

# 1. Pendahuluan

Pada bab ini akan dijelaskan latar belakang, perumusan masalah, tujuan, beserta metodologi penyelesaian masalah pada Tugas Akhir.

## 1.1 Latar Belakang

Pada saat ini perkembangan teknologi informasi semakin meningkat seiring berkembangnya media elektronik yang menyediakan berbagai informasi secara *on-line*. Salah satu bentuk informasi tersebut adalah dokumen maupun artikel berita. Kebutuhan *user* akan dokumen yang berupa berita menyebabkan *user* membutuhkan waktu yang lebih lama untuk membaca keseluruhan dokumen. Oleh karena itu, dibutuhkan informasi yang singkat dan padat yang merepresentasikan isi dokumen. Sehingga dikembangkan sebuah sistem yang dapat meringkas dokumen yaitu, peringkas teks otomatis (*automatic text summarization*).

Banyak metode maupun pendekatan dalam meringkas teks diantaranya berdasarkan cara pengambilan ringkasan adalah metode ekstraksi dan abstraksi. Ekstraksi adalah metode peringkasan kalimat dengan langsung mengekstrak kalimat yang dianggap penting dari teks aslinya, sedangkan abstraksi mengandung *reformulation* kalimat dari teks asli [3]. Metode ekstraksi lebih mudah diterapkan dalam peringkasan teks karena tidak melibatkan *natural language processing*. Kemudian dilihat dari dokumen sebagai masukan, terdapat dua macam dokumen, yaitu *single-document* dan *multi-document*. Pada *single-document* sumber ringkasan hanya terdiri dari sebuah dokumen, sedangkan *multi-document* sumber ringkasan terdiri dari minimal dua dokumen atau lebih yang mempunyai topik yang sama [3].

Pada tugas akhir ini, dibangun dan diimplementasikan *Shortest path algorithm* yang merupakan algoritma berbasis graf untuk melakukan peringkasan dengan metode ekstraksi pada *single-document*. Metode ini telah berhasil digunakan untuk meringkas teks berbahasa Inggris dan mudah dalam implementasi, serta tidak bergantung pada suatu bahasa tertentu. Karena kelebihan yang tidak bergantung pada suatu bahasa tertentu, maka metode peringkasan dengan menggunakan algoritma lintasan terpendek akan coba diimplementasikan pada teks berbahasa Indonesia. *Shortest path algorithm* berbasis pada pencarian lintasan terpendek dari kalimat pertama hingga kalimat terakhir pada sebuah graf yang merepresentasikan teks asli. Simpul pada graf mewakili kalimat dan sisi mewakili kemiripan antar kalimat [7]. *Word overlap* atau jumlah kata yang sama antara dua kalimat digunakan untuk menentukan kemiripan. Oleh karena itu, jika ada dua kalimat yang memiliki minimal satu kata yang sama maka akan ada sisi diantara nya [7].

Kemudian setelah graf terbentuk dan *cost* pada setiap sisi telah dihitung, ringkasan dapat diekstrak berdasarkan kalimat-kalimat yang terpilih pada lintasan terpendek. Dokumen sebagai data *input-an* sistem merupakan artikel berita yang didapat dari beberapa situs internet.

## 1.2 Perumusan Masalah

Pada tugas akhir ini masalah yang diselesaikan yaitu:

1. Bagaimana mengimplementasikan algoritma lintasan terpendek untuk menghasilkan ringkasan pada peringkasan dokumen.
2. Melakukan analisis, apakah metode ini menghasilkan ringkasan yang mempresentasikan isi secara jelas dan juga memberikan informasi-informasi yang penting dari teks asli?

Adapun batasan masalah pada tugas akhir ini yaitu:

1. Data yang digunakan adalah dokumen tunggal berita berbahasa Indonesia.
2. Peringkasan dilakukan secara *off-line*.
3. Hasil ringkasan merupakan ekstraksi dari kalimat-kalimat yang terdapat pada teks asli.
4. Judul setiap teks menjadi kalimat pertama atau kalimat awal teks tersebut.
5. Representasi graf yang digunakan adalah graf berarah dan berbobot.
6. Hasil ringkasan sistem yang diperoleh dibandingkan dengan hasil ringkasan manual yang dibuat oleh manusia.

## 1.3 Tujuan

Tujuan dari Tugas Akhir ini adalah :

1. Membangun sistem peringkasan otomatis dengan menggunakan *shortest path algorithm* yang dapat menghasilkan ringkasan sebagai representasi dari dokumen yang diringkaskan.
2. Mengevaluasi hasil ringkasan yang dihasilkan oleh sistem peringkasan otomatis yang dibangun dari *shortest path algorithm* dengan membandingkan terhadap ringkasan manusia menggunakan *tool ROUGE*.

## 1.4 Metodologi Penyelesaian Masalah

Metodologi yang dilakukan untuk menyelesaikan Tugas akhir ini adalah :

1. Melakukan studi literatur  
Khususnya mengenai peringkasan teks otomatis, *shortest path algorithm*, dan algoritma yang berbasis graf.
2. Pencarian dan Pengumpulan data  
Mengumpulkan dokumen-dokumen berita yang diperlukan untuk mendukung penyelesaian masalah. Data diperoleh dari situs-situs berita seperti kompas dan detik.
3. Analisis kebutuhan dan perancangan
  - a. Mempelajari *shortest path algorithm* untuk peringkasan teks yang akan digunakan dalam implementasi perangkat lunak.
  - b. Melakukan perancangan sistem yang akan dibangun.
  - c. Mempelajari *tool ROUGE* yang akan digunakan sebagai analisa hasil ringkasan sistem.

4. Implementasi
  - a. Membangun sistem yang telah dirancang sebelumnya dengan bahasa pemrograman PHP.
  - b. Dokumen sebagai masukan sistem berekstensi file (.txt).
5. Pengujian dan analisis hasil  
Melakukan pengujian sistem dan menganalisa hasil keluaran sistem yang berupa ringkasan teks, apakah dapat merepresentasikan informasi utama pada teks asli.
6. Pengambilan kesimpulan dan penyusunan laporan tugas akhir.