

# **BAB I**

## **PENDAHUALUAN**

### **1.1 Latar Belakang Masalah**

Wabah Covid-19 yang telah terjadi sepanjang dua tahun terakhir berdampak banyak pada tatanan hidup bermasyarakat di banyak negara. Berbagai upaya penerapan kebijakan dilakukan pemerintah-pemerintah di tiap negara sebagai bentuk upaya menekan laju angka kasus positif Covid-19, pemerintah Indonesia sendiri telah mengeluarkan berbagai macam kebijakan mulai dari pemberlakuan kebijakan PSBB (Pembatasan Sosial Bersekala Besar) yang diatur oleh PP nomor 21 tahun 2020, hingga penerapan PPKM (Pemberlakuan Pembatasan Kegiatan Masyarakat).

Kebijakan-kebijakan yang dikeluarkan oleh pemerintah dalam upaya menekan angka positif Covid-19 di Indonesia ini, mendapatkan berbagai macam respon dari masyarakat, beberapa dari mereka berpandangan positif, bahwa penerapan kebijakan-kebijakan ini harus digiatkan sebagai bentuk usaha untuk percepatan penurunan angka positif Covid-19 agar dapat kembali menjalani kehidupan normal, namun demikian tidak semua masyarakat berpandangan demikian, beberapa dari masyarakat yang merasakan dampak negatif secara langsung seperti penurunan pendapatan beberapa pedagang, pemutusan hubungan kerja, dan banyak lagi, juga mengeluhkan kebijakan yang diterapkan oleh pemerintah. Untuk melihat pola sentimen dari kedua sentimen positif dan negatif salah satu metode yang bisa digunakan adalah metode analisis sentimen.

Analisis sentiment merupakan salah satu cara untuk mengumpulkan pendapat orang banyak terhadap sesuatu seperti layanan publik, isu, kinerja pemerintahan atau hal lain berkaitan. Analisis sentiment dapat digunakan sebagai salah satu cara untuk melakukan evaluasi terhadap layanan yang telah diberikan. Analisis sentiment dapat dilakukan melalui berbagai cara salah satunya adalah dengan mengumpulkan pendapat orang banyak melalui media social (Ramadhan, Novianty, dan Setianingsih, 2017), termasuk salah satunya melalui *Twitter*, sebagai salah satu

media sosial yang masif digunakan oleh masyarakat, untuk mengemukakan opini dan pendapat mereka.

Twitter adalah layanan jejaring sosial dan mikroblog daring yang memungkinkan penggunanya untuk mengirim dan membaca pesan berbasis teks. *Twitter* sendiri didirikan pada bulan Maret 2006 oleh Jack Dorsey, dan situs jejaring sosialnya diluncurkan pada bulan Juli 2006. (*Twitter*, 2020). *Twitter* mengalami pertumbuhan yang pesat dan dengan cepat meraih popularitas di seluruh dunia. Hingga bulan Januari 2020, pengguna aktif media sosial ini mencapai 340 juta pengguna (Ramadhan, Februari 13, 2020).

Dalam rilis Statista diaktakan bahwa Indonesia merupakan negara dengan pengguna aktif *Twitter* ketiga terbesar di dunia, dalam rilis ini disebutkan juga bahwa pengguna aktif *Twitter* Indonesia per Mei 2016 mencapai 24,34 juta, tTingginya popularitas *Twitter* menyebabkan layanan ini telah dimanfaatkan untuk berbagai keperluan dalam berbagai aspek, misalnya sebagai sarana protes, kampanye politik, sarana pembelajaran, dan sebagai media komunikasi darurat. Semua pengguna dapat mengirim dan menerima *tweet* melalui situs *Twitter*, aplikasi eksternal yang kompatibel (telepon seluler), atau dengan pesan singkat (SMS) yang tersedia di negara-negara tertentu (*Twitter*, 2013).

Saat ini pendapat atau pandangan yang disampaikan melalui kicauan di *Twitter* menjadi penting untuk dipertimbangkan, sebab aktifitas, perasaan, dan pendapat masyarakat tentang suatu peristiwa, pelayanan atau kualitas suatu produk banyak diungkapkan melalui media sosial salah satunya *Twitter*. Setiap ekspresi di *Twitter* baik berupa foto, komentar, dan tagar adalah data yang apabila dianalisa dengan baik dan tepat maka akan sangat bermanfaat. Selain menjadi aset berharga dalam penelitian tentang big data, di sisi lain menjadi nilai jual untuk perusahaan penyedia produk dan jasa. Ini tentunya dapat membuka peluang penelitian lebih banyak dan penyediaan produk dan jasa yang lebih *up to date* karena sesuai dengan perkembangan jaman.

Ada beberapa metode klasifikasi yang digunakan dalam analisis sentimen yang biasanya dilakukan untuk mengetahui opini masyarakat terhadap produk, layanan,

politik, atau topik lainnya, metode klasifikasi yang dimaksud diantaranya adalah *Naive Bayes Classifier* dan *K Nearest Neighbour*.

*Naive Bayes Classifier*, yang merupakan algoritma untuk mencari nilai probabilitas tertinggi untuk mengklasifikasi data uji pada kategori yang paling tepat (Feldman & Sanger, 2007), Metode *Naive Bayes Classification* merupakan salah satu metode yang dapat mengklasifikasikan teks. Kelebihan NBC adalah algoritmanya sederhana tetapi memiliki akurasi yang tinggi.

*K Nearest Neighbour* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama, yaitu berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada. (Kusrini dan Luthfi, 2009). Algoritma K-NN adalah suatu metode yang menggunakan algoritma supervised.

Nihru dalam penelitiannya mengungkapkan, Dalam pembuatan suatu sistem dibutuhkan sebuah metode pendukung yang ada didalamnya. Alasan untuk mengapa memakai Metode KNN ini, karena memiliki keunggulan dapat mengklasifikasikan data yang tidak diketahui dengan adanya data latih dan data uji. KNN dapat menprosedur yang berbasis matematis untuk mengevaluasi nilai kriteria-kriteria tersebut menjadi sebuah keterangan klasifikasi. Metode ini dapat mengklasifikasikan data secara akurat dengan memilih terlebih dahulu nilai K tetangga terdekat dengan tepat. KNN juga bisa memilah kumpulan data calon yang dapat digolongkan menjadi baik, paling baik, dan kurang baik.

Metode *Naive Bayes Classifier* (NBC) banyak digunakan pada penelitian *Text Mining* karena algoritmanya yang sederhana, sehingga mudah dipahami dan memiliki akurasi yang tinggi (Rish, 2006), dalam penelitian yang dilakukan Khotimah dan Wasono (2019) *Sentimen Analysis of E-Commerce Brand Review Using Multinomial Text Naive Bayes* juga menemukan tingkat akurasi menggunakan metode *Naive Bayes Multinomial* mencapai 91.95%. sedangkan Metode *K-Nearst Neighbour* memiliki kelebihan salah satunya yaitu dalam hal data training yang cukup banyak (Asiyah, 2016).

Dari kedua metode diatas yang memiliki kelebihan masing masing, maka peneliti pada penelitian ini memberikan pandangan. Berdasarkan penjabaran latar belakang diatas, peneliti melakukan penelitian dengan judul “Perbandingan Metode

*Naïve Bayes Classifier* dan *K-Nearest neighbor* (K-NN) Terhadap Analisis Sentimen Pada Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat (PPKM)”.

## 1.2 Perumusan Masalah

Masalah yang dibahas dalam penelitian ini adalah :

1. Bagaimana model *K-Nearest Neighbor* dan *Naïve Bayes* diterapkan dalam mengklasifikasikan data opini pengguna *Twitter* tentang PPKM (Pemberlakuan Pembatasan Kegiatan Masyarakat)?
2. Bagaimana perbandingan hasil dari klasifikasi dengan model *K-Nearest Neighbor* dan *Naïve Bayes* mengenai PPKM?

## 1.3 Tinjauan Penelitian

Tujuan dari penelitian ini adalah :

1. Mengetahui hasil kinerja dari penggunaan model klasifikasi *K-Nearest Neighbor* dan *Naïve Bayes* dalam analisis sentimen pengguna *Twitter* terhadap Pemberlakuan PPKM (Pemberlakuan Pembatasan Kegiatan Masyarakat).
2. Mengetahui perbandingan dari penerapan model klasifikasi *K-Nearest Neighbor* dan *Naïve Bayes* dalam analisis sentimen pengguna *Twitter* terhadap Pemberlakuan PPKM (Pemberlakuan Pembatasan Kegiatan Masyarakat).

## 1.4 Manfaat Penelitian

Manfaat dari *penelitian* ini adalah :

1. Manfaat Teoritis

Sebagai pijakan dan referensi pada penelitian-penelitian selanjutnya yang berhubungan dengan analisis *sentimen* text mining dengan menerapkan beberapa metode yang berkaitan serta menjadi bahan kajian lebih lanjut.

2. Manfaat Praktis

Manfaat skripsi ini secara praktis adalah sebagai berikut :

- a. Bagi Pemerintah Indonesia Penelitian ini dapat digunakan sebagai bahan pertimbangan dalam pengambilan kebijakan dengan mengetahui sentimen publik mengenai Pemberlakuan Kebijakan ini.

- b. Bagi Peneliti Penelitian ini diharapkan menjadi sarana untuk menerapkan ilmu pengetahuan serta menambah pemahaman mengenai analisis sentimen menggunakan *Naïve Bayes Classifier* dan *K-Nearest Neighbor*.

### 1.5 Batasan Penelitian

Adapun batasan masalah yang ditentukan dalam penelitian ini adalah:

1. Data *tweet* yang diambil dari *Twitter* yang dari rentang 1 Januari 2021 – 1 September 2021.
2. Metode yang digunakan dalam penelitian ini adalah metode *Naïve Bayes Classifier* dan *K-Nearest Neighbor*.
3. Dalam pengukuran performa klasifikasi cara yang digunakan adalah dengan menghitung akurasi, *precision*, *recall* dan *f-measure*.

