

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Objek yang diambil dalam penelitian ini adalah Twitter. Twitter diambil sebagai objek karena teks tweet tersebut diposting lebih singkat dan tujuannya lebih spesifik. Bahan penelitian yang digunakan dalam penelitian ini adalah tweet teks yang diposting di media sosial Twitter. Kegiatan yang dilakukan dalam pengambilan data tweet adalah dengan mengunjungi situs Twitter (Rizal Tjut Adek, 2018).

Twitter merupakan salah satu media sosial yang banyak digunakan di Indonesia. Berdasarkan data, lebih dari 22,8% netizen Indonesia menggunakan tweet pada tahun 2019. Saat ada berita hangat yang banyak diperbincangkan (trending) di Twitter dapat dijadikan sebuah acuan Sesuatu yang serius serta receh kerap kali dibahas di *Twitter* diawali dengan topik politik, agama, hiburan, skandal, sampai cerita menarik serta inspiratif (Aditya, D. R. 2022).

*Twitter* menggambarkan salah satu media sosial yang tidak kalah jauh terkenal digunakan pada masa saat ini. Pengguna *tweet* leluasa mengunggah serta mengutarakan apa saja yang tercantum pada argumennya. Unggahan pada *twitter* dapat berbentuk kenyataan, anjuran, data, serta kritik terhadap suatu. Aplikasi Mirip dengan Twitter, ia menjadi platform komunikasi yang sangat populer di kalangan netizen Indonesia. Pada tahun 2019, Indonesia menjadi negara ketiga dengan pengguna Twitter aktif terbanyak di dunia (Akhmad Muzaki, A. W. 2021).

Lebih dari 50% penduduk Indonesia menggunakan media sosial seperti Facebook, Instagram, Twitter, LINE, WhatsApp, YouTube, dll. Pesatnya perkembangan teknologi membuat banyak hal dicoba secara digital. Selain itu, munculnya pandemi Covid-19 varian omikron membuat segala aktivitas harus dilakukan dari rumah, sehingga jumlah pengguna media sosial terus tinggi.

Varian *Omicron* jenis penyakit yang menyerang pernafasan manusia, varian *Omicron* ini bisa dibilang semacam virus corona (*covid- 19*) bertepatan berjalannya waktu penyebaran varian *Omicron* terus menjadi sangat cepat serta mulai memakan korban. Pemerintah berusaha mencari solusi untuk mencegah varian omicron dengan menciptakan peraturan serta kebijakan seperti mengurangi kegiatan masyarakat di luar, memenuhi protokol kesehatan, serta memberikan vaksinasi pada masyarakat secara sedikit demi sedikit agar terhindar dari bakteri virus.

Salah satu hal yang menarik perhatian di Twitter adalah kesadaran masyarakat terhadap varian Omicron dari Covid-19. Saat ini, Omicron sedang menjadi topik hangat, dengan banyaknya peredaran opini, rumor, dan gosip yang tidak jelas, sebagai akibatnya menyebabkan perdebatan yang saling bertentangan di masyarakat mengenai varian *Omicron*. Di jejaring sosial *Twitter*, tidak ada fungsi untuk mengklasifikasikan *tweet* atau komentar dengan memasukkan emosi positif, negatif, atau netral (Harun Sujadi, S. F. 2022)

Manfaat *twitter* tidak cuma sebagai media yang berbagi informasi dengan berbagai macam postingan di *twitter*, Twitter juga sering digunakan untuk komunikasi antar pengguna dan mengkomunikasikan perasaan dan pendapat mereka mengenai topik atau topik yang sedang hangat diperbincangkan. Untuk memperoleh nilai opini dari sebuah pendapat kita perlu menggunakan sentimen analisis terhadap data pendapat pengguna *twitter* yakni salah satu cara yang menggunakan metode data mining. Dalam data mining dan visualisasi data merupakan hal yang biasa kita bicarakan. Saat ini, meningkatnya komunikasi di media sosial mempengaruhi pembentukan opini publik dan tumbuhnya peran berbagai jenis berita di media sosial.

Oleh karena itu dapat digunakan untuk menganalisis ada tidaknya pendapat masyarakat bersifat positif, negatif, atau netral. Ada juga metode untuk menganalisis opini publik disebut analisis sentimen. Analisis sentimen adalah bidang penelitian yang menganalisis opini, pandangan, penilaian, sikap dan perasaan masyarakat tentang entitas seperti produk, layanan, organisasi, orang, kasus, insiden, objek, serta properti. Pada dasarnya, NLP dan penambangan teks adalah teknologi kecerdasan buatan (AI) yang memungkinkan pengguna mengubah konten utama dokumen teks sebagai data kuantitatif pada saat cepat. Data kuantitatif kemudian dapat digunakan kembali dan diproses sesuai harapan pengguna. Untuk mempertimbangkan besar kecilnya nilai opini dari sebuah opini, kami menggunakan Algoritma Naïve Bayes. Algoritma Naive Bayes mengacu pada algoritma pemrosesan data yang menggambarkan pemrosesan bahasa alami seperti program *Natural language processing* ini untuk melacak keadaan publik tentang segala hal yang dituliskan didalam twitter, baik tentang produk maupun opini tertentu. Selama analisis sentimen, opini yang dianalisis kemudian dikategorikan sedemikian rupa sehingga orang dapat melihat dengan jelas apa kelebihan dan kekurangan dari varian *Omicron*.

Analisis sentiment bermanfaat untuk menganalisis opini dari pengguna yang memberikan opini perihal berbagai informasi dan berita yang sedang trending. Kabar terkait versi Omicron saat ini menjadi salah satu berita yang paling banyak diperbincangkan di

kalangan pengguna Twitter Indonesia. Oleh karena itu, perlu dicoba analisis sentimen untuk mengklasifikasikan *tweet* yang terkait dengan varian *omicron*.

Sebuah penelitian (Pang, Lee, Rd, Jose, 2002) Perbandingan algoritma klasifikasi Naive Bayesian, entropi maksimum, dan mesin vektor dukungan yang berkinerja terbaik, atau mesin vektor dukungan (SVM). Kelebihan Support Vector Machine (SVM) adalah dapat diterapkan pada data berskala besar, namun kelemahannya adalah sulit diterapkan pada data berskala besar (Nugroho, Witarto, & Handoko, 2003).

Investigasi berikut telah dilakukan (Ety sutanty dan Meilani B Siregar, 2019) dengan menggunakan judul “Analisis sentimen opini publik mengenai *Covid-19* melalui platform *twitter* dengan metode *Naïve Bayes Classifier*, pada hasil penelitian ini menunjukkan bahwa data prediksi sangat mendekati dengan data nyata, recall sebesar 99,59% data pada kelas positif memprediksi hampir dengan kesalahan dan akurasi di atas 100% menunjukkan bahwa data untuk peristiwa positif yang diprediksi dengan benar mewakili dari total prediksi peristiwa (sutanty, 2019). Penulis pada studi ini akan melakukan analisis sentimen pada pengguna twitter terhadap varian *Omicron* menggunakan kombinasi metode *Term frequency-Invers Document Frequency* (TF- IDF). dengan input berupa data *tweet* dalam bahasa indonesia dan output berupa tiga jenis sentimen yakni positif, negatif dan netral. dalam situasi diatas, kasus yang mungkin terjadi adalah prosedur buat menentukan hasil komentar publik tentang varian *Omicron* dari pengguna *Twitter* sesuai kata kunci *Omicron*. Pada penelitian ini mencoba menentukan persentase opini positif, negatif dan netral dengan proses klasifikasi menggunakan algoritma *Naïve Bayes Classifier* (NBC). *Naïve Bayes Classifier* (NBC) memiliki kelebihan yaitu, cepat dalam proses perhitungan, algoritma yang sederhana dan akurasi yang tinggi. hasil dari penelitian ini dapat dijadikan sebagai salah satu pertimbangan untuk meningkatkan kesadaran masyarakat dan mengembangkan kebijakan pemerintah untuk mengalahkan varian *Omicron* di masa mendatang.

dalam klasifikasi text dapat menggunakan metode ANN, SVM, NBC. pada (Ahmad Fathan Hidayatullah, 2015) telah menunjukkan bahwa metode *Naïve Bayes* memiliki akurasi yang didapatkan dengan nilai yang begitu besar 91,96%. dan (Fajar Ratnawati, 2018) yang dimana pengguna metode *Naïve Bayes Classifier* memiliki pengujian pengelompokan dataset yang digunakan menerapkan metode 5-folds cross validation yang menghasilkan rata-rata akurasi sebesar 82,4%, precision 83,8%, recall 82,4%, dan f-measure 82,2%. Akurasi yang mempunyai nilai tinggi menandakan bahwa sistem tersebut melakukan pencapaian klasifikasi dengan sangat baik. Sedangkan dalam (Ety sutanty, 2019) menunjukkan bahwa data prediksi sangat mendekati dengan data nyata, recall sebesar 99,59% data pada kelas positif

memprediksi hamper dengan kesalahan dan akurasi di atas 100% menunjukkan bahwa data untuk peristiwa positif yang diprediksi dengan benar mewakili dari total prediksi peristiwa. Melihat beberapa penelitian sebelumnya, penulis berkeinginan melakukan klasifikasi terhadap komentar netizen tentang *Covid-19* varian *Omicron* di Twitter dengan menggunakan metode *Naïve Bayes Classifier* (NBC). algoritma *Naive Bayes Classifier* yang berfungsi untuk mengklasifikasikan setiap opini menjadi suatu sentimen positif, sentimen negatif maupun sentimen netral. dengan dilakukannya penelitian ini penulis bermaksud agar bisa mengetahui bagaimana tanggapan masyarakat mengenai *Omicron*, memberikan ilmu dan menambah informasi khususnya mengenai cara pemutusan dan pencegahan dari Virus tersebut.

Berdasarkan dari masalah di atas penulis mencoba untuk mengangkat sebuah topik yang berjudul “**Analisis Sentimen Komentar Netizen Terhadap Covid-19 Varian Omicron dari Twitter Menggunakan Algoritma Naïve Bayes**”

## **1.2 Rumusan Masalah**

Berlandaskan latar belakang diatas hingga ada rumusan permasalahan, rumusan permasalahan yang hendak dibahas merupakan:

1. Bagaimana melakukan penambangan informasi (retrieval information) dari *Twitter* yang berkenaan dengan *Covid-19* dengan varian *Omicron*?
2. Bagaimana mengimplimentasikan algoritma *Naïve Bayes* dalam melakukan klasifikasi sentiment di *Twitter* mengenai *Covid-19* dengan varian *Omicron*?

## **1.3 Batasan Masalah**

Supaya ulasan lebih terencana bersumber pada formulasi permasalahan serta cocok dengan batas keahlian penulis, hingga batas permasalahan dalam studi ini dipecah jadi 5 bagian, ialah:

1. Informasi pendapat memakai bahasa Indonesia.
2. Informasi sentimen dikumpulkan dari komentar netizen terhadap varian *omicron* di aplikasi *twitter*.
3. Tidak menganalisis kata yang tidak dibutuhkan seperti hashtag (#), username (@) serta url (<http://namasitus.com>). serta emoticon(:v)
4. Algoritma yang digunakan dengan menggunakan algoritma *Naïve bayes*.
5. Penelitian ini hanya berada pada ruang lingkup opini pada *twitter* yang berhubungan dengan varian *omicron*.
6. Data yang digunakan 1000 – 2000 *Tweet*.

#### **1.4 Tujuan Penelitian**

Tujuan yang ingin dicapai dalam penataan laporan tugas akhir ini yaitu:

1. Melaksanakan klasifikasi terhadap sentimen positif, netral, ataupun negatif tentang opini masyarakat mengenai *Omicron* pada twitter.
2. Melaksanakan pengukuran akurasi analisis sentimen terhadap pendapat pengguna *twitter* dengan memakai cara klasifikasi *Naïve Bayes*.

#### **1.5 Manfaat Penelitian**

Penelitian ini diharapkan dapat membawa manfaat. berikut ini manfaat yang diharapkan dari hasil penelitian ini, yaitu:

1. Mengetahui tanggapan sentimen masyarakat terhadap pandemi yang terus menerus berlanjut.
2. Memberikan ilmu dan menambah informasi khususnya mengenai cara mengatasi dan pencegahan dari Varian *Omicron*.
3. Memberikan informasi sentimen positif, negatif atau netral terhadap munculnya virus varian baru ini.