

BAB I PENDAHULUAN

I.1 Latar Belakang

Pada tahun 2019 Dunia dikejutkan oleh wabah mematikan yang menyebarkan diseluruh penjuru dunia. Wabah ini bernama *Corona Virus Disease* tahun 2019 (COVID-19) yang muncul pertama kali di kota Wuhan, Republik Rakyat Tiongkok. Wabah ini akhirnya masuk ke Indonesia dan telah merenggut ratusan nyawa. Pemerintah Indonesia mengeluarkan Peraturan Pemerintah (PP) No. 2 Tahun 2020 tentang Pembatasan Sosial Berskala Besar dalam Rangka Percepatan Penanganan *Corona Virus Disease* 2019 (COVID-19) dan Keputusan Presiden Republik Indonesia No. 11 Tahun 2020 tentang Penetapan Kedaruratan Kesehatan Masyarakat *Corona Virus Disease* 2019 (COVID-19) (Khadapi et al., 2020). Salah satu upaya dalam menangani wabah COVID-19 adalah dengan mengembangkan aplikasi untuk menanggulangi dan mencegah pandemi COVID-19. Aplikasi ini dikembangkan oleh Kementerian Kominfo, Kementerian Kesehatan, Kementerian BUMN, Badan Nasional Penanggulangan Bencana (BPNB) dan dinamakan PeduliLindungi. Aplikasi PeduliLindungi memudahkan pemerintah dalam mengidentifikasi masyarakat yang membutuhkan penanganan lebih lanjut. Pengguna PeduliLindungi juga akan mendapatkan notifikasi jika suatu kawasan ramai dan kawasan tersebut akan ditandai dengan zona merah. Aplikasi ini juga membantu masyarakat dalam melakukan vaksinasi COVID-19. Aplikasi ini dapat digunakan untuk mengakses layanan publik seperti restoran, bioskop, transportasi umum dengan melakukan *scan* melalui fitur *QR Code* (Sudiarsa & Wiraditya, 2020).

Penelitian ini menggunakan analisis sentimen terhadap aplikasi PeduliLindungi karena penelitian ini akan berfokus pada ulasan pengguna aplikasi PeduliLindungi untuk mengetahui sentimen masyarakat terhadap aplikasi PeduliLindungi. Analisis sentimen merupakan teknik atau cara untuk mengidentifikasi bagaimana sebuah sentimen bisa dikategorikan sebagai sentimen positif, netral atau negatif (Medhat et al., 2014). Semenjak munculnya aplikasi PeduliLindungi sebagai salah satu cara dalam mencegah dan menanggulangi penyebaran COVID-19 muncul berbagai macam opini dari masyarakat terhadap kinerja aplikasi PeduliLindungi.

Penelitian akan berfokus pada ulasan yang diberikan pengguna dalam pemakaian aplikasi PeduliLindungi. Analisis sentimen juga mempelajari analisis emosi, mendeteksi sarkasme dan analisis sentimen multibahasa (L. Zhang et al., 2018). Penelitian ini sangat diperlukan karena dengan menggunakan analisis sentiment kita dapat mengklasifikasikan dan mengetahui pendapat masyarakat terhadap aplikasi PeduliLindungi. Dengan mengetahui informasi yang didapatkan dari penelitian ini stakeholder dari aplikasi PeduliLindungi dapat mengambil kesimpulan dan keputusan sesuai hasil dari penelitian ini (Brugha & Varvasovszky, 2000). Karena analisis sentimen akan dilakukan maka penelitian ini juga akan menggunakan metode *text mining*.

Sebagian besar informasi yang tersedia muncul dalam bentuk teks yang tidak terstruktur, karena hal itu maka dibutuhkan teknik khusus yang beroperasi pada data tekstual untuk mengekstrak informasi pada kumpulan teks (Rajman & Besançon, 1998). *Text mining* merupakan suatu proses menambang data yang berupa teks (Hotho et al., 2005). *Text mining* memiliki proses yang sangat penting dinamakan *Preprocessing*. *Preprocessing* ini diperlukan untuk melakukan klasifikasi sentiment (Krouska et al., 2016). *Preprocessing* memiliki beberapa tahap seperti *stemming*, *stopwords removal*, dan *tokenization* (Vijayarani et al., 2015). Metode ini diperlukan karena data yang akan digunakan pada penelitian berupa teks

Untuk melakukan analisis sentimen digunakan *machine learning* yang bertujuan untuk mendeteksi pola dan mengklarifikasikan data pada suatu model. *Machine learning* merupakan mesin yang diciptakan untuk belajar layaknya manusia dan mampu mempelajari suatu data yang diberikan. Dalam *machine learning* terdapat tiga metode utama. Metode yang pertama adalah *supervised learning* dimana terdapat label pada data yang akan ditambahkan pada model *machine learning*. *Supervised learning* banyak digunakan untuk memprediksi suatu pola dari data yang sudah lengkap menjadi pola yang terbentuk dari hasil pembelajaran data lengkap tersebut. Yang kedua yaitu *Unsupervised learning* dimana metode ini tidak memiliki label tetapi menggunakan kesamaan dari atribut-atribut yang dimiliki. Metode terakhir yaitu *Reinforced Learning* merupakan metode yang

diterapkan pada *machine learning* yang pada akhirnya suatu program dapat bekerja secara otomatis dan memberikan hasil yang tepat (Mahesh, 2020).

Salah satu algoritma pada *supervised learning* adalah *K-Nearest Neighbor* yang nanti akan digunakan pada penelitian ini. Algoritma *K-Nearest Neighbor* adalah salah satu metode paling populer dalam teknologi klasifikasi *data mining* (Gao & Li, 2020). Algoritma *K-Nearest Neighbor* bertujuan untuk mengklasifikasikan pengamatan yang tidak berlabel dengan menetapkannya kedalam kelas dari contoh berlabel yang paling mirip (Z. Zhang, 2016). Algoritma ini termasuk mudah untuk diimplementasikan karena hanya membutuhkan parameter tetangga terdekat dan fungsi jarak. Beberapa fungsi jarak tersebut yaitu adalah *Chebyshev Distance*, *Euclidean Distance*, *Manhattan Distance*, dan *Minkowski Distance* (Gao & Li, 2020). Adapun beberapa kesulitan saat menggunakan algoritma ini seperti jumlah data yang besar, susah bekerja dengan informasi yang tidak lengkap dan biaya komputasi yang tinggi. Tujuan algoritma ini adalah untuk mengklasifikasikan objek baru berdasarkan data *training*.

I.2 Perumusan Masalah

Rumusan masalah yang mendasari penelitian ini adalah:

- a. Bagaimana melakukan klasifikasi komentar pengguna aplikasi PeduliLindungi?
- b. Bagaimana mengetahui performa model algoritma *K-Nearest Neighbor* dalam pengklasifikasian teks?
- c. Bagaimana melakukan evaluasi terhadap performa model algoritma *K-Nearest Neighbor*.

I.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

- a. Menganalisis komentar pada aplikasi PeduliLindungi menggunakan algoritma *K-Nearest Neighbor*.
- b. Mengklasifikasikan komentar positif, netral dan negatif pengguna aplikasi PeduliLindungi.

- c. Menguji tingkat akurasi algoritma *K-Nearest Neighbor* dalam pengklasifikasian *text mining*.

I.4 Batasan Penelitian

Adapun batasan masalah dalam penelitian ini sebagai berikut :

1. Metode yang digunakan untuk pengklasifikasian adalah metode *K-Nearest Neighbor*.
2. Penelitian ini berfokus pada analisis sentimen yang berupa ulasan positif, netral dan negatif dari pengguna aplikasi PeduliLindungi.
3. Dataset yang diambil dari ulasan pengguna aplikasi PeduliLindungi yang terdapat pada *Google Play* sejumlah 10.000 data.

I.5 Manfaat Penelitian

Manfaat penelitian ini:

1. Bagi pengembang aplikasi PeduliLindungi, penelitian ini dapat membantu mengoptimalkan aplikasi PeduliLindungi
2. Bagi masyarakat, penelitian ini bermanfaat dalam peningkatan kenyamanan penggunaan aplikasi yang digunakan oleh masyarakat.