

BAB I

PENDAHULUAN

1.1. Latar Belakang

Anamnesis merupakan pemeriksaan tahap awal terhadap pasien yang dilakukan lewat suatu percakapan oleh dokter baik secara langsung atau melalui orang terdekat pasien untuk penegakan diagnostik. Tujuan anamnesis adalah untuk memperoleh informasi mengenai pasien agar dokter mampu menyimpulkan dugaan organ atau sistem organ yang terganggu beserta rumusan masalah kliniknya, serta untuk menjalin hubungan yang baik antara dokter dengan pasiennya (Markum, 2007). Ada dua jenis anamnesis, yaitu autoanamnesis dan alloanamnesis atau heteroanamnesis. Autoanamnesis adalah anamnesis yang dilakukan langsung terhadap pasien, sedangkan alloanamnesis atau heteroanamnesis adalah anamnesis yang dilakukan terhadap orang lain yang mengenal pasien. Alloanamnesis atau heteroanamnesis biasa dilakukan jika pasien dianggap kurang mampu melakukan wawancara seperti pasien yang pingsan, bisu, sakit jiwa, cacat mental, maupun pasien anak-anak. Terkadang autoanamnesis bisa dilakukan bersama-sama dengan alloanamnesis juga. Data anamnesis yang berisi identitas pasien, riwayat penyakit pasien, keluhan utama pasien, riwayat penyakit keluarga pasien, serta informasi pasien ini nantinya disimpan di dalam berkas rekam medis (Kementerian Kesehatan Republik Indonesia, 2008).

Rekam medis elektronik (RME) adalah sistem informasi komputer untuk layanan kesehatan yang berisi data medis dan data sosial pasien (Andriani dkk., 2017). Dibandingkan rekam medis berbasis kertas, rekam medis elektronik lebih mudah untuk diolah oleh komputer karena datanya tersimpan dalam format teks *digital* yang dapat dibaca langsung oleh komputer. Rekam medis elektronik juga tidak memiliki kelemahan-kelemahan yang terdapat pada rekam medis kertas. Kelemahan-kelemahan yang dimaksud antara-lain seperti mudah rusak atau hilang, membutuhkan ruangan besar untuk penyimpanannya, dan tulisannya sulit dibaca oleh sebagian orang (Ajami & Bagheri-Tadi, 2013). Data teks keluhan pada rekam medis elektronik umumnya ditulis menggunakan format penulisan bebas atau format penulisan tidak terstruktur yang mempersulit proses penggalian teks pada rekam medis elektronik secara otomatis menggunakan komputer (Pereira dkk., 2015), padahal banyak informasi penting mengenai pasien yang tersimpan pada rekam medis elektronik tersebut yang dapat digunakan sebagai masukan untuk Sistem

Pendukung Keputusan Klinis. Pemeriksaan data rekam medis elektronik secara manual oleh manusia juga sulit dilakukan karena jumlah data rekam medis elektronik yang sangat banyak.

Penelitian ini bertujuan untuk mengembangkan sebuah aplikasi yang mampu menerjemahkan data keluhan pasien yang tersimpan pada rekam medis elektronik ke dalam format teks yang lebih baku dan terstruktur agar teks tersebut dapat lebih mudah dibaca dan diolah oleh komputer. Aplikasi yang dikembangkan pada penelitian ini menggunakan metode-metode Pengolahan Bahasa Alami (Pengolahan Bahasa Alami) dalam melakukan hal tersebut. Pengolahan Bahasa Alami adalah salah satu cabang dari Informatika yang berkaitan dengan pemrosesan cerdas pada bahasa manusia (Ford dkk., 2016). Metode yang digunakan dalam penerjemahan teks keluhan pasien pada aplikasi ini adalah dengan menggunakan metode statistik, dengan metode pencocokan kata-kata atau frasa-frasa kunci, maupun dengan contoh penerjemahan yang diberikan oleh *user*.

Aplikasi yang dikembangkan pada penelitian ini dirancang agar mampu mengenali masukan suara dari pasien maupun petugas kesehatan menggunakan fitur *speech recognition* yang dikembangkan menggunakan Sphinx4 API. Masukan data keluhan pasien selain berupa teks juga dapat dalam bentuk suara yang dimasukkan langsung menggunakan mikrofon maupun dengan mengimpor rekaman suara yang tersimpan di dalam media penyimpanan. *Speech recognition* pada Sphinx4 API menggunakan pendekatan probabilistik. Pada *speech recognition* menggunakan pendekatan pendekatan probabilistik, setiap *input* suara yang dimasukkan bersesuaian terhadap kata atau kumpulan kata tertentu pada kamus dengan sebuah probabilitas tertentu, hasil transkripsi dari *input* suara tersebut adalah kata atau kumpulan kata yang memiliki probabilitas tertinggi (Gruhn dkk., 2011).

1.2. Rumusan Masalah

Berdasarkan latar belakang dari penelitian ini, maka rumusan masalah penelitian ini adalah bagaimana mengembangkan aplikasi yang mampu untuk:

1. Menerjemahkan keluhan pasien dalam format yang kurang baku dan kurang terstruktur menjadi lebih baku dan terstruktur?
2. Mengenali *input* suara pasien atau petugas kesehatan sebagai masukannya?
3. Dilatih agar dapat menerjemahkan keluhan pasien dan mengenali suara pasien atau petugas kesehatan secara lebih akurat?

1.3. Batasan Masalah

Batasan masalah dalam pengembangan aplikasi pada penelitian ini adalah:

1. Aplikasi ini hanya mampu mengenali 600 macam keluhan pasien yang paling umum saja. Hal ini disebabkan karena keterbatasan waktu dan biaya yang digunakan untuk penelitian. Namun *user* dapat melatih aplikasi agar mampu mengenali keluhan-keluhan pasien lainnya.
2. Fungsi *speech recognition* menggunakan Sphinx4 pada aplikasi ini hanya dilatih menggunakan 992 kata atau frasa dan dengan 2 jenis suara saja karena keterbatasan waktu dan biaya. Namun *user* dapat melatih aplikasi agar mampu mengenali suara dengan lebih baik.

1.4. Tujuan Penelitian

Tujuan dari penelitian ini adalah mengembangkan sebuah aplikasi yang mampu untuk:

1. Menerjemahkan keluhan pasien dalam format yang kurang baku dan terstruktur menjadi bentuk yang lebih baku dan terstruktur.
2. Mengenali *input* suara pasien atau petugas kesehatan sebagai masukannya.
3. Dilatih untuk dapat menerjemahkan keluhan pasien dan mengenali suara pasien atau petugas kesehatan dengan lebih akurat

1.5. Manfaat Penelitian

Beberapa manfaat yang dapat diperoleh dari penelitian ini adalah sebagai berikut:

1. Memudahkan pengolahan data pada Sistem Pendukung Keputusan Klinis yang menggunakan data-data keluhan pasien sebagai masukannya.
2. Memudahkan petugas kesehatan dalam menuliskan data keluhan pasien pada rekam medis elektronik sehingga petugas kesehatan tidak perlu menyesuaikan penulisan data keluhan dengan format yang baku.
3. Memudahkan pasien atau petugas kesehatan yang memiliki keterbatasan fisik untuk memasukkan data keluhan pasien cukup dengan menggunakan suara mereka.

1.6. Metodologi Penelitian

Dalam penelitian ini, langkah pertama yang dilakukan adalah melakukan studi literatur. Pada langkah studi literatur ini dipelajari ilmu-ilmu kedokteran dasar dan teori-teori mengenai Pengolahan Bahasa Alami yang nantinya digunakan dalam pengembangan aplikasi. Ilmu-ilmu kedokteran dasar dan teori-teori mengenai Pengolahan Bahasa Alami dipelajari baik dari buku-buku maupun dari internet.

Setelah melakukan studi literatur, dilakukan perancangan aplikasi. Dibuat rancangan tampilan antarmuka dan rancangan fitur-fitur apa saja yang terdapat pada aplikasi yang dibuat. Bahasa Pemrograman Java diputuskan digunakan dalam pengembangan aplikasi karena aplikasi yang dibuat menggunakan Bahasa Pemrograman Java bersifat portabel dan karena Sphinx4 API yang digunakan untuk pengembangan fitur *speech recognition* pada aplikasi ini juga dikembangkan dengan Bahasa Pemrograman Java. Untuk pembuatan desain antarmuka, disewa seorang desainer grafis dari Oceano Design Yogyakarta untuk menggambarkan desain latar belakang pada jendela utama aplikasi. Pada langkah ini juga dicatat daftar data-data yang dibutuhkan untuk penelitian dan bagaimana cara memperolehnya.

Setelah mencatat daftar data-data yang dibutuhkan untuk penelitian, kemudian dilakukan langkah pengumpulan data untuk pengembangan aplikasi. Data-data yang dibutuhkan antara lain data daftar kata-kata yang umum digunakan dalam teks Bahasa Indonesia dan daftar kata-kata yang umum digunakan oleh pasien dalam menulis keluhanannya beserta daftar sinonimnya yang nantinya digunakan untuk pengembangan fitur pengoreksi ejaan dan penerjemahan, daftar keluhan yang umum dialami oleh pasien beserta terjemahan istilah medisnya, lokasinya, dan sistem organnya yang nantinya digunakan untuk proses penerjemahan keluhan pasien, serta kamus fonetis, model bahasa, dan model akustik yang digunakan untuk pengembangan fitur *speech recognition*. Daftar kata-kata yang umum digunakan dalam teks Bahasa Indonesia dan daftar kata-kata yang umum digunakan oleh pasien dalam menulis keluhanannya didapatkan dari penggalan teks pada situs Wikipedia Indonesia dan beberapa situs-situs konsultasi kedokteran seperti klikdokter.com, alodokter.com, dokter.id, dll. serta untuk daftar sinonimnya didapatkan dari kamus-kamus Tesaurus baik kamus cetak maupun *online*. Tidak semua kata dapat diperoleh daftar sinonimnya karena keterbatasan waktu dan biaya dalam penelitian. Daftar keluhan yang umum dialami oleh pasien beserta terjemahan istilah medisnya, lokasinya, dan sistem organnya, didapatkan dari situs-situs konsultasi kedokteran, buku-buku

kedokteran dasar, dan dari bertanya kepada beberapa dokter beserta dokter muda dan petugas kesehatan di Rumah Sakit Sardjito Yogyakarta. Data model akustik didapatkan dari merekam suara penulis dan dengan menggunakan fitur *speech synthesis* pada *Google Translate*. Aplikasi hanya dilatih dengan 992 kata atau kumpulan kata yang umum karena keterbatasan waktu dan biaya untuk pembuatan aplikasi, sehingga data kamus fonetis pun hanya berisi 992 kata atau kumpulan kata saja. Data model bahasa dibuat dari daftar kata-kata pada kamus fonetis beserta kombinasi bigram, trigram, maupun 4-gram yang umum dari kata-kata tersebut yang didapatkan dari penggalan teks pada situs health.detik.com serta tambahan dari data-data hasil eksperimen yang dibiarkan tersimpan pada model bahasa.

Setelah data-data tersebut dikumpulkan kemudian dilanjutkan ke langkah perbaikan data. Langkah ini dilakukan untuk memperbaiki data-data yang sudah dikumpulkan seperti perbaikan kesalahan ejaan, pengubahan singkatan-singkatan yang tidak baku ke bentuk kepanjangannya, pengabaian kata-kata dalam bahasa asing, pemilihan rekaman suara yang bagus, dan perbaikan rekaman suara dengan memotong bagian awal dan akhir dari rekaman suara tersebut yang hanya berisi *noise* atau *silence* untuk pembuatan model akustik.

Setelah data-data berhasil dikumpulkan dan diperbaiki, dilanjutkan ke langkah pengembangan aplikasi. Aplikasi dikembangkan menggunakan Bahasa Pemrograman Java. Untuk fitur *speech recognition* digunakan Sphinx4 API yang bersifat *open source*. Untuk fitur *speech synthesis* digunakan aplikasi *open source* eSpeak. Aplikasi *open source* FFmpeg digunakan untuk mengonversi masukan suara langsung dari *user* ke format yang dapat diproses oleh Sphinx4 untuk proses *speech recognition*. Setelah aplikasi berhasil dikembangkan, dilakukan pengujian terhadap aplikasi tersebut oleh penulis dan dilakukan perbaikan pada *source code* atau pada data-datanya, seperti perbaikan pada isi kamus-kamus yang digunakan oleh aplikasi untuk proses penerjemahan dan *speech recognition*, dan perbaikan *bug* serta algoritma yang kurang efisien pada *source code*. Setelah itu lima orang dokter diminta untuk mengujicoba aplikasi tersebut dan memberikan penilaian mereka pada kuesioner terhadap aspek kemudahan penggunaan aplikasi tersebut, tampilan antarmuka, tingkat akurasi dan kecepatan pemrosesan fitur penerjemahan, tingkat akurasi dan kecepatan pemrosesan fitur *speech recognition*, serta kemampuan aplikasi ini untuk membantu petugas kesehatan. Kelima dokter tersebut juga diminta untuk memberikan saran mereka bagi pengembangan aplikasi tersebut pada kuesioner.

1.7. Sistematika Penulisan

Berikut adalah sistematika penulisan dan pengelompokan tentang uraian isi dari laporan:

BAB I PENDAHULUAN

Bab ini membahas tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, metodologi penelitian, dan sistematika penulisan laporan.

BAB II TINJAUAN PUSTAKA

Bab ini membahas tentang penelitian terdahulu yang sudah pernah dilakukan, pengertian anamnesis, pengertian bahasa alami, dan pengertian pengolahan bahasa alami.

BAB III METODOLOGI PENELITIAN

Bab ini membahas mengenai data-data yang dibutuhkan dalam penelitian ini serta tahapan-tahapan penelitian yang dilakukan untuk pengembangan aplikasi ini.

BAB IV PEMODELAN SISTEM

Bab ini menjelaskan mengenai rancangan model aplikasi yang dikembangkan yang antara lain berupa gambaran model aplikasi secara keseluruhan, model algoritma pembelajaran untuk penerjemahan, model algoritma penerjemahan menggunakan pembelajaran berdasarkan statistik, dan model algoritma pembelajaran untuk *speech recognition*.

BAB V HASIL DAN PEMBAHASAN

Bab ini berisi penjelasan antarmuka aplikasi dan hasil pengujian dari aplikasi yang telah berhasil dikembangkan beserta evaluasi dan pembahasan dari hasil pengujian tersebut.

BAB VI KESIMPULAN DAN SARAN

Berisi kesimpulan yang diperoleh dari hasil penelitian beserta saran-saran yang bermanfaat untuk penelitian yang akan datang.