

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi (TI) sudah banyak mengubah aspek kehidupan manusia. Konsumsi internet selama aktivitas digital ditunjukkan oleh lalu lintas Global Internet Protocol (IP), yang berfungsi sebagai proxy untuk transfer data, meningkat dari sekitar 100 GB per hari pada tahun 1992 menjadi lebih dari 45.000 GB per detik pada tahun 2017. (Abdillah, Mukti, Puspita, & Suhartini, 2021). Dunia teknologi yang berkembang di Indonesia saat ini pula telah memudahkan pengguna dalam mengakses dan menerima informasi baik dari media sosial, media cetak maupun media *digital*. Dengan berbagai informasi, masyarakat dapat dengan mudah mengeluarkan pendapatnya dan menyebarkan informasi yang diperoleh melalui media sosial yang dimiliki. Media sosial sendiri termasuk platform media yang digunakan untuk melakukan komunikasi satu sama lain yang dilakukan secara online sehingga memungkinkan seseorang dapat berinteraksi secara jarak jauh. Salah satu contoh media sosial di Indonesia yang paling banyak digunakan sampai saat ini adalah Twitter. Berdasarkan data dari wearesocial.com dan Hootsuite, perusahaan analisis media sosial yang berasal dari Kanada bahwa penggunaan media sosial aktif di Indonesia pada tahun 2020 sudah mencapai 59% atau 160 juta jiwa dari jumlah total penduduk Indonesia yang mencapai 272,1 juta jiwa dan pengguna internet sudah mencapai 64% atau 175,5 juta jiwa dan Twitter berada pada urutan kelima sebagai media sosial yang banyak digunakan di Indonesia. Dengan banyak media sosial Twitter saat ini *pengguna dapat mengeluarkan pendapatnya terhadap isu yang sedang hangat dibicarakan saat ini*. (Wulandari, Saedudin, & Andreswar, 2021).

Analisis sentimen merupakan ilmu yang berguna untuk menganalisis pendapat seseorang, sentimen seseorang, evaluasi seseorang, sikap seseorang

dan emosi seseorang ke dalam bahasa tertulis. Tugas dasar dalam analisis sentimen adalah mengelompokkan polaritas dari teks yang ada dalam dokumen, kalimat, atau pendapat yang bersumber dari media *online*. Polaritas mempunyai arti apakah teks yang ada dalam dokumen, kalimat, atau pendapat memiliki aspek positif, negatif atau netral. Opini negatif disini mengandung arti bahwa kata atau opini yang dapat menimbulkan permusuhan, penghinaan, perdebatan dan perselisihan di dunia maya. Sedangkan opini positif yaitu kata atau opini yang sifatnya positif dan tidak menimbulkan permusuhan, penghinaan, perdebatan, dan perselisihan di dunia maya. Sedangkan netral adalah kata yang tidak mengandung negatif maupun positif (Semobodo. 2015).

Metode untuk analisis sentimen peneliti menggunakan algoritma *Naïve Bayes*. *Naïve Bayes* merupakan teknik pembelajaran mesin yang berbasis *probabilistic* (Routray, P., Swain, C. K. & Mishra, S. P., 2013). *Naïve Bayes* adalah metode sederhana tetapi memiliki akurasi serta performansi yang tinggi dalam pengklasifikasian teks (Routray, P., Swain, C. K. & Mishra, S. P., 2013). *Teorema Bayes* yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelum (C.T.I, et al., 2017) mnya sehingga dikenal sebagai *Teorema Bayes* (Bustami, 2014). Torema tersebut dikombinasikan dengan *naïve* dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naïve Bayes* diasumsikan bahwa ada atau tidak ada hubungannya dengan ciri khas lainnya.

Baru – baru ini dikejutkan lagi dengan penemuan varian baru dari virus Covid-19 yaitu B.1.1.529 atau Omicron, yang pertama kali ditemukan di Afrika Selatan pada tanggal 24 November 2021. Varian Omicron merupakan sebuah varian terbaru yang saat ini menjadi varian yang secara mayoritas ditemukan di berbagai negara. Dengan demikian, *World Health Organization* (WHO) telah mengklasifikasikannya sebagai "varian yang diwaspadai" (*variant of concern*) (<https://promkes.kemkes.go.id/>, 2022).

Berdasarkan penemuan baru mengenai Covid-19 di dunia saat ini peneliti ingin melakukan analisis mengenai pendapat masyarakat indonesia tentang adanya virus Covid-19 yaitu B.1.1.529 atau Omicron menggunakan pendekatan klasifikasi data *text* akan dilakukan dengan menggunakan perangkat lunak RapidMiner. RapidMiner merupakan *software/perangkat*

lunak untuk pengolahan data. Dengan menggunakan prinsip dan algoritma data mining, RapidMiner mengekstrak pola-pola dari data set yang besar dengan mengkombinasikan metode statistika, kecerdasan buatan dan database. RapidMiner memudahkan penggunaanya dalam melakukan perhitungan data yang sangat banyak dengan menggunakan operator-operator (CTI dkk., 2017).

Berdasarkan uraian diatas maka peneliti akan melakukan penelitian mengenai Analisis Sentimen masyarakat terhadap Virus Covid-19 varian baru Omicron pada media sosial *Twitter* untuk diambil data sentimen, selanjutnya penelitian ini mencoba melakukan analisa sentimen menggunakan *Naïve Bayes* untuk mengetahui bagaimana polaritas pengelompokan tersebut apakah opini yang diberikan positif, netral atau negatif. Maka judul yang akan diambil yaitu **"Analisis Sentimen Terhadap Opini Masyarakat Tentang Varian Covid-19 Omicron Menggunakan NBC dan RapidMiner"**.

1.2. Rumusan Masalah

Berdasarkan latar belakang permasalahan yang telah diuraikan di atas, maka peneliti merumuskan masalah yaitu bagaimana menganalisis sentimen terhadap opini masyarakat tentang varian Covid-19 Omicron menggunakan NBC dan Rapidminer ?

1.3. Batasan Masalah

Agar permasalahan tidak meluas dan lebih terarah, maka dalam penelitian ini penulis membatasi permasalahan yaitu :

1. Data sentimen yang diambil dari media sosial seperti *Twitter*.
2. Algoritma untuk analisis sentimen menggunakan *Naïve Bayes Classification*.
3. *Editor* untuk menghasilkan keputusan dari algoritma *Naïve Bayes Classification* menggunakan Rapidminer.

1.4. Tujuan dan Manfaat

1.4.1. Tujuan Penelitian

Tujuan dari penelitian ini adalah menganalisis sentimen terhadap opini masyarakat tentang varian Covid-19 Omicron menggunakan NBC dan Rapidminer pada media sosial *Twitter*.

1.4.2. Manfaat Penelitian

Pada penelitian ini diharapkan dapat memberikan manfaat seperti berikut :

1. Memberikan informasi mengenai respon masyarakat tentang virus terbaru yaitu Varian Covid-19 Omicron.
2. Dapat dijadikan alat dalam melakukan pemantauan serta tanggapan masyarakat terhadap fenomena dalam dunia kesehatan khususnya perkembangan virus Covid-19 di media sosial.
3. Hasil dari penelitian ini dapat memberikan pembelajaran mengenai metode *Naïve Bayes Classification* dalam melakukan analisis sentiment pada media social *twitter* mengenai respon masyarakat tentang virus terbaru yaitu Varian Covid-19 Omicron.

1.5 Metodologi Penelitian

1.5.1 Waktu Penelitian

Waktu penelitian dilakukan selama 6 bulan yaitu mulai dari bulan Februari 2022 sampai dengan Agustus 2022.

1.5.2 Metode Penelitian

Metode sampling yang digunakan dalam penelitian ini menggunakan metode Non Random Sampling atau Sampel Tidak Acak. Non Random Sampling adalah metode dimana setiap elemen atau anggota populasi tidak mempunyai peluang yang sama untuk dijadikan sampel dikarenakan

pengambilan sampel mempunyai kriteria tertentu. Non Random Sampling memiliki beberapa teknik, salah satunya adalah sampling kuota (*quota sampling*). *Quota Sampling* merupakan teknik sampling yang menentukan jumlah sampel dari populasi yang memiliki ciri tertentu hingga jumlah kuota yang diinginkan tercapai. Pengambilan sampel berdasarkan pertimbangan peneliti saja, besar dan kriteria sampel telah ditentukan (Sari, F. V., & Wibowo, A. 2019).

1.5.3 Metode Pengumpulan Data

Untuk memperoleh gambaran mengenai data yang dibutuhkan dalam pembuatan laporan ini metode yang digunakan yaitu Langkah-langkah yang dilakukan untuk memperoleh data dan informasi yang berhubungan dengan penelitian ini adalah :

1. Pengamatan (*Observasi*)

Metode ini dilakukan dengan cara mengamati langsung mengenai perkembangan media sosial tentang virus terbaru yaitu Varian Covid-19 Omicron.

2. Studi Pustaka

Mengumpulkan data dengan cara mencari dan mempelajari data-data dari buku-buku ataupun dari referensi lain yang berhubungan dengan penulisan laporan penelitian proposal. Buku yang digunakan penulis sebagai referensi, adapun metode yang digunakan penulis dalam merancang dan mengembangkan dapat dilihat pada daftar pustaka.

3. *Crawling*

Metode ini dilakukan dengan cara membuat *script* untuk mengambil data komentar di *tweets* mengenai covid19, omicron kemudian nantinya data inilah yang akan di kelola untuk menentukan sentiment masyarakat yang tergolong dalam sentimen Positif, Netral dan Negatif.

1.5.4 Sumber Data

Menurut (Sugiyono, 2018), mengatakan sumber data terbagi menjadi:

1. Sumber data Primer

Merupakan sumber data yang diperoleh secara langsung dari tempat atau objek dan memberikan data langsung kepada pengumpulan data, seperti dokumen-dokumen instant yang berupa sejarah perkembangan perusahaan, struktur organisasi dan lainnya yang berhubungan dengan penelitian.

2. Sumber data Sekunder

Merupakan sumber data yang tidak langsung memberikan data kepada pengumpulan data, misalnya melalui orang lain.

1.6. Sistematika Penulisan

BAB I PENDAHULUAN

Pada bab ini berisi latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Dalam bab ini akan dijelaskan landasan teori yang berhubungan dengan masalah yang diteliti, seperti pengertian analisis, *corona virus disease 19*, *text mining*, dan perangkat *tools* yang digunakan untuk menasikkan analisa.

BAB III METODELOGI PENELITIAN

Bab ini berisikan mengenai analisis sistem yang meliputi proses analisis data sentiment seta penjelasan mengenai metode *Naïve Bayes Classification* pada dokumen yang akan di analisis.

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini menguraikan mengenai hasil dan pembahasan dari metodologi penelitian yang telah dilakukan.

BAB V KESIMPULAN DAN SARAN

Dalam bab ini menjelaskan secara garis besar mengenai kesimpulan dan saran dari hasil penelitian.