BAB 1
PENDAHULUAN

1.1. Latar Belakang Masalah

Berita adalah cerita atau keterangan mengenai kejadian atau peristiwa yang hangat. Berita haruslah sesuai dengan kenyataan yang ada, tidak dibuat-buat (fiktif), dan terbaru/terkini. Berita merupakan salah satu cara berkomunikasi melalui peristiwa penting, terbaru, dan menarik. Berita dapat dijumpai dimana saja seperti di koran, majalah, internet, televisi, radio, bahkan di mading sekolah. Sebuah berita harus mengandung unsur 5W+1H (*What, Who, When, Where, Why,*dan *How*) supaya pembaca dapat mengetahui lebih banyak tentang suatu kejadian.

Dalam tahun ke tahun jumlah berita semakin bertambah banyak, maka proses pencarian dan penyajian dokumen menjadi lebih sukar atau sulit sehingga minimbulkan beberapa berita yang tidak terklasifikasi kategori sesuai isi berita tersebut. Klasifikasi dokumen teks adalah permasalahan yang mendasar dan penting. Didalam dokumen teks, tulisan yang terkandung adalah bahasa alami manusia, yang merupakan bahasa dengan struktur yang kompleks dan jumlah kata yang sangat banyak. Oleh karena itu, dibutuhkan suatu sistem yang dapat mengklasifikasikan berita secara otomatis.

Berdasarkan latar belakang di atas, maka akan dibangun suatu sistem yang berfungsi untuk mengklasifikasikan berita secara otomastis. Dimana sistem tersebut menjadi salah satu aplikasi dalam melakukan klasifikasi berita. Terdapat beberapa metode yang di gunakan dalam klasifikasi berita seperti menggunakan algoritma K-NN, *naïve Bayesian*, SVM, dan sebagainya, dalam kasus ini metode yang digunakan dalam klasifikasi teks berita bahasa Indonesia adalah metode TF-Idf dan *Vector Space Model*. Diharapkan sistem yang

dibangun menggunakan metode TF-Idf dan VSM akan memudahkan dalam melakukan klasifikasi berita bahasa Indonesia.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, maka didapat rumusan masalah, yaitu : Terdapat beberapa berita yang belum adanya pengklasifikasian kategori.

1.3. Maksud dan Tujuan

Maksud dari penulisan penelitian ini adalah membangun sistem yang dapat membantu dalam menentukan kategori klasifikasi menggunakan metode TF-IDF dan *Vector Space Model*.

Adapun tujuan dari penelitian ini adalah untuk mengetahui berita yang tidak terklasifikasi kategori sebelumnya.

1.4. Batasan Masalah

Berdasarkan latar belakang yang telah diuraikan, maka dibuat batasan masalah agar ruang lingkup penelitian ini jelas batasannya.

Adapun batasan masalah yang dibuat adalah sebagai berikut :

1. Input :
   - Format file berita adalah .docx
   - Berita yang digunakan adalah berita bahasa Indonesia.
   - Berita yang diambil dari media online https://www.kompas.com.
   - Kategori dibatasi 3 jenis yaitu politik, olahraga, dan teknologi.
2. Proses :
   - Perangkat lunak untuk membuat aplikasi ini yaitu visual studio 2010
   - Menggunakan bahasa pemrograman C#.
   - Metode TF-IDF dan *Vector Space Model* yang digunakan pada proses pengklasifikasian.
3. Output :
   - Output yang dihasilkan adalah berupa informasi.

1.5. Metodologi Penelitian

Penelitian ini menggunakan metode penelitian deskriptif yaitu suatu metode dalam meneliti status sekelompok manusia, suatu objek, suatu set kondisi, suatu sistem pemikiran, ataupun suatu kelas perisitiwa pada masa sekarang. Tujuan dari penelitian ini adalah untuk membuat deksripsi, gambaran, atau lukisan secara sistematis, aktual dan akurat mengenai fakta – fakta serta hubungan antar fenomena yang diteliti [1].

Metodologi yang digunakan adalah sebagai berikut :

1. Tahap Pengumpulan Data

Metode pengumpulan data yang digunakan adalah sebagai berikut:

a. Studi literatur

Pengumpulan data dengan cara mengumpulkan literatur, jurnal, buku fisik maupun online (e-book) dan bacaan - bacaan lainnya yang berkaitan dengan klasifikasi menggunakan metode TF-IDF dan *Vector Space Model* dan artikel berita bahasa Indonesia.

2. Tahap pembangunan perangkat lunak

Metode yang digunakan dalam pembuatan perangkat lunak ini menggunakan model *waterfall. Waterfall* adalah model klasik yang bersifat sistematis, berurutan dalam membangun *software* [2]. Motode *waterfall* melakukan pendekatan secara sistematis dan terurut, dimana tahap demi tahap yang akan dilalui harus menunggu tahap sebelumnya selesai. Tahap dari model *waterfall* adalah sebagai berikut :

1. Analisis Kebutuhan

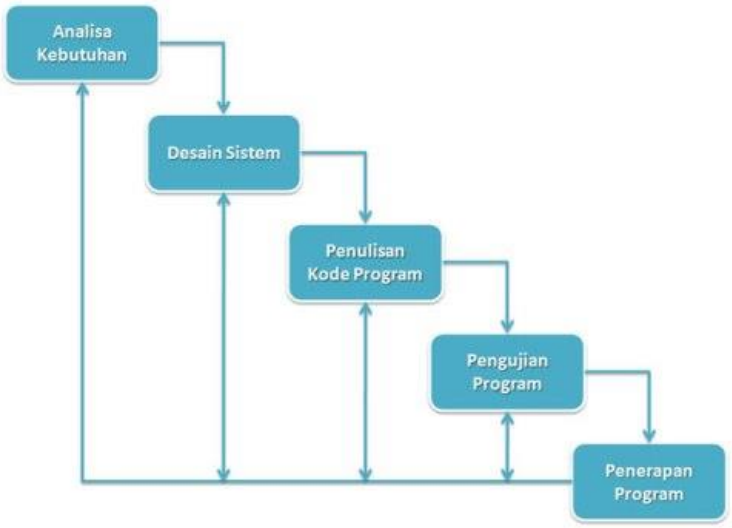
   Merupakan tahap menganalisa hal-hal apa saja yang diperlukan dalam membangun sistem. Hal yang dibutuhkan yaitu data uji berita seperti judul berita, penerbit, dan isi berita dan data *training* berita yang akan digunakan sebagai acuan klasifikasi.
2. Perancangan

   Proses perancangan data hasil analisis kebutuhan dan studi literatur dijadikan gambaran secara umum yang akan dibangun secara lengkap dalam bentuk diagram dan diimplementasikan pada tahap pengkodean.
3. Pengkodean

   Setelah tahap perancangan selesai maka diterjemahkan kedalam bahasa pemrograman.

4. Penerapan

   Tahap ini merupakan tahap pengujian sistem yang dibangun.



Gambar 1. 1 Metode Waterfall [2].

1.6. Sistematika Penulisan

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

BAB I PENDAHULUAN

Menguraikan tentang latar belakang permasalahan, rumusan masalahan, maksud dan tujuan, yang kemudian diikuti dengan batasan masalah, metodologi penelitian, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Bab ini berisikan tentang konsep dasar dan teori yang mendukung dalam pembangunan aplikasi, seperti pengertian berita, TF-IDF dan *Vector Space Model* dan apa saja yang digunakan pada penelitian ini.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Bab ini berisi analisis dalam pembangunan sistem yaitu gambaran umum sistem, analisis basis data, analisis kebutuhan fungsional dan analisis

kebutuhan non fungsional. Pada perancangan berisi mengenai perancangan data, perancangan menu, perancangan antarmuka dari aplikasi yang akan dibuat

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini menjelaskan mengenai hasil implementasi dari hasil analisis dan perancangan sistem yang telah dibuat di sertai 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.