *Naïve Bayes* pada daerah yang sudah ditentukan.

1.5 Manfaat Penelitian

Manfaat dari penelitian ini melihat kinerja algoritma *naïve bayes* dan rujukan terhadap sentimen analisis tentang sentimen positif dan sentimen negatif pada *twitter* di 2 wilayah yang sudah ditentukan.

1.6 Luaran yang Diharapkan

Luaran yang diharapkan yaitu berupa hasil akurasi dari penggunaan metode *Naïve Bayes* informasi yang membedakan sentimen positif dan sentimen negatif terhadap pandangan masyarakat mengenai situasi pada saat Covid-19 di daerah yang terdampak.

1.7 Sistematika Penulisan

Dalam sistematika penulisan proposal ini, disusun berdasarkan aturan penulisan yang terdiri dari beberapa bagian, yaitu :

BAB 1 PENDAHULUAN

Pada Bab 1 Pendahuluan, menjelaskan latar belakang mengenai topik permasalahan yang diangkat, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, luaran yang diharapkan, serta sistematika penulisan dalam menyusun proposal ini.

BAB 2 LANDASAN TEORI

Pada Bab 2 Landasan Teori, menjelaskan teori- teori mendasar yang berkaitan dengan penelitian ini.

BAB 3 METODOLOGI PENELITIAN

Pada Bab 3 Metodologi Penelitian, menjelaskan metode yang digunakan dalam menyelesaikan masalah sehingga penelitian ini dapat mencapai tujuan.

BAB 4 HASIL DAN PEMBAHASAN

Pada Bab 4 Hasil dan Pembahasan, menjelaskan bagaimana tahapan proses yang dilakukan dalam penelitian dari masalah yang terkait sehingga mencapai hasil dan tujuan sesuai dengan yang diteliti.

BAB 5 PENUTUP

Pada Bab 5 Penutup, menjelaskan mengenai kesimpulan dari hasil penelitian yang diteliti, serta saran sebagai sarana pemecahan masalah untuk penelitian berikutnya.

DAFTAR PUSTAKA**RIWAYAT HIDUP****LAMPIRAN**