BAB 1
PENDAHULUAN

1.1. Latar Belakang Masalah

Membuat pertanyaan dari sebuah teks untuk mengevaluasi pemahaman pembaca mengenai apa yang disampaikan didalam teks merupakan hal yang sudah tidak asing lagi bagi terutama bagi seorang pengajar untuk melakukan evaluasi terhadap siswanya. Akan tetapi manusia memiliki keterbatasan dalam membuat pertanyaan dari teks yang berisi banyak kata. Faktor ketahanan kerja menyebabkan ketidak konsistenan dalam proses pembuatan pertanyaan. Hal ini lah yang membuat para peneliti melakukan penelitian tentang bagaimana pemanfaatan mesin dalam pelaksanaan *intelligent tutoring*.

Intelligent Tutoring merupakan salah satu bahasan dibidang kecerdasan buatan yang dikaji oleh para ilmuwan dibidang *computer science* dengan harapan bahwa manusia bisa belajar dengan lebih cepat mendapatkan pemahaman menggunakan bantuan mesin [1]. Pembangkit pertanyaan otomatis termasuk kedalam studi yang dibahas dalam *intelligent tutoring* dengan tujuan untuk mengevaluasi apakah pengguna mampu memahami isi dari sebuah teks yang telah dipelajari.

Question Generation (QG) merupakan proses yang dilakukan untuk membangkitkan pertanyaan yang dapat diterima oleh manusia dari sebuah input dengan bentuk terstruktur (database) ataupun tidak terstruktur (text) [2]. *Question Generation* merupakan spesialisasi dari *Natural Language Processing* (Pemrosesan Bahasa Alami) yang didalamnya mempelajari bagaimana mesin bisa mengerti dan berinteraksi dengan menggunakan bahasa manusia.

Berdasarkan hasil penelitian sebelumnya dari M. Fachrurrozy dan Novi Yusliani, studi mengenai pembuat pertanyaan otomatis sudah lama menjadi topik dalam Pemrosesan Bahasa Alami. Dari pertanyaan yang dibangkitkan secara otomatis oleh mesin *Question Generator* yang dibuat, 55,95% dari pertanyaan

tersebut dapat diterima [3]. Pertanyaan yang dihasilkan oleh QG memuat pertanyaan yang *nonfactoid* dengan kendala bahwa kesesuaian dengan konteks materi yang dimasukkan sebagai input masih memiliki kekurangan sehingga menurunkan akurasi pertanyaan yang benar. Untuk itu dibangun Question generator yang mampu membuat ringkasan dari materi agar sesuai dengan konteks kemudian dibuat pertanyaan dari hasil peringkasannya.

Dalam penelitian yang akan dilakukan untuk menambah presentase pertanyaan yang dapat diterima agar tidak keluar dari konteks yang dibahas dari input dengan membuat ringkasan dari dokumen input terlebih dahulu. Sehingga diharapkan pertanyaan yang dihasilkan oleh *Semantic Question Generator* (SQG) lebih akurat dalam menghasilkan Pertanyaan yang berhubungan denga inti dari dokumen inut. Dengan tujuan mengevaluasi apakah pengguna memahami inti dari dokumen yang dipelajari dan dibuat sebagai input didalam penggunaan SQG.

Automatic Text Summarization adalah sebuah proses pembuatan ringkasan dari sebuah dokumen secara otomatis dengan mengambil inti dari sebuah dokmen. Untuk mendapatkan ringkasan yang baik dari sebuah text, digunakan metode *lexical chain* dan *correlation of sentences* yang dinilai mampu membuat ringkasan lebih baik dengan tingkat keakuratan (*precision*) yang tinggi [4].

Dengan dilakukannya penelitian ini diharapkan metode *lexical chain* dan *correlation of sentences* yang diterapkan dalam SQG dapat menghasilkan pertanyaan yang baik serta dapat dijadikan sebagai acuan dalam penelitian selanjutnya.

1.2. Rumusan Masalah

Berdasarkan uraian latar belakang yang telah dijelaskan sebelumnya, maka rumusan masalah yang ditemukan adalah sebagai berikut :

1. Bagaimana membuat pertanyaan yang dapat diterima dalam Bahasa Indonesia secara otomatis dari sebuah dokumen text?
2. Bagaimana *Automatic Text Summarization* dapat diimplementasikan dalam pembangkitan pertanyaan secara otomatis?

1.3. Maksud dan Tujuan

Maksud dari dilaksanakannya penelitian ini adalah untuk membuktikan bahwa metode yang digunakan dapat diimplementasikan dengan akurasi ketepatan kalimat tanya mengenai konteks bacaan yang bisa diterima oleh manusia untuk membantu pemahaman pembaca dengan memberikan sebuah pertanyaan secara otomatis tentang inti dari sebuah dokumen.

Adapun tujuan dari penelitian ini adalah :

1. Dapat membangkitkan pertanyaan yang dapat diterima dari sebuah teks berbahasa Indonesia.
2. Mengetahui tingkat akurasi pertanyaan dengan mengimplementasikan peringkasan teks otomatis terhadap pembangkit pertanyaan otomatis.

1.4. Batasan Masalah

Batasan masalah merupakan ruang lingkup pekerjaan yang akan dilakukan terhadap permasalahan yang ada. Adapun batasan dalam penelitian sebagai berikut:

1. Jenis file dokumen yang digunakan untuk input berbentuk txt.
2. Berfokus pada teks berbahasa Indonesia.
3. Metode peringkasan yang digunakan adalah metode *lexical chain*.
4. Teks masukan merupakan teks mata pelajaran Ilmu Pengetahuan Alam untuk Sekolah Dasar kelas 6.
5. Kamus kata dasar merupakan kumpulan kata dasar dari kata yang digunakan dalam pelajaran Ilmu Pengetahuan Alam sekolah dasar.
6. Stop list yang digunakan adalah stop list dari Tala [15].
7. Algoritma stemming yang digunakan menggunakan algoritma nazief dan adriani.
8. Parameter dari ekstraksi yang digunakan yaitu panjang kalimat, bobot kalimat, posisi kalimat, dan kemiripan antar kalimat.
9. Kalimat dibatasi dengan tanda baca “. , ! , ? “.
10. Inputan teks diasumsikan merupakan paragraf deduktif.
11. Pertanyaan yang akan dibangkitkan adalah pertanyaan yang bersifat *non-factoid*

1.5. Metode Penelitian

Metode penelitian yang dilakukan adalah metode pengumpulan data dan metode pembangunan perangkat lunak.

1.5.1. Metode pengumpulan data

Metode pengumpulan data yang digunakan dalam membantu penelitian ini menggunakan dua cara, yaitu dengan cara studi literatur dan dokumen.

1. Studi Literatur

   Mengumpulkan literatur, jurnal, dan bacaan-bacaan yang berhubungan dengan *lexical chain*, *automatic text summarization,* dan *Question generator*.
2. Pengumpulan Data

   Data dikumpulkan dari teks dokumen dan artikel berbahasa Indonesia untuk masukan pada sistem.

1.5.2. Metode pembangunan Perangkat Lunak

Pembangunan perangkat lunak ini menggunakan model prototipe dimana model ini ditujukan untuk membuat sebuah pendekatan terbaik dalam suatu kasus dengan membuat sebuah prototipe dari suatu permasalahan [5].

1. Pengumpulan Kebutuhan

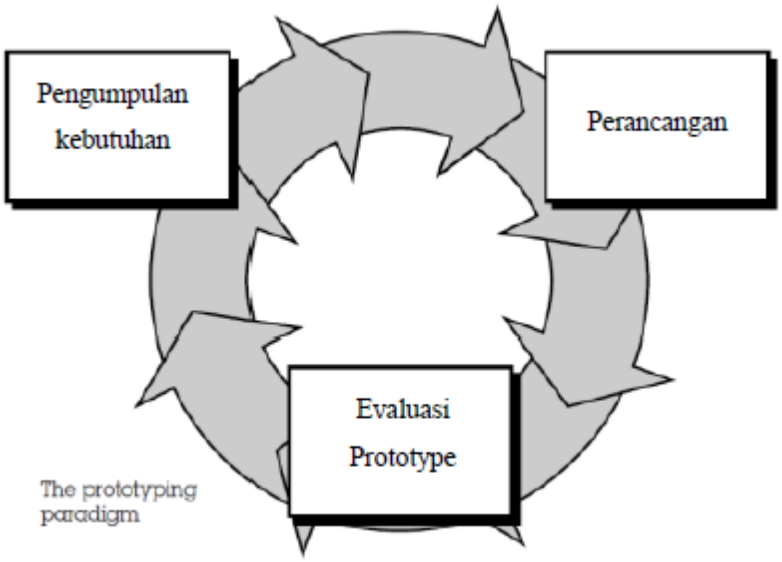
Langkah ini merupakan analisis terhadap kebutuhan software, dan tahap untuk mengumpulkan data dengan melakukan studi dari literatur yang ada sebelumnya berfokus pada literatur mengenai *Question Generator*, *lexical cohesion*, *automatic text summarization* (ATS).

2. Perancangan

Proses perancangan merupakan proses lanjutan dari tahap pembuatan perangkat lunak dengan menggunakan model *prototyping.* Pada tahap ini dilakukan perancangan data dan antar muka perangkat lunak *Automatic Question Generator*.

3. Evaluasi *Prototyping*

Tahap evaluasi *prototyping* membuktikan apakah dengan menggunakan pendekatan semantik dan metode *lexical cohesion* dan *Correlation of sentences* menghasilkan keluaran yang lebih efisien dalam pembuatan pertanyaan.



Gambar 1.1 Model *Prototyping*

1.6. Sistematika Penulisan

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

BAB 1 PENDAHULUAN

Bab ini menguraikan latar belakang permasalahan, menguraikan rumusan permasalahan, menentukan maksud dan tujuan pembatasan masalah, menguraikan batasan masalah, dan sistematika penulisan dari penelitian mengenai pembangkit pertanyaan otomatis dari ringkasan dokumen.

BAB 2 LANDASAN TEORI

Bab ini akan menjelaskan tentang bahasan ilmiah mengenai *Natural Language Processing* (NLP), pembangkit pertanyaan otomatis, peringkas dokumen

teks otomatis, metode *lexical chain*. *Object Oriented Programming* (OOP), *Unified Modeling Language* (UML), Skenario Pengujian.

BAB 3 ANALISIS DAN PERCANGAN SISTEM

Bab ini menganalisis masalah peringkasan sebuah dokumen secara otomatis dengan menghasilkan ringkasan yang baik untuk kemudian dibuat pertanyaan dari ringkasan tersebut agar menghasilkan pertanyaan yang lebih akurat mengenai isi dari sebuah dokumen dengan meminimalisir potensi pertanyaan yang tidak sesuai dengan topik utama dalam sebuah dokumen.

BAB 4 PENGUJIAN SISTEM

Bab ini menjelaskan penerapan bab sebelumnya yakni menguji apakah semua hipotesis yang telah dilakukan untuk pembuatan pertanyaan secara otomatis oleh mesin ini mampu bejalan sesuai dengan yang diharapkan.

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi mengenai kesimpulan tentang keseluruhan dari penelitian dan saran untuk kemungkinan penelitian selanjutnya.