

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Menurut *World Health Organization* (WHO) definisi atau pengertian kesehatan adalah kondisi sehat sejahtera baik secara fisik, mental maupun sosial yang ditandai dengan tidak adanya gangguan-gangguan atau gejala penyakit, misalnya keluhan kesakitan fisik, dan keluhan emosional [1].

Dengan hadirnya Internet sebagai salah satu fenomena paling berpengaruh dalam beberapa dekade terakhir ini, baik bagi perkembangan teknologi itu sendiri maupun bagi perkembangan disiplin ilmu lain seperti ilmu kesehatan, kita dapat memperoleh banyak kemudahan banyak informasi tentang kesehatan yang dapat kita akses melalui internet.

Walaupun internet sudah dapat menyediakan berbagai macam informasi tentang penyakit dan kebiasaan hidup sehat, masih banyak masyarakat yang belum mempunyai kepedulian akan hal tersebut. Dari 262 juta jiwa di Indonesia, hanya sekitar 52 juta orang yang memiliki kepedulian terhadap kebersihan lingkungan sekitar dan dampaknya terhadap kesehatan. Hal ini mengakibatkan setidaknya 2,2 juta jiwa terutama anak-anak meninggal dunia, penyebab dari kematian tersebut adalah karena kurangnya kepedulian akan perilaku hidup bersih dan sehat sehingga menimbulkan berbagai macam penyakit [2].

*World Health Organization* (WHO) atau Organisasi Kesehatan Dunia didirikan pada tanggal 7 April 1948 dan bermarkas di Jenewa, Swiss mempunyai standar minimum untuk satu dokter per populasi yaitu 1 dokter untuk 1000 penduduk (1:1000) atau idealnya dua puluh tiga dokter, perawat, dan bidan per 10.000 penduduk [3].

Indonesia sebagai negara kepulauan dan negara yang tergolong dalam kriteria negara berkembang mempunyai masalah dalam pendistribusian dokter dan tenaga kesehatan. Menurut *World Health Organization's Global Health*

*Workforce Statistics, OECD, supplemented by country data* nilai dokter per kapita Indonesia hanya berada di nilai 0.4:1000 pada tahun 2018 [4].

Dengan angka tersebut Indonesia berada di bawah standar yang ditetapkan WHO, menempati urutan 150 dari 197 negara di dunia, dan tertinggal jauh dari negara tetangga Malaysia yang memiliki nilai dokter per-kapita sebesar 1.54:1000 dan berada di urutan ke 102 dari 197 negara [4].

Hal ini berbanding terbalik dengan jumlah pengguna internet di Indonesia yang sudah mencapai angka 150 juta jiwa yang artinya 56% dari total penduduk Indonesia sudah dapat mengakses internet. Jumlah tersebut hanya selisih sedikit dengan jumlah pengguna pengguna aktif media sosial berbasis *mobile* yang berjumlah 142.8 juta jiwa atau dengan persentase sebesar 53% dari jumlah penduduk di Indonesia [5].

Salah satu sosial media atau *messaging app* yang banyak digunakan adalah aplikasi telegram. Telegram menjadi aplikasi *non-game* yang paling banyak diunduh di seluruh dunia pada bulan Januari 2021. Aplikasi telegram mencapai angka 63 juta penginstalan atau 3,8 kali lipat lebih banyak dari jumlah unduhan pada Januari 2020.

Negara dengan jumlah instalasi pengguna telegram terbesar adalah India dengan jumlah 24%, lalu diikuti oleh Indonesia dengan jumlah 10%. Lonjakan persentase pengunduhan Telegram terjadi setelah salah satu kompetitor yaitu WhatsApp merilis persyaratan layanan terbarunya tentang peningkatan masalah privasi [6].

Solusi yang diusulkan untuk masalah tersebut salah satunya dengan cara merancang sebuah aplikasi konsultasi kesehatan yang berupa *chatbot* dengan pendekatan *Natural Language Processing* (NLP).

*Chatbot* merupakan suatu implementasi dari aplikasi perangkat lunak yang dapat berfungsi menyerupai orang asli untuk melakukan percakapan dan, penalaran. Tujuan utama dari *chatbot* secara umum adalah untuk mengurangi waktu yang dihabiskan pengguna untuk mencari informasi yang benar [7].

Terdapat berbagai macam jenis *chatbot* yang sudah dikembangkan dan juga diimplementasikan di berbagai sektor, salah satu *chatbot* yang populer dan juga mudah atau lebih familiar digunakan untuk berbagai kalangan dengan usia dan juga latar belakang yang berbeda-beda adalah *Rule-based chatbot* [8].

Rafael Mellado-Silva dan rekan nya melakukan sebuah penelitian yang dipublikasikan dengan judul “*Learning tax regulations through rules-based chatbots using decision trees: a case study at the time of COVID-19*” [9]. Penelitian tersebut dipublikasikan pada tahun 2020, tujuan penelitian ini adalah untuk mengembangkan sebuah media pembelajaran tentang regulasi pajak yang digunakan untuk dapat membantu kegiatan belajar mengajar saat masa pandemi covid-19. Hasil dari penelitian tersebut, *rule-based* yang digunakan untuk media belajar *online* dapat meningkatkan kemampuan peserta didik sebesar 17%, sehingga dapat disimpulkan bahwa *rule-based chatbot* merupakan metode yang efektif sebagai media belajar karena dapat dengan mudah dipahami dan dioperasikan.

Bidang NLP memiliki cakupan yang sangat luas meliputi sektor keuangan, pemasaran, kesehatan, dan juga dapat diimplementasikan pada aplikasi layanan pelanggan. Teknologi layanan *online* merupakan salah satu bentuk aplikasi *Natural Language Processing*. Dengan luasnya bidang cakupan dan penerapan NLP, menjadikan NLP sebagai bidang yang pertumbuhannya sangat pesat yaitu sebesar 22% antara tahun 2020 dan 2030 [10].

NLP merupakan salah satu bidang ilmu Kecerdasan Buatan (*Artificial Intelligence*) yang mempelajari komunikasi antara manusia dengan komputer melalui Bahasa alami. Model komputasi seperti ini berguna untuk memudahkan komunikasi antara manusia dengan komputer dalam hal pencarian informasi kesehatan, sehingga dapat terjadi suatu interaksi antara keduanya dengan menggunakan bahasa alami. Aplikasi ini dirancang agar masyarakat dapat mencari tahu tentang informasi kesehatan berupa deskripsi dari suatu penyakit. Pengguna memasukan pertanyaan berupa *keyword*, kemudian *system* akan menganalisa kemiripan kata kunci pengguna dengan *database* yang ada lalu memberikan respon berupa jawaban atas pertanyaan tersebut.

Dengan luasnya bidang implementasi dari NLP, maka tentu terdapat pula banyak cara atau algoritma yang dapat digunakan, untuk proses *string matching* atau pencocokan kalimat terdapat banyak cara yang tersedia diantaranya dengan menggunakan algoritma *Levenshtein Distance*, *Jaro-Winkler Distance*, *Hamming distance*, *Jaccard Similarity*, dan algoritma-algoritma lainnya.

Pada tahun 2019 terdapat penelitian mengenai efisiensi *fuzzy string similarity* yang ditulis oleh Jin Wang dkk dengan judul “*Efficient Fuzzy String Similarity Join with Multi-level Filtering*” [11]. Penelitian ini bertujuan untuk membandingkan berbagai metode dalam mendeteksi *string matching*, setelah melakukan penelitian dengan cara membandingkan metode *fuzzy string matching*, *jaccard similarity* dan juga *dice similarity* menggunakan dataset yang sama, diperoleh nilai presisi sebesar nilai presisi sebesar: *Fuzzy* : 99%, *Jaccard* : 97% dan, *Dice* : 53%. Dari pengujian tersebut dapat disimpulkan bahwa *Fuzzy similarity* memiliki keunggulan dalam nilai presisi jika dibandingkan dengan dua metode lain jika diuji menggunakan variabel yang sama.

Pengujian perbandingan terhadap algoritma *Levenshtein Distance* dengan *Jaro-Winkler Distance* untuk koreksi kata yang dilakukan oleh M. Adnan Nur didapat hasil bahwa algoritma *Levenshtein Distance* memiliki hasil nilai akurasi 72,40%, *recall* 72,07% dan *f1 score* 80,46%, dimana skor tersebut lebih baik dibandingkan dengan algoritma *Jaro-Winkler* yang memiliki hasil nilai akurasi 70%, *recall* 69,87% dan *f1 score* 79,11%. Dengan hasil ini menunjukkan bahwa algoritma *Levenshtein Distance* lebih optimal digunakan dalam metode koreksi kata [12].

Kemudian Berdasarkan uji coba yang dilakukan pada penelitian rata-rata waktu yang dilakukan oleh Susi Rianti, pencarian terhadap algoritma *Edit Distance* = 47,55 ms, *Levenshtein Distance* = 13,125 ms, *Hamming Distance* = 14,25 ms, dan *Jaccard similarity* = 37,125 ms Ini berarti *Levenshtein Distance* memiliki waktu pencarian tercepat [13].

Dengan Bantuan dari NLP dan dari banyak penelitian sebelumnya pada bidang tersebut, salah satunya dari penelitian yang dilakukan oleh Sugiarto pada tahun 2020, penggunaan metode *fuzzy string matching* khususnya algoritma

*levenshtein* dapat mengantisipasi kesalahan penulisan kata (*typo*) pada *input*, dalam penelitian tersebut *Levenshtein Distance* dapat mengantisipasi *typo* dalam dua kata [14].

Dari latar belakang yang sudah disampaikan, maka dalam penelitian skripsi ini penulis mengambil judul “Implementasi Metode *Rule based* dan *Fuzzy string matching* Pada Aplikasi Konsultasi Kesehatan Berbasis Telegram” dengan harapan penggunaan metode *rule based* dapat menyediakan interaksi antara pengguna dengan program yang dimana pada implementasinya program akan mengarahkan pengguna melalui pertanyaan iya dan tidak sehingga proses menentukan informasi penyakit dari gejala yang dirasakan pengguna menjadi lebih optimal karena proses tersebut mengimitasi proses konsultasi dengan dokter. Kemudian untuk penggunaan *fuzzy string matching* diharapkan bisa menanggulangi berbagai macam *input* yang salah (*typo*) sehingga nantinya *chatbot* akan lebih efektif saat digunakan.

Dengan tersedianya informasi berupa media konsultasi kesehatan yang mudah untuk diakses sehingga dapat meningkatkan tingkat kepedulian kesehatan diri dan masyarakat yang merupakan hasil implementasi dari dua metode yaitu Metode *Rule based* dan *Fuzzy string matching* menjadi pembeda penelitian yang akan dirancang ini dengan penelitian lain.

## 1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini adalah bagaimana merancang sebuah aplikasi layanan masyarakat dengan menggunakan metode *rule based* dan *fuzzy string matching* dengan pendekatan *Natural Language processing* (NLP) dengan deskripsi sebagai berikut:

1. Bagaimana mengimplementasikan metode *rule-based*, *Natural Language processing* (NLP) dan juga *fuzzy string matching* pada *chatbot* sebagai sarana penyedia informasi dan konsultasi kesehatan berupa diagnosis penyakit?
2. Bagaimana tingkat rata-rata akurasi metode *fuzzy string matching*, (menggunakan algoritma *Levenshtein Distance*) yang dihasilkan dalam pencarian gejala penyakit ?

3. Bagaimana efektifitas dan kinerja bot dalam menyampaikan informasi?

### 1.3. Tujuan

Pada Penelitian “Implementasi Metode *Rule based* dan *Fuzzy string matching* Pada Aplikasi Konsultasi Kesehatan Berbasis Telegram” ini bertujuan untuk:

1. Mendapatkan produk berupa *chatbot* yang dapat menyediakan informasi dan konsultasi kesehatan berupa diagnosis penyakit.
2. Mengetahui tingkat rata-rata akurasi algoritma *Levenshtein Distance* dalam membandingkan dua kalimat pada *chatbot* konsultasi kesehatan.
3. Mengetahui tingkat kinerja dan efektifitas waktu eksekusi pencarian data oleh *user* pada *chatbot* konsultasi kesehatan.

### 1.4. Manfaat

Adapun tujuan dari pembuatan sistem konsultasi kesehatan ini yaitu:

- a. Untuk dapat merancang aplikasi yang dapat memprediksi suatu penyakit berdasarkan keluhan pengguna dan memberikan deskripsi tentang penyakit tersebut.
- b. Untuk mengimplementasikan metode *rule based* dan *fuzzy string matching* pada *Chatbot* sebagai konsultasi kesehatan virtual.
- c. Menghadirkan berbagai informasi tentang kesehatan pada media sosial telegram.
- d. Dapat menjadi salah satu program asistensi kesehatan yang gratis dan dapat mudah diakses.
- e. Dapat menjadi referensi untuk penelitian selanjutnya yang berhubungan dengan *Rule based* ataupun *Fuzzy string matching*.

### 1.5. Batasan Masalah

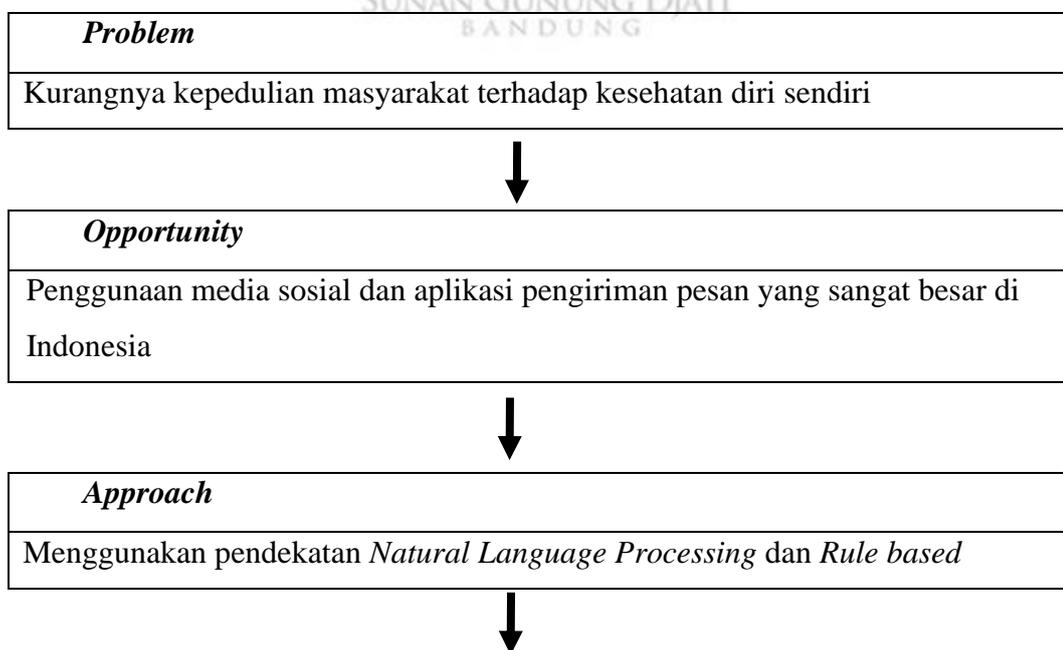
Batasan masalah dari penelitian ini bertujuan agar penelitian ini menjadi lebih terarah dan sesuai dengan tujuan yang diinginkan. Berikut merupakan beberapa batasan masalah dari pembuatan sistem konsultasi kesehatan ini yaitu:

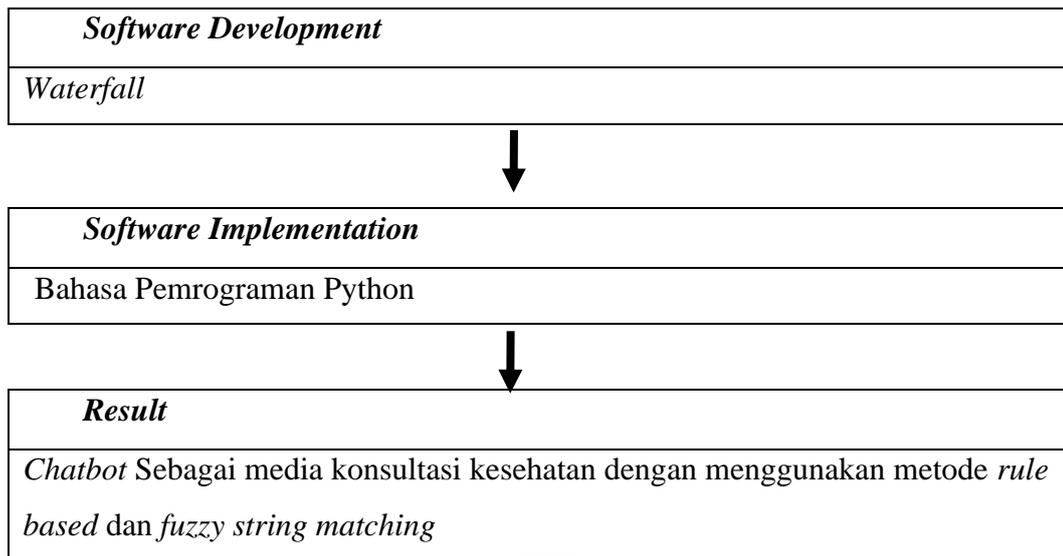
- a. Dataset yang digunakan dalam pengembangan aplikasi ini bersumber dari buku “*The British Medical Association BMA, FAMILY DOCTOR HOME ADVISER*” yang ditulis oleh Dr. Michael Peters [15].

- b. Jawaban yang diberikan aplikasi kepada *user* menggunakan *format text-message*.
- c. Aplikasi berbasis telegram
- d. Bahasa pemrograman yang digunakan adalah Bahasa Python.
- e. Aplikasi tidak dapat merekap identitas berupa jenis kelamin dan, umur.
- f. Dalam percakapan, aplikasi hanya menanggapi *keyword* dalam Bahasa alami dan tidak melayani masukan berupa karakter-karakter khusus.
- g. Aplikasi dapat digunakan bila pengguna terhubung ke internet.
- h. Informasi yang diberikan hanya dibatasi mengenai informasi tentang prediksi penyakit, rekomendasi obat, dan *treatment*.
- i. Aplikasi hanya dapat menjawab pertanyaan yang memiliki kata kunci yang sama dengan kata kunci yang terdapat di dalam dataset.
- j. Menggunakan Algoritma *Levenshtein Distance* untuk mengukur rasio perbedaan kalimat pada metode *Fuzzy string matching*.
- k. Bot bekerja dengan cara memanfaatkan algoritma *Levenshtein Distance* untuk memberikan toleransi terhadap *typo* sehingga dapat melakukan pencarian pada gejala umum, setelah itu akan dilanjutkan dengan metode *rule-based* untuk menggali kemungkinan diagnosanya.

### 1.6. Kerangka Pemikiran

Berikut merupakan kerangka pemikiran dari penelitian tugas akhir ini :





### 1.7. Metodologi Penelitian

#### 1. Tahap pengumpulan data

Pada tahap pengumpulan data ini, penulis menggunakan data sekunder sebagai acuan. Metode pengumpulan data sekunder sering disebut metode penggunaan bahan dokumen, karena dalam hal ini peneliti tidak secara langsung mengambil data sendiri tetapi meneliti dan memanfaatkan data atau dokumen yang dihasilkan oleh pihak-pihak lain. Dalam hal ini, data sekunder merupakan data primer yang diperoleh oleh pihak lain atau data primer yang telah diolah lebih lanjut dan disajikan baik oleh pengumpul data primer atau oleh pihak lain (buku “Pertolongan Pertama Dokter di Rumah Anda” yang disunting oleh Dr. Tony Smilth.) yang pada umumnya disajikan dalam bentuk tabel-tabel atau diagram-diagram.

#### 2. Tahap pengembangan sistem

Untuk tahapan pengembangan sistem terdapat banyak sekali metode yang dapat digunakan, metode pengembangan aplikasi yang diterapkan pada penelitian ini adalah Metode *Waterfall*.

### 1.8. Sistematika Penulisan

Berikut merupakan sistematika penulisan pada penelitian ini yang disusun ke dalam lima bab :

## **BAB I PENDAHULUAN**

Bab pendahuluan memuat Latar Belakang, Rumusan Masalah, Tujuan dan Manfaat, Batasan Masalah, Kerangka Pemikiran, Metodologi Penelitian, dan Sistematika Penulisan.

## **BAB II STUDI PUSTAKA**

Bab studi pustaka berisi landasan teori yang menyokong penelitian sehingga menjadi terarah.

## **BAB III METODOLOGI PENELITIAN**

Bab perancangan berisikan uraian perancangan sistem yang akan dibuat mulai dari pemahaman data, pengumpulan data dan proses pengolahan data dari “Implementasi Metode *Rule based* dan *Fuzzy string matching* Pada Aplikasi Konsultasi Kesehatan Berbasis Telegram”.

## **BAB IV HASIL DAN PEMBAHASAN**

Bab implementasi berisi hasil dari sistem yang telah dirancang yang kemudian dievaluasi.

## **BAB V SIMPULAN DAN SARAN**

Bab penutup adalah tahapan akhir yang isinya adalah kesimpulan dari penelitian ini.

## **REFERENSI**

Daftar pustaka berisikan sumber-sumber tertulis yang dipakai menjadi acuan pada penelitian ini.