

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Di era teknologi ini, media sosial merupakan salah satu media berkomunikasi untuk berekspresi, berpendapat, dan mencari berbagai informasi. Salah satu media sosial yang saat ini populer masyarakat khususnya di Indonesia adalah twitter. Twitter adalah salah satu media dimana pengguna dapat mencari dan membahas isu-isu yang terkini. Menurut laporan dari Twitter peningkatan jumlah *Daily Active User* (DAU) pada kuartal-IV 2019 mencapai 152 juta, dimana naik hampir 21% dari jumlah 126 juta yang tercatat pada periode sebelumnya (Salsabila, 2020).

Pengguna Twitter saat ini hanya dapat mengirim sebuah pesan text dengan batas maksimal 140 karakter dan dapat mengirimkan sebuah pesan gambar, pesan tersebut biasa dikenal dengan tweet (PUTRA, 2017). Tweet adalah sebuah teks status pengguna yang digunakan untuk menulis tentang aktifitas mereka (Faishol Nurhuda, 2013). Seperti isu di masa *pandemic covid-19* ini banyak masyarakat pengguna twitter ramai-ramai mengunggah pesan pendek sebagai bentuk dukungan kepada para dokter dalam menangani virus covid-19, tak sedikit juga pesan pendek bernarasi miring yang dikirim sebagai bentuk kekecewaan yang dirasakan selama masa *pandemic*. Seperti halnya *hashtag* yang pernah trending di twitter yaitu “#terserah”.

Sentimentt analysis merupakan bagian dari *opinion mining*, mengutip dari Kamus Besar Bahasa Indonesia (KBBI), *sentimentt* adalah “pendapat atau pandangan yang didasarkan pada perasaan yang berlebih-lebihan terhadap sesuatu (bertentangan dengan pertimbangan pikiran)”. *Sentimentt analysis* digunakan untuk melihat pendapat terhadap sebuah masalah atau dapat juga digunakan untuk identifikasi kecenderungan hal di pasar (Pang, et al., 2002).

Dalam penelitian ini dibutuhkan metode klasifikasi *sentiment* untuk menentukan *sentiment* dari sebuah tweet. Metode *Naive bayes* merupakan metode klasifikasi yang dapat dibagi menjadi dua, yaitu *Multinomial Naive bayes* dan *Bernoulli*. Pada penelitian ini kedua metode tersebut penulis gunakan untuk mencari nilai akurasi terbaik yang dihasilkan.

Penelitian ini penulis mencoba memanfaatkan media sosial yaitu Twitter untuk menganalisis tweet berbahasa Indonesia yang membahas sebuah isu tentang *covid-19*. Opini melalui tweet inilah yang dapat dimanfaatkan untuk menganalisis bagaimana *sentimentt* yang dimunculkan pada pengguna dalam menanggapi isu tentang *covid-19*. Pesan yang

menjadi sumber data yaitu adalah pesan mengenai isu tentang *covid-19*. Pada penelitian ini proses pengumpulan data dilakukan menggunakan sebuah teknik *Crawling* untuk memperoleh pesan-pesan tersebut dari Twitter. Tahap selanjutnya pesan yang sudah dikumpulkan akan memasuki tahap *preprocessing*. Tahap *preprocessing* ini bertujuan untuk mempersiapkan pesan agar dapat dianalisis dengan menggunakan algoritma *Naive bayes*.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, maka dapat disimpulkan perumusan masalah sebagai berikut:

1. Bagaimana cara mengumpulkan pesan twitter berbahasa Indonesia mengenai isu pandemik *covid-19* ?
2. Berapakah akurasi *learning rate* yang didapat dari hasil implementasi algoritma *Naive bayes* ?

1.3 Batasan Masalah

Adapun Batasan pada penelitian ini adalah:

1. Tweet yang digunakan hanya berbahasa Indonesia, dan mengenai isu *covid-19*.
2. Bahasa pemrograman yang digunakan adalah bahasa Python.
3. Dataset yang digunakan hasil dari *Crawling* menggunakan API twitter dengan “#dirumahaja” dan “#covid19”
4. Metode yang digunakan untuk pengolahan data adalah *cleansing, case folding, stopwords, Tokenizing*.
5. Algoritma yang digunakan *Naive bayes*.

1.4 Tujuan dan Manfaat Penelitian

Tujuan dari penelitian ini adalah sebagai berikut

1. Untuk merancang sistem identifikasi tweet negative, positif, atau netral berbahasa Indonesia dan menerapkan algoritma *Naive bayes*.
2. Untuk mengetahui kinerja algoritma *Naive bayes* dalam mengidentifikasi tweet negatif, positif, atau netral pada penelitian ini.

3. Untuk merancang dan mengimplementasikan algoritma *Naive bayes* dalam teks *mining* .

Adapun manfaat dari penelitian ini adalah sebagai berikut :

1. Dapat membantu menganalisis *sentimentt* pada twitter dengan menggunakan algoritma *Naive bayes*.
2. Dapat mengetahui akurasi klasifikasi algoritma *Naive bayes* dalam melakukan *analysis sentimentt* pada twitter.
3. Penelitian ini dapat digunakan sebagai pembanding dalam penelitian-penelitian sejenis