

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Di awal tahun 2020, seluruh dunia dikejutkan dengan adanya penyakit Coronavirus-19 atau biasa disebut COVID-19 yang ditemukan pertama kali di daerah Wuhan, China [1]. Virus COVID-19 ini menyebabkan hampir seluruh dunia mengalami *lockdown* dan menyebabkan kematian lebih dari 6,22 juta jiwa pertanggal 26 April 2022 [2]. Menyikapi hal tersebut, pemerintah Indonesia memutuskan untuk mendapatkan vaksin COVID-19 bagi masyarakat dengan harapan yang sangat besar agar bisa kembali beraktivitas [3]. Pada tanggal 13 Januari 2021, Presiden Indonesia bersedia untuk mendapatkan vaksin Sinovac agar dapat meningkatkan kepercayaan masyarakat terhadap vaksin COVID-19 dan sudah melalui proses uji klinis tahap III dan sudah mendapatkan EUA (Emergency use authorization) dari Badan Pengawas Obat dan Makanan (BPOM) RI [4]. Akan tetapi, banyak pihak yang masih memperdebatkan pro dan kontra dari vaksin COVID-19, bahkan ada yang menolak kehadiran vaksin walaupun pemerintah sudah memberikannya secara gratis kepada masyarakat [5]. Media sosial dan kebebasan berpendapat membuat sebuah informasi yang belum tentu benar lebih cepat menyebar di lingkungan masyarakat [3]. Hal tersebutlah yang menimbulkan pro dan kontra di antara masyarakat [3].

Sebagian besar masyarakat Indonesia pengguna layanan digital mengakses informasi mengenai COVID-19 melalui media sosial [51]. Sebanyak 81,5% diantaranya masih bersinggungan dengan berbagai bentuk postingan-postingan yang memuat teori konspirasi [51]. Mayoritas masyarakat Indonesia menyakini bahwa vaksin COVID-19 dibuat untuk mendapatkan keuntungan pihak farmasi ataupun untuk memasukkan *microchip* dalam tubuh manusia [51]. Ditambah lagi pada awal Maret tahun 2022, dunia digemparkan dengan kemunculan *hepatitis* varian baru yang penyebabnya masih belum diketahui. Sejalan dengan hal tersebut, ditengah-tengah lingkungan masyarakat muncul informasi yang mengatakan bahwa *unknown hepatitis* ini disebabkan oleh vaksin COVID-19 [52]. Kemudian pada akhir maret 2022, Presiden Joko Widodo menetapkan wajib untuk mendapatkan

vaksin *booster* untuk dapat melakukan mudik lebaran membuat polemik ditengah masyarakat. Walaupun hal ini merupakan strategi pemerintah agar masyarakat Indonesia segera mendapatkan vaksin COVID-19 secara lengkap, banyak masyarakat yang merasa terbebani dengan keputusan yang dibuat oleh pemerintah [53]. Masyarakat beranggapan bahwa vaksin booster sebagai syarat mudik ini bersifat tidak adil terhadap masyarakat, hal ini dikarenakan banyak kegiatan yang menimbulkan kerumunan seperti acara konser, liburan tahun baru, dan pagelaran MotoGP yang diadakan di Mandalika, dan tidak ada persyaratan untuk vaksinasi kedua [57]. Apalagi saat ini, penerima vaksin juga belum merata dan masih banyak masyarakat yang baru menyelesaikan vaksin kedua [57]. Agar kehadiran vaksin COVID-19 dapat diterima oleh seluruh masyarakat, maka Pemerintah perlu memperkenalkan dan penyuluhan secara intens dan menyeluruh mengenai pentingnya vaksin COVID-19 bagi kesehatan [5]. Oleh karena itu, perlu dilakukan analisis sentimen pendapat masyarakat dan mengklasifikasikannya kedalam kelas positif, negatif, atau netral.

Untuk melihat opini masyarakat terhadap vaksin COVID-19 dapat dilihat dari berbagai media salah satunya yaitu media sosial Twitter [6]. Pengguna aktif media sosial Twitter di Indonesia saat ini cukup banyak, Indonesia dinilai sebagai salah satu negara dengan pertumbuhan pengguna aktif harian media sosial Twitter paling banyak berdasarkan laporan finansial Twitter tahun 2019 yang digunakan untuk berkomunikasi berbagai informasi dengan pengguna lain [7]. Isi tweet ini dapat dijadikan sebagai sumber data pendapat dan sentimen masyarakat terhadap vaksin COVID-19 yang dapat digunakan untuk analisis dan bahan evaluasi bagi pemerintah Indonesia dalam memberikan edukasi kepada masyarakat mengenai program vaksin COVID-19. Pada penelitian ini, data diambil menggunakan teknik *crawling* dengan menggunakan bahasa pemrograman *python* dengan *library tweepy*. Analisis sentimen ini nantinya akan mengklasifikasikan data menjadi kelas yang bersifat positif, negatif, atau netral.

Penelitian yang berkaitan dengan analisa sentimen telah dilakukan oleh Miftahul Ihsan, Benny Sukma Negara, dan Surya Agustian mengenai pendapat mengenai vaksin COVID-19 pada media sosial Twitter dengan perolehan nilai

akurasi sebesar 66% dengan menggunakan metode *Long Short-Term Memory* (LSTM) dengan model *word embeddings* Word2Vec [60]. Analisa sentimen juga dilakukan oleh Nadia Ristya Dewi, Eva Yulia Puspanigrum, dan Hendra Maulana dengan menggunakan RNN dengan Metode TF-IDF dan Word2Vec yang membahas mengenai pendapat masyarakat Indonesia mengenai vaksin COVID-19 dengan perolehan nilai akurasi dengan metode *word embedding* Word2Vec dengan RNN sebesar 53%, sedangkan perolehan nilai akurasi dengan metode RNN dengan TF-IDF menghasilkan nilai akurasi sebesar 51% [61].

Berdasarkan latar belakang tersebut, diusulkan metode *Long Short-Term Memory* (LSTM) dengan *word embedding* untuk menganalisis sentimen masyarakat Indonesia terhadap vaksin COVID-19. Penelitian menggunakan pencarian pada Twitter dengan kata kunci “vaksin” dan “vaksin covid”. Data yang diperoleh dari proses crawling tersebut akan diklasifikasikan menjadi kelas positif, negatif, atau netral. Algoritma *machine learning Long Short-Term Memory (LSTM)* yang digunakan pada penelitian ini diharapkan mampu memberikan hasil analisis data yang sesuai dengan akurasi yang terbaik dan lebih baik dari penelitian sebelumnya [60]. Hasil yang diharapkan pada penelitian ini yaitu dapat membantu pemerintah Indonesia untuk mengetahui respon masyarakat terkait vaksin COVID-19 serta bisa digunakan sebagai bahan evaluasi dan mempermudah pemerintah dalam menentukan langkah selanjutnya yang terkait dengan edukasi kepada masyarakat Indonesia mengenai vaksin COVID-19 yang saat ini tengah digencarkan pemerintah untuk vaksin *booster* bagi masyarakat.

## 1.2. Rumusan Masalah

Berdasarkan latar belakang diatas, maka rumusan masalah yang akan dibahas pada penelitian ini yaitu :

1. Bagaimana melakukan klasifikasi sentimen masyarakat Indonesia terhadap vaksin COVID-19 menggunakan algoritma *Long Short-Term Memory (LSTM)*?

2. Bagaimana mengetahui hasil klasifikasi sentimen masyarakat Indonesia terhadap vaksin COVID-19 menggunakan algoritma *Long Short-Term Memory (LSTM)*?
3. Bagaimana melakukan pengukuran akurasi untuk klasifikasi sentimen masyarakat Indonesia terhadap vaksin COVID-19 menggunakan algoritma *Long Short-Term Memory (LSTM)*?
4. Bagaimana membangun antar muka pengguna menggunakan *framework website flask* dari hasil klasifikasi sentimen masyarakat Indonesia terhadap vaksin COVID-19 menggunakan *flask python*?

### 1.3. Batasan Masalah

Bagian batasan masalah pada penelitian ini yaitu:

1. Data yang digunakan untuk klasifikasi kategori sentimen pada penelitian ini yaitu *data collection* dari Twitter menggunakan *library tweepy* bahasa pemrograman Python dengan kata kunci “vaksin” dan “vaksin covid”.
2. Proses Modelling dengan menggunakan algoritma *Long Short-Term Memory (LSTM)*.
3. Implementasi model yaitu *website* dan *framework flask python*.
4. Membagi opini ke dalam tiga jenis klasifikasi yaitu positif, negatif, atau netral.
5. Dataset yang digunakan adalah dataset berbahasa Indonesia.

### 1.4. Tujuan

Tujuan dari penelitian ini yaitu :

1. Mengetahui cara kerja klasifikasi sentimen masyarakat Indonesia terhadap vaksin COVID-19 menggunakan algoritma *Long Short-Term Memory (LSTM)*.
2. Mengetahui hasil dari klasifikasi sentimen masyarakat Indonesia terhadap vaksin COVID-19 menggunakan algoritma *Long Short-Term Memory (LSTM)*.

3. Mengetahui cara untuk melakukan pengukuran akurasi untuk klasifikasi sentimen masyarakat Indonesia terhadap vaksin COVID-19 menggunakan algoritma *Long Short-Term Memory (LSTM)*.
4. Mengetahui cara membangun antar muka pengguna menggunakan *framework website flask* klasifikasi pendapat masyarakat Indonesia terhadap vaksin COVID-19.
5. Membangun aplikasi penentuan sentimen positif, negatif, dan netral untuk mengetahui pendapat masyarakat Indonesia mengenai vaksin COVID-19.

## 1.5. Manfaat

Manfaat dari penelitian ini yaitu :

1. Memperoleh hasil sentimen positif, negatif, dan netral dari pendapat masyarakat Indonesia mengenai vaksin COVID-19 sehingga diharapkan bisa ditindaklanjuti secara dalam oleh kepentingan yang bersangkutan.
2. Aplikasi Sentimen Pendapat Masyarakat Indonesia Terhadap Vaksin Covid-19 dapat diimplementasikan dan digunakan oleh kepentingan yang bersangkutan.

## 1.6. Metodologi Penelitian

### 1.6.1 Metode Pengumpulan Data

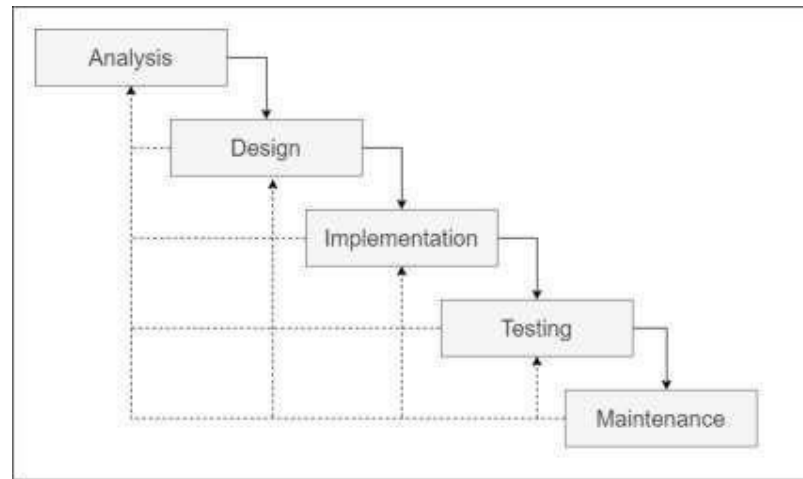
Metode pengumpulan data pada penelitian ini yaitu menggunakan bahasa pemrograman Python dengan menggunakan library *tweepy*. Data yang dikumpulkan sebanyak 4000 data dari rentang tanggal 20 Maret 2022 sampai dengan 2 April 2022. Kemudian dilakukan tahap *labelling* untuk pemberian sentimen terhadap setiap kalimat yang dilakukan secara manual oleh penulis. Dan dilakukan proses text preprocessing terhadap data sebelum dilakukan proses modelling menggunakan algoritma LSTM (*Long Short-Term Memory*).

### 1.6.2 Metode Pengembangan Sistem

Metode pengembangan sistem pada penelitian ini yaitu menggunakan model SDLC (*System Development Life Cycle*). SDLC merupakan proses yang meliputi tahap pengembangan, pemeliharaan, serta penggunaan sistem informasi yang digunakan untuk mengembangkan suatu sistem [44]. SDLC terdiri dari beberapa

tahap yaitu rencana (planning), analisis (analysis), desain (design), implementasi (implementation), uji coba (testing), dan pengelolaan (maintenance) [45].

Model SDLC yang dipakai dalam penelitian ini yaitu model Waterfall. Model waterfall merupakan pengembangan sistem yang tahapnya dilakukan dengan diselesaikannya tahap demi tahap yang dilalui dan berjalan secara berurutan [46].



Gambar 1.1 Tahapan Model Waterfall [47]

*a. Requirements analysis and definition*

Kebutuhan sistem, kendala, dan tujuan yang sesuai dengan kebutuhan pengguna akan didefinisikan secara rinci dan berfungsi sebagai spesifikasi sistem [50].

*b. System and software design*

Perancangan perangkat lunak tidak lepas dengan identifikasi dan penggambaran sistem dasar perangkat lunak (*software design*). Tahapan dari perancangan design mengalokasikan kebutuhan-kebutuhan yang dibutuhkan oleh sistem, baik kebutuhan perangkat keras maupun perangkat lunak [50]. Pada penelitian ini sistem akan dibangun dengan menggunakan algoritma Long Short-Term Memory (LSTM) dengan *word embedding*. Kemudian hasil dari modelling diimplementasikan melalui *framework flask* python berbasis *website*.

*c. Implementation and unit testing*

Pada tahap ini, rancangan perangkat lunak akan direalisasikan menjadi unit program. Tahap *testing* dilakukan dengan memverifikasi bahwa setiap unit pada program berfungsi sesuai dengan kebutuhannya [50].

*d. Integration and system testing*

Unit-unit program akan digabung dan diuji sebagai suatu program yang lengkap guna untuk memastikan apakah sudah sesuai dengan kebutuhan perangkat lunak atau tidak. Setelah melalui tahap *testing*, perangkat lunak akan dikirimkan kepada *customer* [50].

*e. Operation and maintenance*

Tahapan ini merupakan tahapan yang paling di beberapa penelitian. Sistem akan digunakan secara nyata oleh *customer*. *Maintenance* berguna untuk pembetulan kesalahan-kesalahan yang tidak ditemukan pada pengujian yang telah dilakukan sebelumnya, meningkatkan implementasi dari sistem, serta meningkatkan layanan pada sistem sebagai kebutuhan fungsional yang baru [50].

**1.7. Jadwal Pelaksanaan**

**Tabel 1.1** Jadwal Penelitian

No	Kegiatan	Waktu Penelitian						
		Juli	Agustus	September	Oktober	November	Desember	Januari
1.	Studi Literatur							
2.	Analisis Kebutuhan (Planning System)							
3.	Pengumpulan Data							
4.	Pelabelan Data							
5.	Preprocessing Data							

6.	Desain sistem dan software							
7.	Pembuatan Model Klasifikasi <i>Long Short-Term Memory</i> (LSTM)							
8.	Implementasi dan Testing setiap unit							
9.	Proses Hasil Klasifikasi							
10.	Pembuatan <i>framework website</i> dengan <i>flask python</i> .							
11.	Integrasi dan testing sistem							
12.	Pengujian <i>website</i>							
13.	Pemeliharaan sistem							
10.	Penulisan Tugas Akhir							

### 1.8. Sistematika Penulisan

Sistematika penulisan dalam penelitian ini terbagi menjadi beberapa bab yaitu :

#### **BAB I Pendahuluan**

Bab ini berisi mengenai latar belakang, rumusan masalah, tujuan dan manfaat, batasan masalah, metode penelitian, dan jadwal pelaksanaan.

#### **BAB II Tinjauan Pustaka**

Bab ini berisi mengenai penjelasan dari teori yang digunakan dalam pembuatan tugas akhir ini yang meliputi Analisis Sentimen, *Natural Language Processing*, *Long Short-Term Memory (LSTM)*, *Text Preprocessing*, *Twitter*, *Confusion Matrix*, *website*, *metode pengujian sistem*.

#### **BAB III Analisis dan Perancangan Sistem**



Bab ini berisi mengenai analisis sentimen pendapat masyarakat Indonesia mengenai vaksin COVID-19 pada media sosial Twitter dengan menggunakan algoritma *Long Short-Term Memory (LSTM)* untuk melakukan klasifikasi kelas sentimen menjadi tiga kategori yaitu positif, negatif, atau netral. Analisis sistem ini meliputi analisis masalah dan kebutuhan arsitektur umum dan perancangan sistem.

#### **BAB IV Hasil Percobaan dan Analisis**

Bab ini berisi mengenai hasil dari penelitian yang telah dilakukan. Topik ini berisikan implementasi dari kode program yang telah dibuat.

#### **BAB V Kesimpulan dan Saran**

Bab ini berisi kesimpulan dari penelitian yang telah dilakukan.