

# BAB 1

## PENDAHULUAN

Bab ini mencakup beberapa hal yang menjadi dasar penelitian seperti: latar belakang, identifikasi masalah, perumusan masalah, batasan masalah, tujuan, dan manfaat dari penelitian. Dari uraian tersebut, diharapkan gambaran umum mengenai permasalahan dan pemecahan yang diambil pada penelitian ini dapat dipahami.

### 1.1 Latar Belakang

Tahun 2020 merupakan tahun yang mengkhawatirkan seluruh negara, tanpa terkecuali negara Indonesia. Hal itu disebabkan munculnya wabah virus Corona yang bermula dari Kota Wuhan China dan menyebar ke seluruh penjuru dunia (Burhanuddin & Abdi 2020; Yunus & Rezki 2020). Wabah ini kemudian disebut sebagai COVID-19 serta kemudian dinyatakan sebagai pandemi oleh badan kesehatan dunia (WHO). Selanjutnya, dengan teridentifikasinya penularan virus di Indonesia, pemerintah segeramemberikan reaksi cepat berupa sosialisasi pencegahan meluasnya penularan. Dewasa ini, pemerintah Indonesia terus melakukan upaya-upaya guna meminimalisir jumlah orang yang terinfeksi COVID-19.

Berbagai cara dilakukan untuk mengantisipasi bertambahnya jumlah penderita virus Corona di Indonesia, di antaranya dengan memberikan kebijakan membatasi aktivitas keluar rumah, kegiatan belajar dilakukan secara daring, bekerja dari rumah (*work from home*), bahkan kegiatan beribadah pun dihimbau tidak dilakukan beramai-ramai untuk mencegah kerumunan. Hal ini sudah menjadi kebijakan pemerintah berdasarkan pertimbangan-pertimbangan yang sudah dianalisa dengan maksimal tentunya. Hampir seluruh kegiatan di luar rumah dibatasi untuk mencegah kerumunan. Dengan mencegah terjadinya kerumunan dapat membantu menghindari penyebaran virus Corona ke suatu wilayah, sehingga masyarakat yang berada di suatu wilayah tersebut diharapkan dapat terhindar dari wabah yang cepat menyebar tersebut. Kebijakan ini hanya dapat dilakukan oleh pemerintah, dengan terlebih dahulu melakukan pemeriksaan secara ketat sebelumnya ke beberapa wilayah dan mempertimbangkan konsekuensi secara matang, baik dari segi ekonomi maupun sosial.

Penggunaan media sosial di masyarakat telah menyebar luas dan sangat pesat dalam berbagai macam lapisan masyarakat, tidak hanya sebagai media untuk bersosialisasi, dan berkomunikasi namun juga untuk menyampaikan aspirasi dan mempresentasikan tentang apa saja yang terjadi dan dirasakan masyarakat (Rosenthal et al. 2017). Banyak masyarakat yang

mengutarakan pendapatnya melalui media sosial *Twitter* dalam menanggapi suatu kejadian akan hal-hal yang terjadi di lingkungannya (Kaur & Sharma 2020). Berbagai opini yang diutarakan masyarakat dapat menjadi sebuah acuan untuk mengetahui sentimen masyarakat terhadap hal-hal yang terjadi di lingkungan atau tempat tinggal masyarakat tersebut. Jumlah opini yang sangat banyak dan terus meningkat menjadi permasalahan untuk mengetahui sentimen masyarakat sehingga diperlukan analisis terhadap opini masyarakat secara otomatis. Salah satu solusi yang dapat digunakan dalam mengatasi permasalahan tersebut adalah *text mining*.

*Text mining* dapat didefinisikan sebagai proses pengambilan informasi yang bersumber dari beberapa dokumen. Dalam *text mining* terdapat beberapa tujuan penggunaan yang khas diantaranya adalah *text categorization*, *text clustering*, dan analisis sentimen. Dalam pengklasifikasian opini yang telah dijelaskan fungsi *text mining* yang digunakan adalah analisis sentimen.

Analisis sentimen atau *opinion mining* merupakan proses memahami, mengekstrak dan mengolah data eksternal secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini (Priyono et al. 2016). Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang apakah cenderung berpandangan positif atau negatif (Kaur & Sharma 2020; Rasool et al. 2019; Elbagir 2019).

Terdapat beberapa penelitian sebelumnya yang membahas analisis sentimen dokumen teks. Penelitian (Wardhana et al. 2016) menggunakan analisis sentimen untuk menilai kepuasan seseorang melalui opini berdasarkan dokumen teks. Dokumen dipecah berdasarkan teks yang penyusun, kemudian masing-masing teks dinilai untuk menentukan opini memiliki sentimen positif atau negatif. Penelitian (Rozi et al. 2018) menggunakan sentimen analisis untuk mengetahui tingkat kepuasan masyarakat terhadap produk dan layanan dari SAMSAT Kota Malang. Dokumen yang digunakan adalah *tweet* masyarakat yang membahas produk dan layanan tersebut. Pada penelitian tersebut, Naive Bayes menunjukkan hasil yang relatif tinggi. Penelitian (Pandey et al. 2017) memanfaatkan metode Hybrid Cuckoo Search untuk menganalisis sentimen dengan menggunakan data berupa opini pada sosial media *Twitter*. Penelitian lainnya menggunakan Naive Bayes dan K-Means untuk menganalisis sentimen pada *Twitter* dengan topik pemilihan presiden 2019 (Kurniawan & Susanto 2019). Penelitian (Rachmat et al. 2017) melakukan sentimen analisis terhadap kasus pemilihan calon legislatif 2014 melalui data Facebook Page dengan memanfaatkan metode Naive Bayes. Metode lain

yang dikembangkan untuk sentimen analisis adalah Textblob (Kaur & Sharma 2020) dan Backpropagation (Habibi & Setyohadi 2016).

Berdasarkan latar belakang tersebut, pada penelitian ini akan diajukan analisis sentimen mengenai dampak Corona melalui dokumen dalam media sosial *Twitter*. Metode yang digunakan dalam penelitian ini adalah Naive Bayes dengan pertimbangan tingginya *accuracy* hasil klasifikasi pada penelitian sebelumnya. Selain itu, Naive Bayes merupakan metode pengklasifikasian yang sering digunakan dalam sentimen analisis karena sederhana dan mudah dalam melakukan pengklasifikasian dokumen. Metode ini diharapkan mampu menghasilkan nilai *accuracy* yang tinggi untuk klasifikasi dokumen sosial media pada kasus dampak virus Corona.

## 1.2 Identifikasi Masalah

Berdasarkan latar belakang penelitian, masalah yang teridentifikasi dan akan dijadikan bahan penelitian sebagai berikut:

Opini mengenai virus Corona berkembang pesat di masyarakat melalui media sosial seiring meningkatnya jumlah masyarakat yang terkena wabah Corona.

Identifikasi sentimen sebuah opini mengenai virus Corona sulit dilakukan mengingat jumlah opini yang sangat banyak.

## 1.3 Perumusan Masalah

Dengan melihat uraian latar belakang di atas, maka dapat dirumuskan masalah sebagai berikut:

1. Bagaimana memanfaatkan Naive Bayes untuk melakukan klasifikasi *tweet* mengenai dampak virus Corona?
2. Bagaimana cara mengetahui sentimen masyarakat melalui dokumen *Twitter* secara presisi dan akurat?

## 1.4 Batasan Masalah

Batasan masalah dibuat agar penelitian tetap berfokus pada tujuan awal dan untuk menghindari melebarnya lingkup penelitian. Beberapa batasan masalah pada penelitian ini adalah:

1. Data *tweet* mengenai dampak virus Corona, dengan *hashtag* #DampakVirusCorona, #DampakCorona, #DiRumahSaja, #LawanCorona, #Covid19, #BersatuLawanCovid19, dan #NewNormal.

2. Metode yang digunakan adalah Naive Bayes Classifier.
3. Media sosial yang digunakan adalah *Twitter*
4. *Tweet* yang digunakan hanya dalam bahasa Indonesia.

### 1.5 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah:

1. Untuk mengetahui cara memanfaatkan Naive Bayes untuk melakukan klasifikasi *tweet* mengenai dampak virus Corona.
2. Untuk mengetahui sentimen masyarakat pada *Twitter* secara presisi dan akurat.

### 1.6 Manfaat Penelitian

Adapun manfaat yang bisa diperoleh bagi peneliti dari penelitian ini adalah sebagai tugas akhir dan syarat kelulusan. Selain itu, peneliti bisa mengetahui manfaat Naive Bayes dalam melakukan klasifikasi *tweet* guna mendapatkan sentimen positif dan negatif masyarakat terhadap dampak virus Corona yang saat ini dihadapi oleh banyak negara. Bagi universitas, ilmu yang diperoleh dari hasil penelitian dapat dijadikan sebagai pertimbangan untuk mengembangkan topic ini sehingga bisa melahirkan sebuah solusi yang diperuntukkan bagi masyarakat dalam menghadapi pademi virus Corona.

