BAB 1
PENDAHULUAN

1.1. Latar Belakang Masalah

Seiring diterbitkannya data survei yang dikeluarkan *Markplus Insight* dan Majalah *Online Marketeers* 2013, terdapat 98% dari 2015 responden menjadikan sumber informasi utama adalah *internet*. Sebanyak 54,2% dari persentase responden tersebut, konten yang dijadikan rujukan sumber informasi adalah berita [1]. Namun, survei yang dilakukan terhadap sampel 41 pembaca menunjukan sekitar 58,84% menganggap bahwa pencarian berita untuk mendapatkan kronologis peristiwa membutuhkan waktu yang lama. Padahal, sebanyak 95% pembaca memiliki tujuan dalam aktivitas pencarian berita adalah untuk mendapatkan alur kronologis suatu peristiwa. Disamping itu, hasil pencarian yang diperoleh melalui media online, terutama pada mesin pencari tidak memahami pola semantik atau tren dari sejumlah data yang telah diperoleh [2]. Sehingga, hasil pencarian tidak terstruktur dan tidak memiliki hubungan sebab akibat antara teks dokumen satu dengan yang lainnya.

Teks representatif terhadap suatu dokumen, secara umum diterapkan sebagai fitur dalam mesin pencari. Namun, teks representatif tersebut masih berdasarkan kata kunci saja. Sehingga secara kualitas belum memenuhi representasi dokumen. Untuk mendapatkan teks representatif, penelitian yang dilakukan oleh Mustaqhfiri [3], teks representatif berupa ringkasan dihasilkan oleh metode *MMR* merupakan hasil ekstraksi kalimat. Metode ini digunakan untuk mengurangi redudansi dalam perangkingan kalimat pada multi dokumen. Kualitas ringkasan yang dihasilkan dari metode *MMR* ini ditunjukan dengan hasil berupa *recall*, *precission* dan kombinasi keduanya berupa *f-measure*. Dari data uji coba yang diambil dari surat kabar berbahasa Indonesia online sejumlah 30, menunjukan hasil rata-rata berupa *recall 60%, precision 77%* dan *f-measure 66%.*

Sekumpulan teks representatif dengan penggunaan metode *MMR* yang terfokus terhadap *query*, merupakan sekumpulan *node* yang memiliki hubungan berdasarkan topik maupun alur waktu. Strukturisasi kumpulan teks representatif

dapat dilakukan dengan memahami keterkaitan antar dokumen teks tersebut. Untuk menentukan sebuah jaringan tanpa sirkuit (struktur pohon tanpa cabang) yang menghubungkan seluruh *node* dapat digunakan *Steiner Tree* [4]. Penerapan *steiner tree* dalam penelitian yang dilakukan oleh Achmad Baihaqi [5], bertujuan untuk menentukan konstruksi suatu jaringan. Dalam penerapannya digunakan dalam jaringan pipa gas dari suatu graf. Implementasi algoritma *heuristic* yang didasarkan pada kecepatan waktu komputasi dalam algoritma *steiner tree* menghasilkan konstruksi pendistribusian gas yang lebih optimal dalam pertimbangan waktu dan biaya. Hal ini dikarenakan bobot *steiner tree* yang dihasilkan oleh algoritma *heuristic* ini belum tentu merupakan bobot yang minimum, akan tetapi bobot tersebut tidak akan melebihi suatu nilai batas atas.

Oleh karena itu, pada penelitian ini akan disusun sebuah penelitian berjudul "*Implementasi Metode Maximum Marginal Relevance (MMR) dan Algoritma Steiner Tree Untuk Menentukan Storyline Dokumen Berita"*

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan diatas, maka rumusan masalah yang diperoleh adalah bagaimana meningkatkan efektivitas dan efesiensi dalam melakukan pencarian berita oleh pembaca media *online* dengan implementasi metode *maximum marginal relevance* dan algoritma *steiner tree* untuk menentukan *Storyline* (alur kronologis) dari dokumen berita tekstual.

1.3. Maksud dan Tujuan

Berdasarkan permasalahan yang diteliti, maka maksud dan tujuan dari penelitian adalah sebagai berikut:

1.3.1. Maksud

Maksud dari penelitian ini adalah untuk melakukan implementasi metode *maximum marginal relevance* dan algoritma *steiner tree* pada sistem yang akan dibangun.

1.3.2. Tujuan

Adapun tujuan dari penelitian ini adalah mempermudah dan mempercepat proses pencarian yang dilakukan oleh pembaca media berita *online* dalam mencari alur kronologis sebuah peristiwa dengan penerapan metode *maximum marginal relevance* dan algoritma *Steiner Tree*.

1.4. Batasan Masalah

Batasan masalah dilakukan untuk membatasi penelitian ini agar dalam implementasi metode *maximum marginal relevance* dan algoritma *steiner tree* dapat dilakukan secara lebih mendalam sehingga pembahasan tidak terlalu umum. Adapun batasan masalah yang dapat dijelaskan adalah sebagai berikut:

1. Dokumen teks artikel diperoleh dari
    a. *detik.com/internasional*,
    b. *internasional.kompas.com*,
    c. *news.okezone.com/internasional*,
    d. *internasional.sindonews.com*,
    e. *dunia.news.viva.co.id* dan
    f. *voa-indonesia.com*
    g. *unisosdem*
2. Dokumen teks artikel berita diambil dalam satu kategori rubrik politik internasional.
3. Peringkasan teks dilakukan dengan proses peringkasan dokumen tunggal untuk masing-masing dokumen.
4. Batasan cakupan setiap teks hasil dari ringkasan kurang sama dengan 2 kalimat.
5. Cakupan sejumlah teks berita yang menjadi alur kronologis dibatasi oleh penentuan waktu peristiwa pada teks berita itu terjadi.
6. Alur kronologis dibentuk dari berita peristiwa.
7. Jumlah dokumen teks artikel berita berjumlah lebih dari 90 dokumen.
8. Teks Artikel berita yang digunakan adalah dalam Bahasa Indonesia.
9. Bahasa pemrograman yang digunakan adalah bahasa pemograman web.

10. Sistem *query* yang digunakan adalah MySQL.
11. Software yang digunakan dalam penelitian ini adalah Dreamwever.
12. Proses pencarian *steiner tree* dari ringkasan banyak teks artikel berita dilakukan secara *offline.*
13. Representasi *Graph* menggunakan *Adjacency Matrix.*
14. Segmentasi kalimat dilakukan berdasarkan simbol titik untuk membagi artikel kedalam beberapa kaliamt.

1.5. Metodologi Penelitian

Metode penelitian yang akan dilakukan dalam penelitian ini adalah dengan melakukan penelitian deskriptif, yakni metode penelitian yag sistematika dan akurat mengenai keadaan fakta penelitian. Gambaran tersebut diperoleh dengan cara mengumpulkan, mengklasifikasikan, menyajikan seta menganalisis data sehingga dapat ditaring suatu kesimpulan. Metode ini dibagi menjadi dua, yaitu pengumpulan data dan pembangunan perangkat lunak.

1.5.1. Metode Pengumpulan Data

Metode pengumpulan data yang digunakan adalah dengan melakukan studi literatur. Pada tahap ini, penggunaan metode pengumpulan studi literatur adalah untuk mencari informasi yang berkaitan dengan permasalahan sekaligus berkaitan erat dengan teknik peringkasan *maximum marginal relevance* dan penggunaan algoritma *steiner tree* dalam beberapa penelitian. Tahap ini dilakukan dengan mengumpulkan literatur, jurnal, paper, buku-buku dan bacaan-bacaan yang dapat membantu menyelesaikan penelitian ini.

1.5.2. Metode Pembangunan Perangkat Lunak

Metode yang digunakan dalam pembangunan perangkat lunak pada penelitian ini adalah metode *waterfall*. Secara garis besar, metode *Waterfall* meliputi beberapa tahapan proses [6], diantaranya:

1. *Requirement Analysis* (Analisis Kebutuhan)

Pada tahap ini dilakukan analisa terhadap kebutuhan sistem. Pengumpulan data dalam tahap ini dilakukan dengan studi literatur. Kemudian, dilakukan analisis terhadap algoritma yang akan diimplementasikan yakni terdiri dari penggunaan metode *maximum marginal relevance* dan algoritma *steiner tree.* Selain itu, analisis yang dilakukan antara lain analisis non-fungsional (kebutuhan *hardware* dan *software*), serta analisis fungsional meliputi *usecase diagram, class diagram, Activity Diagram* dan *squence diagram*.

2. *System Design* (Perancangan Sistem)

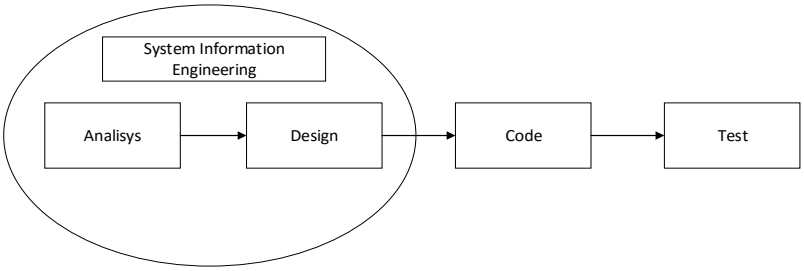
Proses desain akan menerjemahkan syarat kebutuhan kesebuah perancangan perangkat lunak yang dapat diperkirakan sebelum dilakukan *coding*. Tahap ini memiliki tujuan dalam memberikan gambaran apa yang haris dikerjakan dan bagamana tampilannya. Tahap ini serupa dengan perancangan *storyboard* pra implementasi.

3. *Implementation/Coding*

Tahap implementasi yakni berupa penerapan hasil rancangan kepada program merupakan penerjemahan *desain* dalam bahasa yang bisa dikenali oleh komputer. Pada tahap ini dilakukan pemrograman pada sistem yang dibangun, sekaligus melakukan implementasi algoritma didalamnya. Selain itu, pada tahap ini pula dilakukan pemeriksaan, apakah sistem yang sudah dibuat sudah memenuhi fungsi yang diinginkan atau belum.

4. *Integration & Testing*

Tahapan ini bisa dikatakan sebagai tahap final dalam pembuatan sebuah sistem. Karena, pada tahap ini dilakukan pengujian dari segi lojik apakah sistem yang sudah dibangun terdapat kesalahan atau tidak. Maka, dalam tahap pengujian digunakan metode pengujian berupa *whitebox*.



Gambar 1.1 Pemodelan *Waterfall* [6]

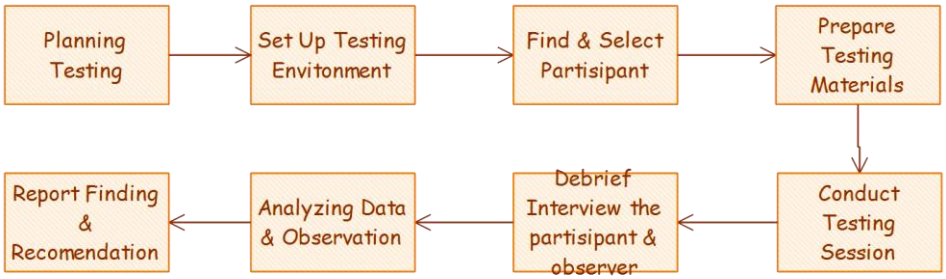
1.5.3. Metode Pengujian

Metode pengujian yang digunakan untuk melakukan evaluasi terhadap perangkat lunak atau sistem pencariann alur kronologis adalah teknik pengujian ketergunaan *(Usability Testing)*. Dengan teknik ini, dilakukan pengujian perangkat lunak dengan langsung kepada pengguna. *Usability* memiliki lima komponen diantaranya *learnability, effeciency, memorability, Errors* dan *Satisfaction*.

Tujuan dari metode pengujian *Usability Testing* ini adalah menginformasikan desain dengan mengumpulkan data untuk mengidentifikasi dan memperbaiki kekurangan kegunaan yang ada dalam produk dan bahan pendukung sebelum rilis, mengeliminasi problem desain dan meminimalisir kesalahan serta menunjukan keuntungan yang diperolah dari sistem [7].

Ada beberapa langkah dalam proses melakukan pengujian ketergunaan *(usability testing)* diantaranya [7]:

1. Mengembangkan perencanaan pengujian
2. Mengatur lingkundan pengujian
3. Mencari dan memilih partisipan
4. Menyiapkan bahan pengujian
5. Melakukan sesi pengujian
6. Mengadakan wawancara terhadap partisipan dan pengamat
7. Analisa data dan observasi
8. Melaporkan temuan dan rekomendasi



Gambar 1.2 Langkah-Langkah Pengujian *Usability Testing*

1.6. Sistematika Penulisan

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

BAB I PENDAHULUAN

Bab ini menjelaskan tentang latar belakang masalah mengenai hal-hal yang menjadi kajian penelitian, merumuskan masalah, menentukan batasan masalah, mengutarakan maksud dan tujuan serta menjelaskan mengenai metode penelitian dan sistematika penulisan yang dimaksudkan untuk memberikan gambaran tentang penyajian laporan ini.

BAB II LANDASAN TEORI

Bab ini menjelaskan tentang metode *maximum marginal* relevance, teori graf, *approximation steiner tree algorithm* dan *Storyline* (alur kronologis) serta bahasa pemograman yang digunakan berikut konsep dasar dan teori-teori yang berhubungan dengan topik penelitian yang akan dilakukan.

BAB III ANALISIS DAN PERANCANGAN

Bab ini menjelaskan tentang analisis masalah berkaitan dengan tahapan dalam penggunan metode *MMR* dan algoritma *steiner tree* dalam menentukan *Storyline* dokumen berita analisis kebutuhan fungsional dan perancangan antarmuka yang menggambarkan bagaimana sistem dapat menentukan *Storyline* dari dokumen berita tekstual melalui *query* yang dimasukan.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab ini menjelaskan implementasi, kebutuhan perangkat lunak, kebutuhan perangkat lunak yang digunakan, implementasi pada sistem yang dibangun, implementasi antarmuka dan pengujian sistem dari penerapan metode *maximum marginal relevance* juga algoritma *steiner tree* yang digunakan.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi kesimpuan yang diperoleh dari hasil implementasi dan pengujian bahwa tujuan awal dalam penelitian ini tercapai atau tidak, serta berisi saran untuk pengembangan terhadap penelitian dan pembangunan sistem dalam menentukan *Storyline* dokumen berita.