

BAB I

PENDAHULUAN

1.1 Latar Belakang

Informasi merupakan salah satu kebutuhan yang sangat penting bagi setiap orang. Informasi dapat diperoleh melalui berbagai media, baik media cetak maupun media digital. Agar informasi penting dapat diketahui pembaca membutuhkan banyak waktu untuk sekedar mendapatkan informasi yang dibutuhkan. Oleh karena itu peringkasan dapat dilakukan untuk menghemat waktu pembaca tanpa menghilangkan informasi penting tanpa harus membaca keseluruhan dokumen tersebut.[1]

Ringkasan merupakan cara mengambil kalimat penting yang dapat mewakili keseluruhan isi informasi dari dokumen dan disajikan secara singkat untuk mempermudah dalam memahami isi dari dokumen tersebut[2]. Proses meringkas keseluruhan dari dokumen/sumber informasi dapat dilakukan dengan membaca terlebih isi dokumen/sumber informasi secara menyeluruh terlebih dahulu. Namun membutuhkan banyak waktu agar dapat memahami keseluruhan isi dokumen/sumber informasi tersebut.

Seiring dengan meningkatnya teknologi informasi untuk mendapatkan ringkasan secara cepat dapat dilakukan tanpa diharuskan untuk membaca keseluruhan isi dari suatu dokumen/sumber informasi dengan teknik *text mining* sehingga dapat menghemat waktu yang diperlukan untuk meringkas keseluruhan isi dari dokumen/sumber informasi[3][4]

Forum diskusi adalah tempat/wadah dalam memberikan pendapat/solusi juga dapat mencari informasi yang lebih akurat dari hasil berbagi yang melibatkan lebih dari satu individu yang saling berdiskusi dalam satu forum[4].

Dengan demikian forum diskusi merupakan salah satu sumber informasi namun dengan semakin banyaknya diskusi yang dilakukan maka akan semakin sulit untuk dapat memahami keseluruhan hasil dari diskusi tersebut. Dengan demikian peringkasan pada forum diskusi diharapkan mampu membantu pengguna dalam mencari informasi hasil diskusi secara cepat.

Penelitian mengenai peringkasan teks menggunakan algoritma *textrank* diantaranya Eris, Viny Christanti M, dan Jeany Pragantha (2017) dengan judul penelitian Penerapan Algoritma *Textrank* Untuk Automatic Summarization Pada Dokumen Berbahasa Indonesia. Penelitian yang dilakukan memperoleh hasil sebesar 82,48% untuk hasil ringkasan konten yang informatif, 50% untuk hasil ringkasan teks, 93,76% hasil konten informatif, dan 75% untuk hasil teks yang dirangkum[1]. Penelitian selanjutnya dilakukan oleh Rivandi Restu Putra dkk (2018) dengan judul penelitian Peringkasan Teks Otomatis Pada Multi Dokumen Menggunakan *Textrank*. Penelitian yang dilakukan memperoleh mencapai hasil 75% *accuration*, 73% *precision*, dan 73% *f-measure*[5]. Penelitian selanjutnya dilakukan oleh Yuzar Marsyah dengan membandingkan algoritma *textrank* dengan algoritma *Lexrank*. Hasil penelitian yang didapat algoritma *textrank* menghasilkan 95.56% mengandung kalimat yang relevan, dan kecepatan dalam menghasilkan ringkasan diperoleh 0,0294 detik oleh algoritma *textrank* dan 0,4321 detik oleh algoritma *lexrank*[6].

Algoritma *textrank* adalah model peringkasan berbasis graf dan termasuk pada metode *unsupervised learning* dimana algoritma *textrank* tidak bergantung pada data latih sebagai acuan untuk melatih sistem. Pada implementasi peringkasan teks metode *textrank* mengekstraksi kalimat yang memiliki rangking tinggi yang kemudian dimasukkan kedalam hasil ringkasan[5]. Proses peringkasan pada algoritma *textrank* meliputi beberapa tahap diantaranya: *preprocessing*, representasi graf yaitu dengan menghitung nilai *content overlap similarity* dan nilai *textrank*. Kalimat dengan urutan terbanyak akan diekstraktif menjadi ringkasan dari dokumen/sumber informasi. Banyak kalimat yang menjadi hasil ringkasan bergantung terhadap presentase nilai kompresi[1]. Algoritma *textrank* juga dipilih karena algoritma ini dapat diimplementasikan juga untuk mengekstraksi kata kunci.[7] Kata kunci atau keyword berperan untuk mengetahui deskripsi dari suatu text, proses indexing, peringkasan text, dan peningkatan performansi pengkategorian.[7][8], selain itu berdasarkan hasil penelitian yang telah dilakukan sebelumnya algoritma ini mampu menghasilkan hasil ringkasan yang baik, serta mampu menghasilkan waktu peringkasan yang cepat.

Berdasarkan uraian tersebut penelitian akan mengembangkan penelitian mengenai peringkasan teks dengan forum diskusi yang menjadi objek penelitian. Penelitian ini diharapkan dapat memudahkan pengguna untuk mencari informasi hasil diskusi secara cepat pada forum diskusi yang disediakan. Maka penelitian mengenai peringkasan teks Otomatis dengan judul “**Peringkasan Teks Otomatis Pada Forum Diskusi Menggunakan Algoritma *Textrank***”

1.2 Perumusan Masalah

Berdasarkan pemaparan latar belakang diatas maka rumusan masalah yang akan dibahas pada penelitian ini yaitu:

- a. Bagaimana mengimplementasikan algoritma *Textrank* untuk membangun sistem yang dapat membuat ringkasan teks otomatis pada forum diskusi online?
- b. Bagaimana kinerja algoritma *Textrank* dalam membuat teks ringkasan pada forum diskusi online?

1.3 Tujuan dan Manfaat

Adapun tujuan dan manfaat yang akan dicapai berdasarkan permasalahan yang diteliti yaitu:

- a. Mengimplementasikan algoritma *Textrank* pada sistem untuk menghasilkan ringkasan dari forum diskusi online.
- b. Mengetahui kinerja algoritma *Textrank* dalam meringkas teks di forum diskusi online.

1.4 Batasan Masalah

Adapun batasan masalah pada penelitian ini yaitu:

- a. Peringkasan teks dilakukan di forum diskusi online tentang islam.
- b. Hasil ringkasan berupa hasil ekstraksi kalimat dari data teks forum diskusi.
- c. Peringkasan hanya dilakukan pada data teks yang berhasil didapatkan (scrapping).

- d. Jumlah data yang dipakai sebanyak 20 data dengan 10 data dari forum diskusi mashara.id, 5 data dari mui.or.id dan 5 data dari muslim.or.id.
- e. Sistem yang dibangun berbasis web dengan bahasa pemrograman *python*.
- f. Aplikasi yang dibangun menggunakan algoritma *textrank* untuk meringkas isi forum diskusi.
- g. Metode pengembangan perangkat lunak menggunakan metode *waterfall*.

1.5 Metodologi Penelitian

Metodologi yang digunakan pada penelitian tugas akhir ini diantaranya:

a. Metode Pengumpulan Data

Pada penelitian ini Metode yang digunakan dalam mengumpulkan data yaitu dengan 2 tahap meliputi:

1) Observasi

Teknik Observasi merupakan teknik pengumpulan data dengan meneliti dan meninjau langsung pada objek yang diteliti

2) Studi Literatur

Yaitu cara mengumpulkan data yang terdapat di kajian literatur, laporan penelitian, jurnal, studi ilmiah, dan bahan bacaan yang memiliki keterkaitan dengan penelitian.

b. Metode Pengembangan Perangkat Lunak

Pada penelitian ini pengembangan perangkat lunak menggunakan metode *waterfall*. Adapun tahapan dalam pengembangan perangkat lunak pada model *waterfall*, diantaranya:

1) Rekayasa Sistem

Kegiatan pengumpulan data untuk mendukung pembangunan sistem dan menentukan arah perangkat lunak yang akan dibangun.

2) Analisis Sistem

Mengumpulkan kebutuhan secara lengkap lalu menganalisis dan mendefinisikan kebutuhan yang harus dipenuhi pada perangkat lunak yang akan dibangun.

3) Perancangan Sistem

Perancangan antarmuka dari hasil analisis kebutuhan yang telah selesai.

4) Pengkodean Sistem

Hasil perancangan sistem diterjemahkan ke dalam source code menggunakan bahasa pemrograman yang sudah ditentukan..

5) Pengujian Sistem

Menyatukan unit-unit program kemudian dilakukan pengujian secara keseluruhan.

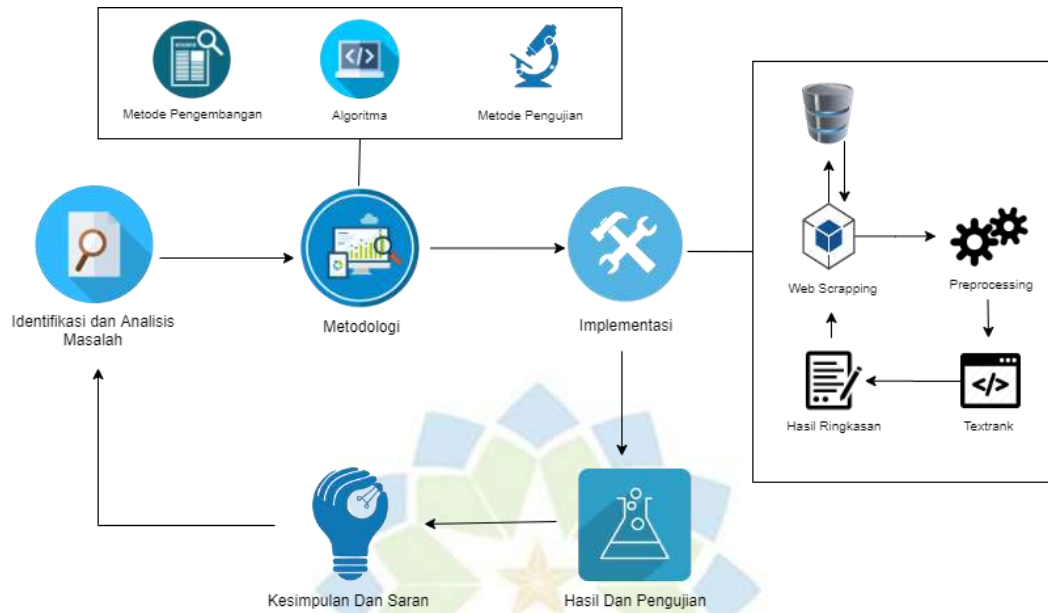
6) Pemeliharaan Sistem

Mengoperasikan perangkat lunak dilingkungannya dan melakukan pemeliharaan, seperti penyesuaian atau perubahan karena adaptasi dengan situasi yang sebenarnya.

7) Umpan Balik.[9]

c. Alur Penelitian

Adapun alur penelitian yang dilakukan pada penelitian yang akan dilakukan digambarkan pada gambar 1.1



Gambar 1.1 Alur Penelitian

Pada gambar 1.1 merupakan gambaran alur penelitian yang dilakukan dimulai dari identifikasi masalah serta menganalisis masalah yang akan dijadikan topik penelitian. Setelah masalah telah diidentifikasi dan di analisis selanjutnya analisis metode yang akan diterapkan pada penelitian.

Pada penelitian ini metode pengembangan menggunakan metode *waterfall*, selanjutnya algoritma yang diterapkan adalah algoritma *texrank*, serta pengujian yang diterapkan untuk mengukur hasil penelitian adalah *ROUGE-L*.

Alur selanjutnya adalah implementasi yang mencakup proses-proses berdasarkan analisis masalah dan metodologi yang digunakan untuk

mendapatkan hasil berdasarkan alur sebelumnya. Adapun langkah yang dilakukan mencakup: langkah pertama yaitu proses mengambil kalimat dari jawaban forum diskusi dengan memasukan url forum diskusi yang akan diringkas, lalu kalimat yang telah diambil dari url akan menjadi data yang akan diproses. Tahap selanjutnya adalah *pre-processing* dimana kalimat dipecah menjadi kalimat-kalimat tunggal, dan membuang kalimat yang tidak diperlukan (*stopword*). Setelah kalimat melalui tahap *pre-processing* kata hasil *pre-processing* akan dicari nilai dari setiap kata dan diurutkan nilai tersebut dari nilai tertinggi hingga terendah untuk menghasilkan *top-rank* menggunakan algoritma *textrank*. Selanjutnya kata akan diekstraksi dan menjadi hasil ringkasan dari algoritma *textrank*. Hasil ekstraksi akan dievaluasi untuk menentukan hasil penelitian. Selanjutnya hasil yang didapat dari implementasi akan dilakukan pengujian untuk mengukur hasil yang didapatkan. Lalu selanjutnya hasil akan disimpulkan dan dapat memberikan saran untuk penelitian selanjutnya.

1.6 Sistematika Penulisan

Penyusunan skripsi ini dibagi menjadi beberapa bab dengan sistematika penulisan sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menjelaskan mengenai Latar Belakang, Rumusan Masalah, Tujuan Penelitian, Batasan Masalah, Metodologi Penelitian dan Sistematika Penulisan

BAB II STUDI PUSTAKA

Pada bab ini berisi mengenai teori-teori penunjang untuk dijadikan landasan tugas akhir serta menyelesaikan masalah yang akan dikaji

BAB III PERANCANGAN SISTEM

Pada bab ini berisikan pembahasan analisis dan perancangan sistem yang akan dibangun

BAB IV IMPLEMENTASI DAN PENGUJIAN

Pada bab ini membahas tentang hasil implementasi perancangan sistem yang dibuat dan berisikan hasil pengujian pada sistem yang sudah dibangun.

BAB V PENUTUP

Pada bab ini berisikan kesimpulan hasil penelitian dan saran untuk penelitian selanjutnya.