

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi seperti dengan adanya internet, lebih memudahkan manusia dalam mencari dan menemukan informasi yang dibutuhkannya. Karena perkembangan itu pula, jumlah dokumen terutama dokumen teks secara eksplosif meningkat dari hari ke hari, dan menjadi sulit untuk mengelola informasi yang ada dengan membaca semua teks (Maharani & Sanjaya, 2013). Proses mendapatkan informasi dari dokumen teks secara tepat dapat dilakukan dengan membaca keseluruhan isi teks. Membaca keseluruhan isi teks tentunya membutuhkan waktu yang lebih lama jika dibandingkan dengan membaca ringkasan dari teks tersebut.

Summary atau ringkasan didefinisikan sebagai sebuah teks yang dihasilkan dari satu atau lebih teks, mengandung informasi dari teks asli dan panjangnya tidak lebih dari setengah teks asli (Hovy, 2001 dalam (Budiono & Solihin, 2014)). Dengan adanya ringkasan, pembaca dapat dengan cepat dan mudah memahami isi sebuah teks tanpa harus membaca keseluruhan isinya. Hal ini dapat menghemat waktu pembaca karena dapat menghindari bacaan yang tidak relevan dengan informasi yang diharapkan oleh pembaca.

Peringkasan teks otomatis (*Automatic Text Summarization*) merupakan proses dimana komputer membuat versi yang lebih pendek dari teks aslinya (atau kumpulan beberapa teks) yang masih menyimpan sebagian besar informasi yang ada dalam teks asli (Nagwani & Verma, 2011). Dengan penyajian informasi berupa inti dokumen teks secara singkat tetapi mencakup semua informasi dokumen, dapat mempermudah pembaca mengetahui isi teks tanpa harus membaca teks secara keseluruhan. Hal ini dapat menghemat waktu pembaca.

Banyak penelitian yang telah dilakukan untuk melakukan peringkasan secara otomatis. Salah satu metode yang dapat digunakan adalah *Latent Semantic*

Analysis (LSA). *Latent Semantic Analysis (LSA)* adalah metode statistik aljabar yang mengekstrak makna kata-kata dan kesamaan kalimat dengan menggunakan informasi tentang penggunaan kata berdasarkan konteks. LSA menyimpan informasi tentang kata-kata yang digunakan dalam sebuah kalimat, sambil menyimpan informasi dari kata-kata umum di antara kalimat-kalimat. Kata-kata yang lebih umum di antara kalimat berarti bahwa kalimat-kalimat itu lebih berhubungan secara semantik. Metode LSA bisa mewakili arti kata dan makna kalimat secara bersamaan (Ozoy, et al., 2010). Pada penelitian (Luthfiarta, et al., 2013), menyatakan bahwa LSA memiliki kinerja *clustering* lebih baik daripada menggunakan metode fitur dan proses *clustering* standar. Pada penelitian (Winanta & Rainarli, 2016) dikatakan bahwa metode LSA mempunyai tingkat akurasi lebih tinggi dibanding dengan metode *relevance measure*.

Berdasarkan uraian diatas, maka peneliti bermaksud melakukan penelitian dengan judul “Peringkasan Teks Bahasa Indonesia Secara Otomatis Menggunakan Metode LSA”. Jenis dokumen berupa teks yang diambil dari *file* pengguna. Diharapkan penelitian tersebut dapat membantu pembaca memahami isi teks secara singkat.

1.2 Perumusan Masalah

Berdasarkan uraian diatas, penelitian ini dapat dirumuskan yakni “Apakah metode *Latent Semantic Analysis (LSA)* dapat meringkas teks Bahasa Indonesia secara otomatis?”

1.3 Batasan Masalah

Untuk memfokuskan dan menjelaskan ruang lingkup dalam penelitian ini ditentukan batasan-batasan sebagai berikut:

- 1) Bahasa pemrograman PHP.
- 2) Maksimal jumlah kalimat dalam teks tidak lebih dari 120 kalimat.
- 3) Input data berupa file *.txt*.
- 4) *Compression rate* ringkasan sebesar 25%.

1.4 Tujuan dan Manfaat Penelitian

1.4.1. Tujuan

Hal-hal yang ingin dicapai sebagai tujuan dari penelitian ini adalah sebagai berikut.

- 1) Menerapkan *Latent Semantic Analysis* (LSA) untuk meringkas teks Bahasa Indonesia.
- 2) Mempermudah pembaca menemukan inti teks Bahasa Indonesia.

1.4.2. Manfaat

Dengan menggunakan aplikasi ini, maka akan diperoleh manfaat-manfaat sebagai berikut.

- 1) Sebagai referensi untuk penelitian lebih lanjut.
- 3) Pembaca dapat mengetahui ringkasan teks Bahasa Indonesia.

1.5 Metodologi Penelitian

Metodologi penelitian yang digunakan dalam melakukan penelitian dan pengembangan sistem peringkasan teks Bahasa Indonesia secara otomatis menggunakan metode LSA ini, secara terperinci dijelaskan sebagai berikut.

1.5.1. Jenis Penelitian

Klasifikasi penelitian berdasarkan tujuan dan manfaatnya, penelitian ini merupakan sebuah penelitian terapan. Penelitian ini dilakukan untuk menerapkan dan menguji teori peringkasan teks otomatis menggunakan metode *Latent Semantic Analysis* (LSA). Aplikasi berbasis *web* dengan menggunakan bahasa pemrograman PHP. Metode LSA diterapkan untuk meringkas teks Bahasa Indonesia.

1.5.2. Metode Pengembangan sistem

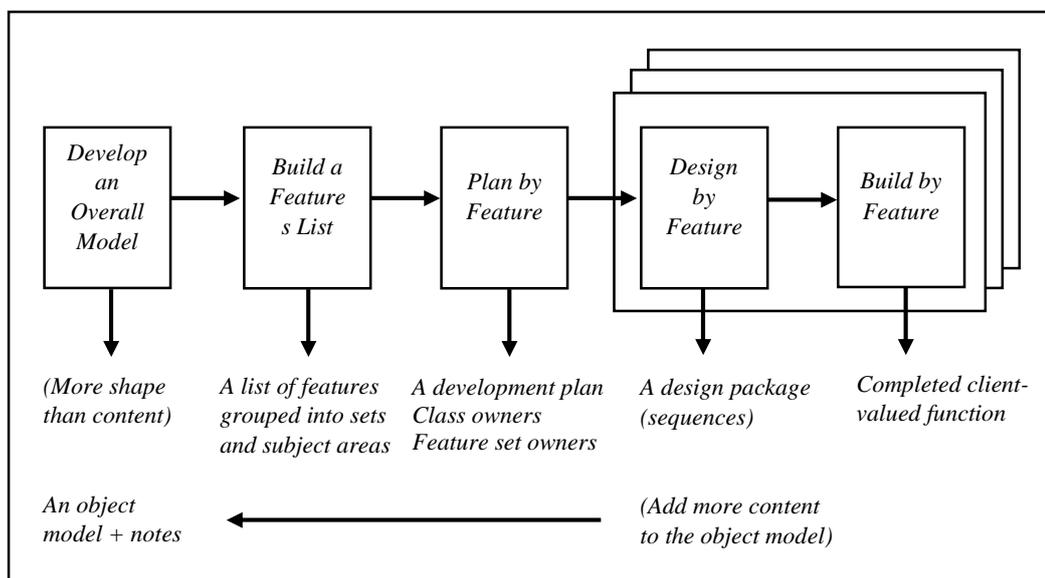
Dalam penelitian ini, pembangunan sistem dilakukan dengan menggunakan metode *Feature-Driven Development* (FDD). *Feature-Driven Development* adalah proses yang didesain dan dilaksanakan untuk menyajikan (*deliver*) hasil kerja secara berulang-ulang dalam waktu tertentu dan dapat diukur. FDD adalah pendekatan yang mengacu pembuatan sistem menggunakan metode yang mudah dimengerti dan mudah diimplementasikan, teknik *problem solving*, dan pedoman pelaporan yang menyediakan informasi yang dibutuhkan oleh *stakeholders* untuk

membuat keputusan yang tepat. FDD memadukan sejumlah praktik terbaik yang diakui industri yang berhasil digunakan oleh Peter Coad, Leff, De Luca dan konsultan lainnya. Praktik ini semua didorong dari perspektif fitur yang dihargai klien (Palmer & Felsing, 2002).

Menurut (Palmer & Felsing, 2002), FDD terdiri dari lima proses.

1. Membuat pemodelan *activity* dari setiap *feature* yang diperlukan.
2. Membuat *feature list* dari pemodelan *activity* yang telah dirancang.
3. Melakukan perencanaan kasar yang tersusun dari kelompok-kelompok kecil berdasarkan tanggung jawab yang ditugaskan.
4. Kelompok kecil melakukan *design and build* yang berlangsung tidak lebih dari dua minggu untuk setiap kelompoknya.
5. Melakukan iterasi *design and build* pada sampai tidak ada lagi kelompok *feature*.

Dalam bentuk skala grafis, metode pengembangan sistem *Feature-Driven Development* (FDD) dapat diperlihatkan seperti gambar 1.1. Model ini dipilih untuk penelitian ini karena pembuatan *feature* akan lebih sesuai dengan kebutuhan *user*.



Gambar 1.1 Lima Proses *Feature-Driven Development*

Sumber: (Palmer & Felsing, 2002)

Tahap-tahap pengembangan model *Feature Driven-Development* dalam penelitian ini dipetakan seperti dalam Tabel 1.1

Tabel 1.1 Tahap Pengembangan Sistem

No	Tahap	Kegiatan
1.	<i>Develop an Overall Model</i>	1. Membuat daftar kebutuhan 2. Membuat daftar konten 3. Melihat cakupan program (<i>scope</i>)
2.	<i>Build a Features List</i>	1. Membuat daftar kelompok-kelompok fitur
3.	<i>Plan by Feature</i>	1. Membuat <i>time-schedule</i> berdasarkan kelompok fitur
4.	<i>Design by Feature</i>	1. Membuat <i>design sistem</i> berdasarkan kelompok fitur
5	<i>Build by Feature</i>	1. Melakukan pemrograman 2. Melakukan <i>testing</i>

1.6 Sistematika Penulisan

Untuk memberikan gambaran secara baris besar mengenai tiap bab yang terdapat dalam tugas akhir, berikut akan diuraikan secara singkat sistematika penulisannya, sebagai berikut:

BAB 1 PENDAHULUAN

Bab ini membahas tentang pendahuluan yang mencakup uraian tentang latar belakang, rumusan masalah, ruang lingkup, tujuan dan manfaat, metodologi penelitian serta sistematika penulisan dijelaskan pada bab ini.

BAB 2 LANDASAN TEORI

Dalam bab ini dibahas mengenai beberapa teori yang dipakai untuk mendukung penelitian.

BAB 3 ANALISIS DAN PERANCANGAN SISTEM

Bab ini menguraikan kebutuhan dasar yang diperlukan selama proses perancangan perangkat lunak (*software*), meliputi pembahasan mengenai *system* peringkasan dengan metode *Latent Sematic Analysis* serta menguraikan tentang gambaran secara umum dari disain dan tampilan-tampilan perangkat lunak (*software*) yang dibangun.

BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini menjelaskan tentang spesifikasi *hardware* dan *software* yang dibutuhkan dalam menjalankan aplikasi, prosedur operasional, implementasi, serta evaluasi dari percobaan yang dilakukan.

BAB 5 KESIMPULAN DAN SARAN

Bab ini kesimpulan yang telah didapat setelah melakukan proses pembuatan aplikasi sistem, serta saran-saran yang diajukan untuk pengembangan sistem.