BAB III
METODOLOGI PENELITIAN

3.1. Pengambilan Data

Untuk mengembangkan aplikasi pada penelitian ini beberapa data yang dibutuhkan adalah sebagai berikut:

1. Data keluhan pasien yang umum beserta istilah medis, lokasi, serta sistem organnya.
2. Data kata-kata yang umum pada teks Bahasa Indonesia dan kata-kata yang umum pada penulisan keluhan pasien beserta sinonimnya.
3. Data model akustik, kamus fonetis, dan model bahasa untuk *speech recognition.*

Data keluhan pasien yang umum beserta istilah medis, lokasi, serta sistem organnya digunakan untuk proses penerjemahan pada aplikasi. Data keluhan pasien yang umum diperolah dari beberapa situs konsultasi kedokteran *online* seperti alodokter.com, dokter.id, klikdokter.com, dan tanyadok.com. Untuk istilah medis, lokasi, dan sistem organ dari keluhan tersebut diperoleh dari membaca-baca buku kedokteran dasar, internet, dan bertanya kepada beberapa dokter dan petugas kesehatan di Rumah Sakit Sardjito Yogyakarta.

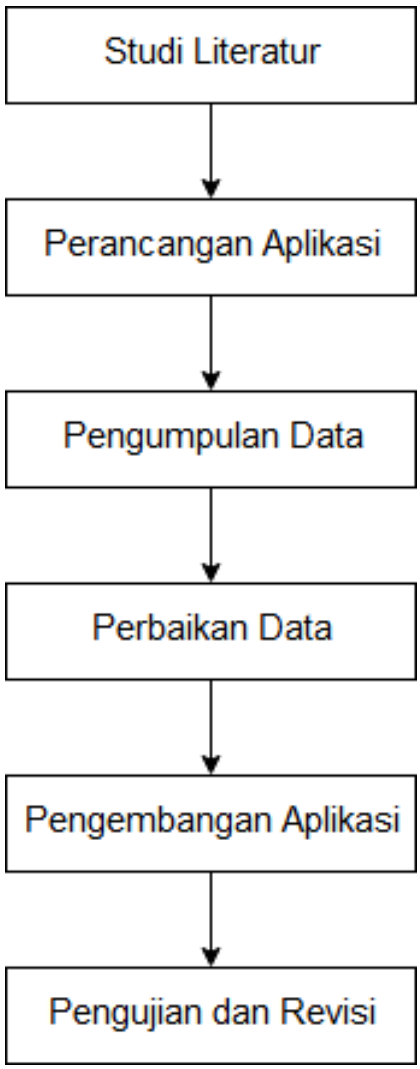
Data kata-kata yang umum pada teks Bahasa Indonesia dan kata-kata yang umum pada penulisan keluhan pasien beserta sinonimnya. Data kata-kata tersebut diperoleh dari penggalian teks pada situs Wikipedia dan beberapa situs konsultasi kedokteran *online* seperti alodokter.com, dokter.id, klikdokter.com, dan tanyadok.com. Untuk sinonimnya diperoleh dari kamus tesaurus Bahasa Indonesia *online* maupun cetak.

Data model akustik, kamus fonetis, dan model bahasa digunakan sebagai basis pengetahuan untuk proses *speech recognition*. Data model akustik diperoleh dari merekam contoh suara oleh penulis dan contoh suara yang dihasilkan oleh fitur speech synthesis pada Google Translate. Hanya 992 kata dan kumpulan kata umum yang dapat diambil contoh suaranya karena keterbatasan waktu dan biaya untuk penelitian, sehingga kamus fonetis hanya berisi 992 kata dan kumpulan kata saja. Data model bahasa berisi 992 kata dan kumpulan kata pada kamus fonetis tersebut termasuk kombinasi 2-gram, 3-gram,

maupun 4-gram yang umum dari kata-kata dan kumpulan-kumpulan kata tersebut yang diperoleh dari penggalian teks pada *website* health.detik.com beserta data-data untuk eksperimen yang sengaja tetap disimpan pada model bahasa.

3.2. Tahapan Penelitian

Gambar 3.1. di bawah menunjukkan alur tahapan yang dilakukan dalam penelitian ini. Berdasarkan gambar 3.1. di bawah, alur tahapan dalam penelitian ini adalah studi literatur, perancangan aplikasi, pengumpulan data, pengolahan data, pengembangan aplikasi, dan diakhiri dengan pengujian dan revisi pada aplikasi. Tiap-tiap tahapan tersebut dijelaskan pada subbab-subbab berikutnya.



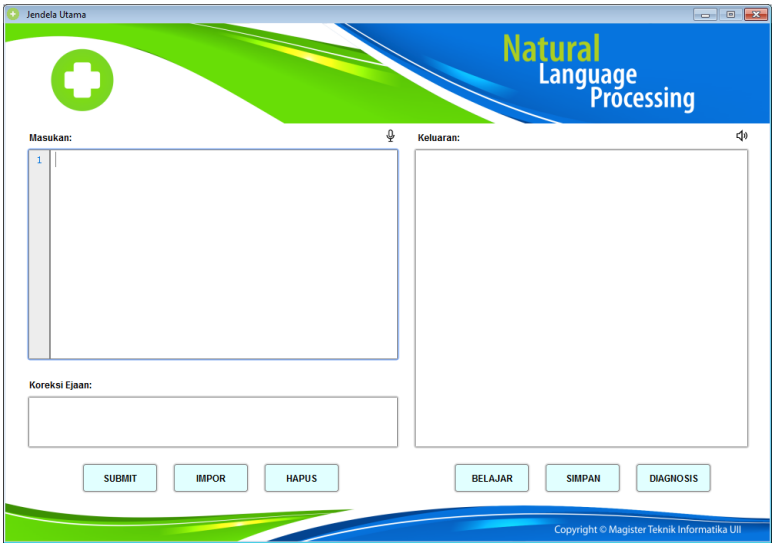
Gambar 3.1. Bagan alur penelitan.

3.2.1. Studi Literatur

Pada tahap studi literatur ini dipelajari beberapa literatur baik dari media *online* maupun media cetak mengenai teori-teori dan aplikasi-aplikasi pemrosesan bahasa alami yang sudah pernah dibuat terutama mengenai *speech recognition*, *speech synthesis*, dan *statistical machine translation*. Pada tahap ini dipelajari juga ilmu kedokteran dasar dari beberapa media *online* dan media cetak agar lebih memahami penyakit-penyakit yang umum diderita di masyarakat

3.2.2. Perancangan Aplikasi

Setelah dilakukan studi literatur, kemudian dilakukan perancangan aplikasi yang ingin dikembangkan. Pada tahapan perancangan aplikasi, dibuat rancangan antarmuka, rancangan algoritma, dan rencana data-data yang dibutuhkan dalam pengembangan aplikasi. Rancangan tampilan antarmuka utama aplikasi yang dikembangkan dapat dilihat pada gambar 3.2. di bawah ini.



Gambar 3.2. Rancangan tampilan antarmuka utama aplikasi.

3.2.3. Pengumpulan Data

Setelah rancangan aplikasi selesai, kemudian dikumpulkan data-data yang dibutuhkan untuk pengembangan aplikasi. Data-data tersebut antara lain adalah data keluhan pasien yang umum yang didapatkan dari situs-situs konsultasi kedokteran *online*,

data kata-kata Bahasa Indonesia yang paling sering digunakan, data model bahasa yang diperoleh dari penggalian teks pada Wikipedia Indonesia dan beberapa situs-situs kedokteran, dan data model akustik yang didapatkan dengan merekam suara penulis sendiri dan dengan menggunakan aplikasi *speech synthesizer* yang ada pada Google Translate.

3.2.4. Perbaikan Data

Data yang sudah dikumpulkan kemudian diperbaiki terlebih dahulu sebelum digunakan oleh aplikasi. Data-data keluhan pasien yang sama persis dengan data-data keluhan pasien lainnya dihapus. Pada daftar kata-kata yang umum pada teks Bahasa Indonesia dan daftar kata-kata yang umum pada penulisan keluhan pasien, kata-kata yang bukan kata Bahasa Indonesia, seperti kata Bahasa Inggris, diabaikan. Data-data suara untuk model akustik dipilih yang paling baik dan dipotong bagian awal dan akhirnya yang hanya berisi *silence* atau *noise*.

3.2.5. Pengembangan Aplikasi

Pada tahap ini aplikasi dikembangkan menggunakan Bahasa Java. Bahasa Java dipilih karena Sphinx4 yang digunakan sebagai fitur *speech recognition* pada aplikasi ini dikembangkan menggunakan Bahasa Java. Bahasa Java juga menghasilkan *file executable* yang memiliki performa cukup baik dan aplikasi yang dihasilkannya bersifat *portable*.

3.2.6. Pengujian dan Revisi

Setelah aplikasi berhasil dikembangkan, aplikasi kemudian diujicoba oleh penulis. Dilakukan pengujian akurasi dan performa penerjemahan pada aplikasi ini dengan memasukkan *input* keluhan sebanyak 30 kalimat secara acak serta pengujian akurasi dan performa fitur *speech recognition* pada aplikasi dengan memasukkan *input* suara sebanyak 30 kali secara acak dengan lingkungan suara sekitar yang berbeda-beda. Jika didapatkan hasil yang kurang memuaskan maka dilakukan revisi baik pada kode program maupun data-data pada program. Setelah proses pengujian oleh penulis dan revisi selesai dilakukan, lima orang dokter masing-masing diminta untuk memasukkan keluhan secara acak untuk 10 penyakit pada aplikasi dan hasil terjemahan oleh aplikasi dicek oleh dokter yang bersangkutan apakah sesuai atau tidak, dari hasil tersebut dihitung tingkat akurasi dan

performa penerjemahan oleh aplikasi. Kelima orang dokter tersebut masing-masing juga diminta untuk mengetes fitur *speech recognition* dengan menggunakan masukan suara mereka sendiri sebanyak 5 kali untuk mengecek tingkat akurasi dan performa fitur *speech recognition* pada aplikasi ini. Setelah itu kelima dokter yang bersangkutan diminta untuk mengisi kuesioner untuk mengetahui penilaian dan saran mereka mengenai aplikasi yang dikembangkan ini. Tampilan rancangan kuesioner yang diberikan kepada dokter-dokter yang mengujicoba aplikasi dapat dilihat pada gambar 3.3. di bawah ini.

Kuesioner Penilaian Program

Nama Dokter :

Tanggal :

	Sangat Tidak Setuju	Tidak Setuju	Ragu-Ragu	Setuju	Sangat Setuju
Program ini mudah dipelajari dan digunakan					
Langkah-langkah pengoperasian program ini mudah diingat					
Program ini memiliki tampilan antarmuka yang cukup bagus					
Tingkat akurasi pengenalan suara pada program ini sudah cukup baik					
Tingkat akurasi penerjemahan program ini sudah cukup baik					
Kecepatan program ini dalam mengenali suara sudah cukup baik					
Kecepatan program ini dalam menerjemahkan masukan sudah cukup baik					
Program ini dapat membantu pekerjaan petugas kesehatan					
Apakah ada saran bagi pengembangan program?					

Tanda Tangan Dokter

( )

Gambar 3.3. Rancangan kuesioner.