

# BAB I PENDAHULUAN

## I.1 Latar Belakang

Sejak awal tahun 2020, masyarakat dunia dihebohkan dengan adanya sebuah virus mematikan yang menyebar luas dalam waktu singkat yaitu *SARS-CoV-2*. *SARS-CoV-2* ini merupakan jenis virus terbaru dari *Coronavirus*. Penyakit yang disebabkan oleh virus ini dinamakan *Coronavirus disease 2019 (Covid-19)*. Berdasarkan data *World Health Organization (WHO)* terdapat sebanyak 241,411,380 kasus di seluruh dunia, dengan 457,013 kasus baru per tanggal 21 Oktober 2021 (*World Health Organization, 2021*).

Awal bulan Maret 2020, Presiden Joko Widodo menyampaikan langsung temuan kasus Covid-19 pertama di Indonesia (*Gitiyarko, 2020*). Pada bulan Oktober 2021 ini sendiri sudah terdapat 4,237,201 kasus, dengan 914 kasus baru. Dimana jika dilihat pada diagram terjadi kenaikan kasus pada bulan Juni-Agustus dan terjadi penurunan pada bulan September hingga saat ini (*World Health Organization, n.d.*).

Tindakan awal yang dilakukan oleh Presiden Joko Widodo saat virus Covid-19 muncul adalah dengan memerintahkan kedutaan Indonesia di China untuk memberi perhatian khusus terhadap WNI yang terisolasi di Wuhan. Setelah adanya kasus pertama di Indonesia, pemerintah berusaha melakukan penanganan cepat dengan membentuk tim satuan tugas penanggulangan Covid-19 yang dipimpin langsung oleh Presiden. Sejak tanggal 15 Maret 2020 Presiden membuat kebijakan kerja dari rumah, belajar dari rumah dan ibadah dari rumah, juga pemerintah dan lapisan masyarakat mengkampanyekan pembatasan sosial (*Social Distancing*) demi mencegah penularan Covid-19. Dalam upaya melindungi masyarakat dari resiko penularan, Presiden Jokowi menetapkan peraturan tentang Pembatasan Sosial Berskala Besar (PSBB) dan mulai berlaku sejak 1 April 2020 (*Gitiyarko, 2020*).

Vaksin menjadi salah satu upaya pemerintah dalam menanggulangi pandemi Covid-19. Pemerintah telah memesan 1,2 juta dosis vaksin Covid-19 buatan *Sinovac Biotech*, rencananya akan menjadi vaksin *Coronavirus* pertama yang tersedia di Indonesia. Uji klinis tahap I dan II vaksin *Sinovac Biotech* ini pun sebenarnya sejak bulan April 2020 hingga Mei 2020. Hasilnya juga telah dipublikasi di jurnal ilmiah *The Lancet* pada 17 November 2020 silam dengan kesimpulan berlanjut ke uji klinis tahap III. Selain vaksin *Sinovac* terdapat beberapa jenis vaksin yang ditetapkan pemerintah dimana rencananya akan diproduksi oleh PT Bio Farma (Persero), *AstraZeneca*, *China National Pharmaceutical Group Corporation (Sinopharm)*, *Moderna*, *Pfizer Inc and BioNTech* dan *Sinovac Biotech Ltd* (Sitanggang, 2020). Berdasarkan data kemenkes Indonesia terdapat 208,265,720 sasaran vakksinasi dan telah tercapai 110,406,777 (53.01%) dosisi pertama serta 65,173,148 (31.29%) dosis kedua (Vaksinasi Covid-19 Nasional, n.d.).

Berdasarkan data tersebut, target vaksin yang ditetapkan pemerintah belum tercapai, dimana baru terdapat 36,59% masyarakat yang telah melakukan vaksin pada bulan November 2021 baik dosis satu maupun dosis dua. Sementara itu, pemerintah meningkatkan vaksinasi 70% pada akhir tahun 2021 terutama untuk vaksin dosis kedua. Respon masyarakat terhadap vaksinasi pemerintah sangat bervariasi. Analisis sentimen dapat memberikan informasi kepada banyak pihak, dimana analisis sentimen merupakan proses pengklasifikasian teks menjadi sentimen positif dan negatif. Pendapat tersebut haruslah dipertimbangkan sebagai bahan evaluasi agar program vaksinasi yang dilaksanakan dapat berjalan dengan baik dan pemerintah dapat melakukan sosialisasi lebih lanjut untuk meningkatkan kesadaran masyarakat akan vaksinasi Covid-19.

Metode yang banyak digunakan untuk melakukan analisis sentimen adalah algoritma *K-Nearest Neighbor* dan *Logistic Regression*. Pada *machine learning* kedua algoritma tersebut merupakan teknik *supervised learning*. Pada sudut pandang parameter algoritma yang harus dioptimasi, *K-Nearest Neighbor* merupakan kategori *non-*

*parametrik* yaitu tidak mengasumsikan permasalahan, sementara *Logistic Regression* merupakan kategori *parametrik* dimana harus dilakukan reduksi permasalahan sebagai optimasi parameter. Sedangkan dari sudut pandang lainnya, *K-Nearest Neighbor* termasuk ke dalam model *non-linear* sementara *Logistic Regression* termasuk ke dalam model *linear* (Putra J. W., 2020).

Berdasarkan perbedaan tersebut maka dilakukan perbandingan performansi dan akurasi yang dihasilkan oleh kedua algoritma tersebut untuk menentukan algoritma mana yang memiliki performansi paling optimal dan akurasi paling tinggi. Pada penelitian ini digunakan dua metode pelabelan yaitu dengan pelabelan *TextBlob* dan *Vadert*. Pemilihan kedua metode pelabelan tersebut karena *TextBlob* dan *Vader* adalah dua *library python* untuk pemrosesan teks. Bedanya di *TextBlob*, polaritasnya ditentukan dengan menghitung jumlah kalimat/ulasan positif dan negatif kemudian diberikan skor polaritas menggunakan fungsi *sentiment()*. Sementara itu, *Vader* menganalisis sepotong teks untuk melihat apakah ada kata dari teks dalam kamus *Vader*, serta penentuan indeks polaritas pada *Vader* menggunakan fungsi *polarity\_score()*. Perbedaan karakteristik tersebut menyebabkan hasil pelabelan yang berbeda, sehingga dilakukan perbandingan untuk menentukan teknik pelabelan mana yang memberikan hasil optimal pada dataset dalam analisis sentimen program Vaksinasi Covid-19 pada Media Sosial *Twitter*.

## **I.2 Perumusan Masalah**

Berdasarkan latar belakang yang ada maka didapatkan rumusan masalah yang mendasari penelitian ini sebagai berikut:

- a. Bagaimana implementasi algoritma *K-Nearest Neighbor* dan algoritma *Logistic Regression* terhadap analisis sentiment Vaksinasi Covid-19 pada media sosial *twitter*?
- b. Bagaimana tingkat akurasi yang dihasilkan algoritma *K-Nearest Neighbor* dan algoritma *Logistic Regression* terhadap analisis sentimen Vaksinasi Covid-19 pada media sosial *twitter*?

- c. Bagaimana hasil perbandingan sentimen positif dan sentimen negatif terhadap Vaksinasi Covid-19 pada media sosial *twitter* dengan menggunakan pelabelan *TextBlob* dan pelabelan *VaderSentiment*?

### **I.3 Tujuan Penelitian**

Penelitian ini bertujuan untuk:

- a. Mengimplementasikan algoritma *K-Nearest Neighbor* dan algoritma *Logistic Regression* terhadap analisis sentimen Vaksinasi Covid-19 pada media sosial *twitter*.
- b. Mendapatkan hasil akurasi dari algoritma *K-Nearest Neighbor* dan *Logistic Regression* terhadap analisis sentimen Vaksinasi Covid-19 pada media sosial *twitter*.
- c. Mengetahui perbandingan sentimen positif dan sentimen negatif terhadap Vaksinasi Covid-19 pada media sosial *twitter* dengan menggunakan pelabelan *TextBlob* dan pelabelan *VaderSentiment*.

### **I.4 Batasan Penelitian**

Batasan dalam penelitian ini, sebagai berikut:

- a. Penelitian ini menggunakan data dari media sosial *twitter* dengan *keyword* yang berkaitan dengan vaksin Covid-19, pengambilan data dilakukan pada bulan November 2022, Desember 2021, Januari 2022, Maret 2022 dan Mei 2022.
- b. Penelitian ini hanya mengambil *tweet* yang menggunakan Bahasa Indonesia untuk di analisis.
- c. Pelabelan pada penelitian ini adalah label positif (*tweet* yang mendukung vaksinasi Covid-19) dan label negatif (*tweet* yang kontra akan vaksinasi Covid-19).
- d. Penelitian ini hanya mengambil kasus vaksinasi Covid-19 di Indonesia.

## **I.5 Manfaat Penelitian**

Manfaat penelitian ini, sebagai berikut:

- a. Bagi peneliti, menambah wawasan dalam proses mengklasifikasikan data menggunakan algoritma *K-Nearest Neighbor* dan algoritma *Logistic Regression*.
- b. Bagi akademik, diharapkan dapat memberikan kontribusi dalam memberikan informasi dan referensi klasifikasi untuk penelitian lebih lanjut
- c. Bagi pemerintah, diharapkan dapat memberikan informasi mengenai respon masyarakat dan dapat meningkatkan sosialisasi lebih untuk mengajak masyarakat vaksin
- d. Bagi masyarakat, diharapkan dapat memberikan informasi mengenai respon masyarakat terkait vaksin Covid-19

## **I.6 Sistematika Penulisan**

Tugas akhir ini diuraikan dengan sistematika penulisan sebagai berikut:

### **Bab I Pendahuluan**

Bab ini berisi uraian tentang konteks masalah, latar belakang masalah, rumusan masalah, tujuan penelitian, batasan penelitian, manfaat penelitian, dan sistematika penulisan.

### **Bab II Tinjauan Pustaka**

Bab ini berisi literatur yang berkaitan dengan masalah yang diteliti dari hasil buku referensi/penelitian/referensi lain yang dapat digunakan untuk merancang dan memecahkan masalah. Di akhir bab ini, dijelaskan alasan pemilihan kerangka kerja penelitian di tugas akhir ini.

### **Bab III Metodologi Penelitian**

Bab ini merupakan uraian tentang metode/konsep/kerangka yang dipilih dalam bab Tinjauan Pustaka, dan juga memuat uraian rinci tentang prosedur penelitian, termasuk model konseptual dan pemecahan masalah yang sistematis.

**Bab IV Analisis dan Perancangan**

Bab ini mencakup fase pengumpulan data dan fase pra-pemrosesan data. Keluaran dari fase ini berupa data yang digunakan dalam penelitian.

**Bab V Hasil dan Evaluasi**

Bab ini berisi uraian tahap implementasi terhadap penelitian, akan dilakukan pengujian hasil hingga evaluasi model.

**Bab VI Kesimpulan dan Saran**

Bab ini dijelaskan kesimpulan dan penyelesaian masalah yang dilakukan serta jawaban dari rumusan permasalahan yang ada pada bagian pendahuluan. Saran dan solusi dikemukakan pada bab ini untuk tugas akhir selanjutnya.