BAB 1
PENDAHULUAN

1.1 Latar Belakang

Banyak kegiatan yang dapat dilakukan untuk mendapatkan informasi yang diinginkan. Salah satu cara untuk mendapatkan informasi tersebut adalah dengan bertanya kepada orang yang lebih mengetahui akan hal yang ditanyakan. Orang yang diberi pertanyaan akan menjawab dan memberikan informasi yang relevan. Kegiatan tersebut dinamakan kegiatan tanya jawab.

Sistem Tanya Jawab (QA) merupakan sebuah sistem pada komputer yang berisi sebuah cara otomatis untuk menjawab pertanyaan menggunakan bahasa alami yang biasa dipakai manusia. Jawaban ini berisi informasi yang berasal dari sebuah sumber basis data [1]. Sebuah sistem tanya jawab, diharapkan dapat membantu orang untuk mendapatkan informasi mengenai hal tertentu tanpa harus bertanya kepada orang lain.

Penelitian mengenai sistem tanya jawab sudah banyak dilakukan. Sangdo Han dkk, pada penelitiannya yang berjudul *Keyword Question Answering System with Report Generation for Linked Data* (2015) menggunakan metode *Keyword Matching* untuk membangun sebuah sistem tanya jawab. Penelitian ini menghasilkan akurasi 95,1% dari 670 pertanyaan yang diajukan oleh 12 orang [2]. Kekurangan dari penelitian tersebut adalah, *keyword* atau kata kunci yang diproses hanya berupa kata kunci yang terdiri dari satu buah kata saja, sementara dalam aturan tata Bahasa Indonesia, sebuah pertanyaan memiliki kata tanya, kata kunci, dan keterangan yang dapat memiliki lebih dari satu buah kata.

Sementara salah satu penelitian mengenai sistem yang menggunakan *rule-based* dilakukan oleh Ivan Karl Bindoff yang menerapkan metode *rule-based Multiple Classification Ripple Down Rules* (MCRDR) pada penelitiannya yang berjudul *An Intelligent Decision Support System for Automated Medication Review* (2005). Pada penelitian tersebut, *rule-based* MCRDR digunakan untuk mengklasifikasikan resep obat dengan data masukan berupa keluhan-keluhan dari

pasien. Hasil penelitian tersebut menunjukkan 60% nilai akurasi dalam klasifikasi resep obat. Penelitian tersebut menyatakan bahwa dengan menambahkan kondisi pada *rules* yang dibuat, akurasi dapat ditingkatkan secara signifikan [3].

Banyak penelitian lain mengenai Sistem Tanya Jawab yang dipublikasikan pada jurnal yang ditulis oleh Ashish Mathur dan M.T.U Haider pada tahun 2015 [4], diantaranya,

- Moldovan (1999) dan Harabagiu (2000) menggunakan WordNet untuk memproses arti semantik dari sebuah pertanyaan. Hasil akurasi terbesar dari penelitian tersebut adalah 77,7%.
- Gaizauskas & Humphreys (2000) menggabungkan *Information Retreival* dengan *Natural Language Processing* yang menghasilkan akurasi terbesar sebesar 53,33%.
- Zhang & Lee (2003) meneliti beberapa metode *machine learning*, seperti *Support Vector Machine* (SVM), *Nearest Neighbors* (NN), *Naive Bayes* (NB), *Decision Tree* (DT), dan *Sparse Network of Winnows* (SNoW). Penelitian yang mereka lakukan menghasilkan akurasi terbesar sebesar 75,6% untuk NN, 77,4% untuk NB, 84,2% untuk DT, 78,2% untuk SNoW, dan 87,4% untuk SVM.
- Stoyanchev, dkk (2008) menggunakan *Part-of-speech Tagger, Named Entity Recognition* dan *Stop Word List* pada sistem tanya jawab, yang menghasilkan akurasi tertinggi sebesar 65,3%.

Semua penelitian tersebut menggunakan bahasa Inggris sebagai data masukannya. Kedua penelitian yang dilakukan oleh Sangdo Han, dkk, dan Ivan Karl Bindoff mengenai *keyword matching* dan MCRDR akan menjadi dasar untuk menerapkan sistem tanya jawab menggunakan Bahasa Indonesia. Berdasarkan semua penelitian yang sudah dilakukan, metode *keyword matching* memiliki akurasi tertinggi, dan MCRDR akan memeriksa *keyword* yang ada dalam sebuah kalimat satu persatu secara berurutan. Sehingga diharapkan penerapan kedua metode tersebut dapat menghasilkan jawaban dari pertanyaan dengan nilai akurasi yang tinggi.

Sistem Tanya Jawab dapat diterapkan pada *front office* untuk menangani pertanyaan yang diajukan pengunjung. Contohnya, pada *front office* UNIKOM, pengunjung sering bertanya mengenai informasi pendaftaran atau penerimaan mahasiswa baru. Pengunjung dapat menanyakan hal-hal sederhana seperti besarnya biaya yang harus dibayarkan, kapan dimulainya ujian saringan masuk, kapan dimulainya yang seharusnya dapat dijawab dengan cepat. Namun dengan banyaknya pengunjung yang menanyakan hal tersebut, waktu yang dibutuhkan untuk melayani akan semakin lama. Sehingga penerapan metode *keyword matching* dan MCRDR dalam sistem tanya jawab akan dibuatkan dalam bentuk *prototype* untuk mengetahui nilai akurasi jawaban sebelum sistem dapat diterapkan oleh *front office*.

1.2 Rumusan Masalah

Identifikasi masalah yang dapat dirumuskan sesuai latar belakang yang dijelaskan sebelumnya adalah akurasi sistem tanya jawab yang menggabungkan metode *keyword-matching* dengan metode *rule-based* MCRDR dalam bahasa Indonesia belum bisa diukur karena belum adanya penelitian yang menggabungkan kedua metode tersebut.

1.3 Maksud

Maksud dari penelitian ini adalah membangun prototype aplikasi tanya jawab (*question & answering*) yang menggunakan metode *rule-based* MCRDR dan *keyword-matching* untuk menerima pertanyaan dari pengunjung dan memberikan jawaban atau informasi yang relevan dengan pertanyaan tersebut.

1.4 Tujuan

Sedangkan tujuan dari penelitian ini untuk mengukur akurasi sistem tanya jawab yang menggabungkan metode *keyword-matching* dengan *rule-based* MCRDR dalam Bahasa Indonesia.

1.5 Batasan Masalah

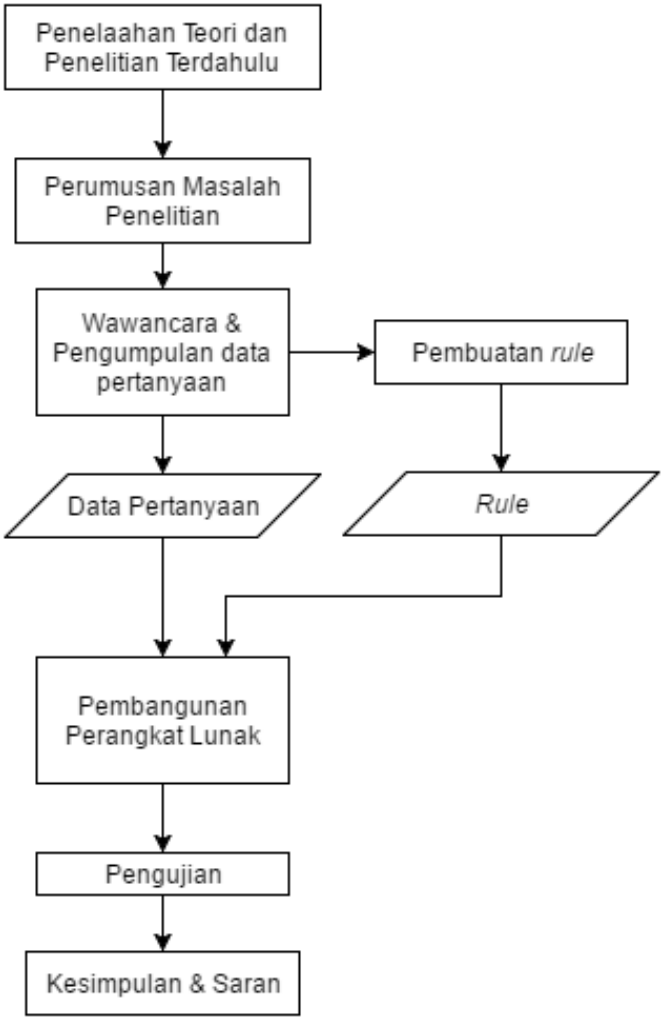
Adapun batasan masalah dalam pembangunan aplikasi ini adalah sebagai berikut.

1. Data Masukan :
   - Pertanyaan dalam Bahasa Indonesia.
   - Cakupan tema pertanyaan adalah informasi-informasi mengenai pendaftaran dan pembayaran kuliah pada kampus UNIKOM.
   - Tema pertanyaan yang diajukan mengenai pendaftaran contohnya adalah pertanyaan mengenai kapan dimulainya pendaftaran, berapa biaya pendaftaran, dan informasi alamat lokasi pendaftaran.
   - Pertanyaan setidaknya harus memiliki minimal satu buah kata tanya dan satu kata kunci.
   - Pengumpulan pertanyaan dihasilkan dari wawancara pada pengunjung *front office* di UNIKOM.
2. Proses Pengolahan Data :
   - *Emoticon* dan semua simbol tanda baca terkecuali spasi dan *dash* (-) tidak akan diproses oleh sistem.
   - Sistem akan memberikan *output* atau keluaran terhadap pengguna berupa jawaban yang akan dihitung relevansinya.
3. Data Keluaran :
   - Keluaran dari penelitian ini adalah nilai akurasi dari *prototype* sistem tanya jawab.

1.6 Metodologi Penelitian

Metode penelitian yang digunakan adalah metode deskriptif. Metode deskriptif adalah suatu metode dalam penelitian untuk membuat gambaran mengenai situasi atau kejadian berdasarkan pencarian fakta dengan interpretasi yang tepat. Metode penelitian ini memiliki dua tahapan, yaitu pengumpulan data, dan pembangunan perangkat lunak.

Metode ini dipilih karena sesuai dengan penelitian ini, dimana *prototype* sistem yang akan dibangun harus bisa menggambarkan situasi atau kejadian proses tanya-jawab berdasarkan fakta-fakta yang didapatkan dari tahapan pengumpulan data.



Gambar 1.1 Skema Metode Penelitian Deskriptif

1.6.1 Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini adalah dengan menggunakan metode studi literatur, kuesioner dan metode wawancara.

Studi literatur adalah pengumpulan data dengan cara mempelajari sumber kepustakaan, yang berupa hasil penelitian, buku referensi, dan jurnal – jurnal yang terkait dengan penelitian mengenai sistem tanya jawab dan metode *rule-based*.

Menurut Sugiyono (2008), kuesioner adalah teknik pengumpulan data yang berisi pertanyaan tentang suatu penelitian untuk dijawab oleh responden [5]. Pada penelitian ini kuesioner digunakan untuk mendapatkan data-data contoh pertanyaan yang diajukan pada bagian *front office* UNIKOM.

Wawancara adalah teknik pengumpulan data yang digunakan peneliti untuk mendapatkan keterangan-keterangan lisan dengan melakukan tanya jawab terhadap petugas *front office* di UNIKOM mengenai situasi atau keadaan tanya jawab yang terjadi.

1.6.2 Pembangunan Perangkat Lunak

Metode pengembangan perangkat lunak yang digunakan adalah metode *prototyping.* Metode ini digunakan dalam penelitian ini karena memungkinkan untuk kembali ke tahapan analisis apabila ada kesalahan yang terjadi pada saat implementasi perangkat lunak dilakukan. [6] Adapun tahapan pada metode ini adalah sebagai berikut.

1. Pengumpulan Informasi

   Mendefinisikan format seluruh perangkat lunak, mengidentifikasikan semua kebutuhan dan garis besar sistem yang akan dibuat, serta melakukan studi literatur terhadap penelitian-penelitan sejenis. Pada tahapan ini dilakukan pengumpulan jurnal-jurnal mengenai sistem tanya jawab dan metode *rule-based*, serta mempersiapkan perangkat lunak yang akan dipakai untuk membangun *prototype* system.
2. Desain *Prototype*

Tahap ini adalah tahap untuk merancang dan mendesain sistem yang akan dibuat, termasuk rancangan antar muka, *flowchart*, dan desain basis data yang disesuaikan dengan *prototype* sistem tanya jawab yang akan dibangun.

3. Pembangunan *Prototype*

   Desain yang sudah dibuat pada tahap 2 diimplementasikan atau dikodekan pada sistem *prototype* yang akan dibangun.

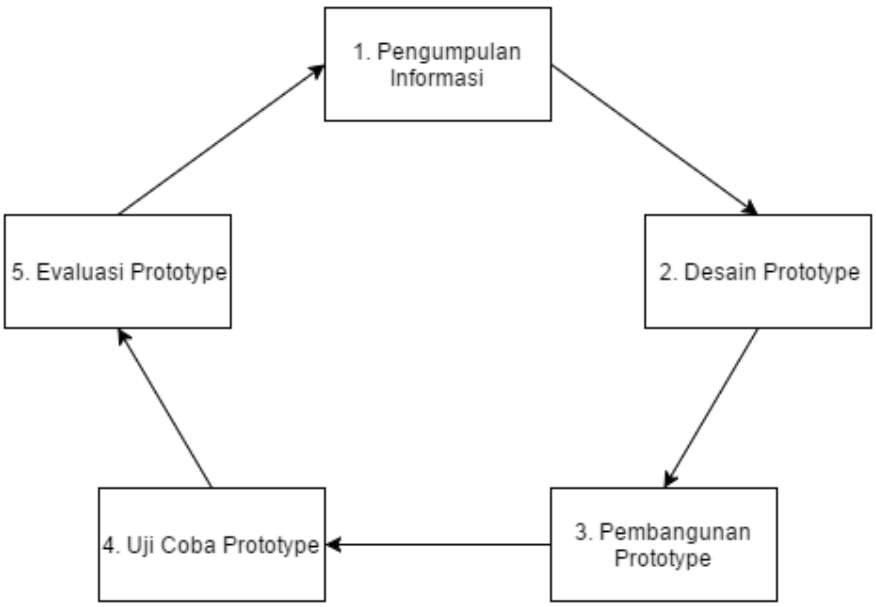
4. Uji Coba *Prototype*

   Prototype yang sudah dibangun diujicobakan dengan data-data pertanyaan yang sudah dipersiapkan sebelumnya.

5. Evaluasi *Prototype*

   Melakukan evaluasi dari hasil uji coba yang dilakukan pada tahap 4 yang dilakukan dengan melihat hasil akurasi. Apabila hasil evaluasi kurang memuaskan, maka pengembang bisa kembali lagi ke langkah pertama.

Secara garis besar, metode *prototyping* bisa dilihat pada gambar bagan berikut.



Gambar 1.2 Bagan Skema Metode Prototyping

1.7 Sistematika Penulisan

Sistematika penulisan laporan tugas akhir ini disusun untuk memberikan gambaran umum tentang penelitian yang dijalankan. Sistematika penulisan laporan tugas akhir ini adalah sebagai berikut.

BAB 1 PENDAHULUAN

Pada bab ini akan dibahas mengenai latar belakang masalah, rumusan masalah, maksud dan tujuan, batasan masalah, metodologi penelitian, dan sistematika penulisan pada laporan ini.

BAB 2 LANDASAN TEORI

Pada bab ini membahas berbagai konsep dasar dan teori-teori yang berkaitan dengan topik penelitian yang dilakukan, seperti *front office*, tanya jawab (*question & answering*), metode yang digunakan, perangkat lunak pembantu, dan sebagainya.

BAB 3 ANALISIS DAN PERANCANGAN

Pada bab ini membahas tentang analisis sistem, analisis pemecahan masalah, analisis kebutuhan fungsional dan non fungsional, serta perancangan sistem yang mencakup perancangan antar muka, alur data, basis data, struktur menu, dan jaringan semantik.

BAB 4 IMPLEMENTASI DAN PENGUJIAN

Merupakan tahapan pembangunan, implementasi dan pengujian perangkat lunak. Pengujian dan evaluasi akan dilakukan secara *online* dengan menghitung jumlah jawaban yang relevan terhadap pertanyaan yang diajukan atau dimasukkan oleh pengguna.

BAB 5 KESIMPULAN DAN SARAN

Pada bab ini akan membahas mengenai kesimpulan dari penelitian ini dan dilengkapi dengan saran-saran yang dapat dijadikan masukan untuk pengembangan penelitian sejenis berikutnya.