BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Perkembangan teknologi semakin pesat menyebabkan kebutuhan akan informasi yang sangat besar dan tidak terbatas, terutama informasi dalam bentuk jurnal, artikel, berita, dan dokumen-dokumen lainnya. Artikel yang sering dijumpai pada umumnya tidak mempunyai ringkasan atau abstraknya. Suatu artikel yang memiliki ukuran yang terlalu panjang dapat mengakibatkan pembaca kesulitan bila harus membaca dan menyerap semua informasi dari artikel tersebut, bahkan jika pembaca harus menggali informasi yang diinginkan dari banyak artikel dengan topik yang sama.

Membuat ringkasan secara manual membutuhkan waktu cukup lama jika dokumen yang diringkas berjumlah sangat banyak. Salah satu langkah untuk mendapatkan ringkasan yang cepat adalah dengan menggunakan aplikasi yang dapat meringkas teks secara otomatis. Dengan adanya peringkas teks otomatis ini maka waktu yang dibutuhkan untuk mendapatkan informasi dari suatu dokumen akan lebih cepat. Teknik yang umum digunakan dalam peringkasan teks adalah mengambil kalimat yang penting dari sebuah artikel tersebut, namun pembentukan ringkasan dengan pemotongan secara paksa atau acak terhadap kalimat dapat menghilangkan bagian penting dari artikel.

Tugas akhir ini menggunakan metode *Vector Space Model* (VSM) yaitu salah satu metode pada *text summarization* untuk melihat tingkat kedekatan atau kesamaan *(smilarity) term* dengan cara pembobotan *term* [1]. Algoritma *Vector Space Model* (VSM) merupakan salah satu metode *extractive summary* yang dapat digunakan untuk meringkas dokumen tunggal dengan memperhitungakan nilai *query relevance* pada setiap kalimat. Sebagai pembanding digunakan metode peringkasan teks dengan metode *Support Vector Machine* (SVM) yang merupakan metode baru yang menjanjikan untuk klasifikasi data, baik linier maupun nonlinier [2]. Metode *Support Vector Machine* (SVM) juga merupakan salah satu metode

machine learning yang memaksimumkan akurasi prediksi dengan mencari bidang pembatas terbaik dari dua kelas dalam ruang fitur**.** Metode SVM pada penelitian ini bekerja dengan mengambil himpunan data masukan, kemudian memprediksi untuk setiap masukan yang diberikan ke dalam salah satu kelas dari dua kelas yang ada yaitu kelas positif dan negatif. Dengan adanya kedua metode ini diharapkan hasil ringkasan akan sesuai dengan keinginan pembaca.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang diatas, permasalahan yang terdapat pada penelitian ini yaitu bagaimana membangun suatu sistem yang dapat menampilkan sebuah ringkasan dengan tidak menghilangkan bagian penting dari artikel.

1.3 Maksud dan Tujuan

Maksud dari penelitian ini adalah untuk mengimplementasikan algoritma *Vector Space Model* dan *Support Vector Machine* untuk peringkas teks otomatis artikel berbahasa Indonesia.

Adapun tujuan dari penulisan penelitian ini yaitu menghasilkan suatu sistem yang dapat menampilkan sebuah ringkasan dengan tidak menghilangkan bagian penting dari artikel.

1.4 Batasan Masalah

Adapun beberapa batasan-batasan dalam tugas akhir ini yaitu :

1. Dokumen yang diringkas adalah dokumen berbahasa Indonesia.
2. Dokumen diambil dari media online www.kompas.com
3. Bentuk masukan ke sistem berupa teks dengan format file .docx.
4. Dokumen masukan ke sistem adalah sebuah dokumen tunggal berbahasa indonesia.
5. Kalimat yang dipilih untuk hasil ringkasan yaitu kalimat yang memiliki nilai *query relevance* terbesar dari hasil perhitungan VSM, sedangkan dari hasil perhitungan SVM yaitu kalimat yang berada pada kelas positif.
6. Menggunakan bahasa pemrograman C#.

1.5 Metodologi Penelitian

Metode penelitian yang digunakan adalah metode deskriptif yaitu metode untuk mencari unsur-unsur, ciri-ciri, sifat-sifat suatu fenomena [3].

Metode yang digunakan dalam penelitian ini adalah :

1. Studi Literatur

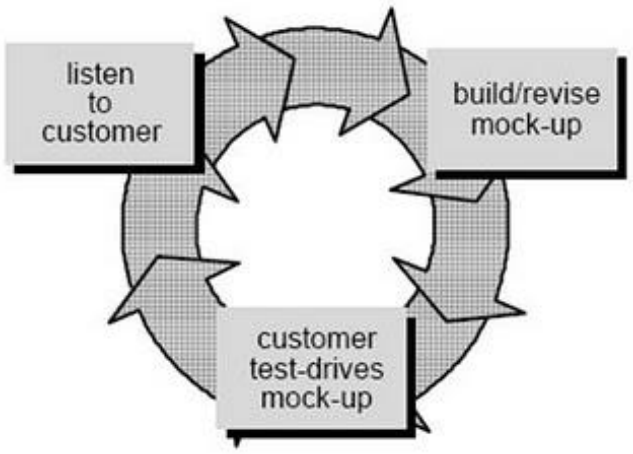
   Studi literatur adalah teknik pengumpulan data dengan cara mengumpulkan literatur, jurnal, buku online dan bacaan-bacaan lainnya serta memanfaatkan media internet yang berkaitan dengan implementasi Algoritma *Vector Space Model*, *Support Vector Machine*, dan pembuatan peringkas teks otomatis.

2. Pengumpulan Data

   Pengumpulan data dilakukan secara manual. Dimana data yang diambil merupakan artikel dengan sumber dari media online. Data dipilih secara acak, disalin, kemudian disimpan dalam format file .docx.

1.5.1 Metode Pembangunan Perangkat Lunak

Dalam membangun aplikasi peringkas teks ini, digunakan metode pembangunan perangkat lunak *Model Prototyping*, karena dapat menggali spesifikasi secara detail aplikasi yang diinginkan oleh pengguna. *Prototyping* adalah proses iteratif dalam pengembangan sistem dimana kebutuhan diubah kedalam sistem yang bekerja (*working system*) yang secara terus menerus diperbaiki melalui kerjasama antar pengguna dan analis [4]. Prototype juga bisa dibangun melalui beberapa *tool* pengembangan untuk menederhanakan proses. *Prototyping* merupakan bentuk dari *Rapid Application Development* (RAD).



Gambar 1.1 Model Prototype menurut *[5]*

1.5.1.1 Tahapan-tahapan *Prototype*

Tahap-tahap pengembangan *Prototype* model menurut [5] adalah :

1. Mengumpulkan kebutuhan pelanggan terhadap perangkat lunak yang akan dibuat.
2. Membangun program prototipe agar pelanggan lebih terbayang dengan apa yang sebenarnya diinginkan.
3. Program prototipe ini dievaluasi oleh pelanggan atau *user* sampai ditemukan spesifikasi yang sesuai dengan keinginan pelanggan atau *user*.

1.6 Sistematika Penulisan

Sistematika penulisan tugas akhir ini disusun untuk memberikan gambaran umum tentang penelitian yang dijalankan, sistematika penulisan tugas akhir ini adalah sebagai berikut :

BAB I PENDAHULUAN

Bab ini menguraikan tentang latar belakang permasalahan, pengimplementasian algoritma Vector Space Model, Support Vector Machine terhadap Text Summarization pada artikel berbahasa Indonesia, identifikasi masalah, maksud dan tujuan, batasan masalah, metodologi penelitian serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini berisikan tentang konsep dasar dan teori yang mendukung dalam pembangunan peringkas teks otomatis, seperti pengertian *text summarization*, tahap *Preprocessing*, pembobotan TF-IDF, perhitungan *cosine similarity*, metode VSM, metode SVM, perhitungan nilai performansi (*recall, precision*, *f-measure*). Teori perancangan aplikasi yang akan dibangun menggunakan pemodelan UML seperti *use case diagram, use case scenario, activity diagram, class diagram, sequence diagram* serta teori-teori tentang *tools* atau *software* pendukung yang digunakan dalam pembangunan aplikasi ini.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini menjelaskan mengenai analisis masalah peringkas teks otomatis, analisis Algoritma, serta perancangan aplikasi yang akan dibangun.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini menjelaskan mengenai hasil implementasi algoritma VSM dan SVM dari hasil analisis dan perancangan sistem yang telah dibuat disertai hasil pengujian sistem yang diterapkan pada aplikasi yang telah dibangun.

BAB V KESIMPULAN DAN SARAN

Bab ini menjelaskan tentang kesimpulan yang diperoleh dari hasil implementasi dan pengujian aplikasi yang telah dibuat, serta saran-saran untuk pengembangan aplikasi ini selanjutnya.