

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Teknologi informasi telah berkembang pesat, sehingga mudah mengakses informasi melalui media berita *online*. (Adek et al., 2022). Perkembangan inovasi data dapat diperoleh melalui portal berita *online*. Portal berita *online* seperti kapanlagi, suara, detik, cnnindonesia, kumparan, kompas, tribunnews, dan sebagainya sangat digemari oleh masyarakat. Kini masyarakat dapat memperoleh berita atau informasi secara nasional maupun internasional dengan cepat dan tanpa biaya yang mahal. Tingginya minat pembaca berita *online* menimbulkan munculnya portal berita *online* yang tersedia.

Pembaca memerlukan waktu yang lama untuk mencari informasi pada suatu berita karena harus memahami isi berita secara rinci. (Widiantoro, 2014). Dalam situasi tertentu, kita memiliki keterbatasan waktu untuk memperoleh informasi dari berita yang kita baca. Hal ini tentu saja dapat menghambat kita untuk memperoleh informasi secara maksimal. Masalah ini dapat diatasi dengan adanya *parafrase* atau ringkasan berita dalam bentuk teks.

*Parafrase* sangat diperlukan untuk menyampaikan informasi yang terdapat dalam sebuah berita secara singkat. *Parafrase* dapat diartikan sebagai ringkasan dari suatu uraian dengan versi singkat serta mengambil inti dari pokok pembahasan tanpa mengubah struktur wacana. Pembaca sangat memerlukan ringkasan untuk memperoleh informasi yang terdapat pada berita dengan mudah dan cepat. (Robiyanto et al., 2019).

Peringkasan teks otomatis (*automatic text summarization*) merupakan teknik yang digunakan untuk meringkas teks secara otomatis dengan menggunakan algoritma tertentu. Peringkasan memberikan *input* teks kemudian memperoleh *output* berupa teks singkat yang merupakan inti permasalahan dari teks aslinya. *Textrank* merupakan salah satu algoritma yang dapat digunakan sebagai peringkasan teks otomatis. (Abdurrohman, 2018).

Penerapan algoritma *textrank* banyak dilakukan untuk melakukan ekstraksi kata kunci karena dianggap menghasilkan ekstraksi kata kunci baru yang tepat dan relevan, sehingga memudahkan pemahaman pengguna. Pada saat ini banyak pengembangan sistem yang telah dibuat. Algoritma *textrank* dipilih karena dapat menghasilkan penerapan graf tanpa memerlukan dataset penelitian, sehingga mempercepat proses analisis data.

Penelitian terkait juga dilakukan oleh Muhammad Adib Zamzam, Cahyo Crysdiyan, dan Khadijah Fahmi Hayati Holle. Penelitian tersebut menyimpulkan bahwa dengan nilai relevansi menggunakan algoritma *textrank* mendapatkan nilai 41,659%. Pada artikel 48 mendapatkan nilai tertinggi yaitu 0,764, sedangkan pada artikel 37 mendapatkan nilai terendah yaitu 0,167. Dengan memanfaatkan sistem tersebut, pengguna mendapatkan ringkasan dengan lebih cepat karena melibatkan perhitungan komputer. Perhitungan tersebut jauh efisien dibandingkan dengan merangkum secara manual. (Zamzam, 2020).

Pada penelitian (Pragantha et al., 2017) menyimpulkan kalimat yang diambil sebagai hasil ringkasan memiliki nilai *content overlap similarity* yang tinggi dibandingkan kalimat yang lain. Pada penelitian tersebut hasil ringkasan dengan panjang 50% dapat memperoleh informasi sebesar 82,48%. Selain itu hasil ringkasan dengan panjang 75% dapat memperoleh informasi sebesar 93,76%. Jadi ringkasan tersebut hampir mencakup seluruh informasi yang terkandung dalam dokumen dengan panjang 50% dan 75% ringkasan. Berdasarkan uraian tersebut, maka penulis mengajukan sebuah penelitian yang menggunakan perhitungan *textrank* untuk menghasilkan peringkasan teks otomatis pada berita *online*. Penelitian ini dirancang dengan judul **“Peringkasan Teks Otomatis (*Automatic Text Summarization*) Pada Berita *Online* Dengan Algoritma *Textrank*”**.

## **1.2 Rumusan Masalah**

Berdasarkan uraian latar belakang tersebut, penulis mendapatkan permasalahan yang akan diselesaikan pada penelitian ini.

1. Bagaimana mengimplementasikan teknik peringkasan teks otomatis dengan algoritma *textrank*?

2. Bagaimana membuat peringkasan teks otomatis pada berita *online* berbasis web?
3. Bagaimana cara mempermudah pembaca mendapatkan informasi berita dengan cepat?

### **1.3 Batasan Masalah**

Penelitian ini memiliki batasan-batasan masalah sebagai berikut :

1. Informasi atau berita diambil dari detik.com, kompas.com, dan serambi Indonesia.
2. Teks berita menggunakan bahasa Indonesia.
3. Penelitian ini akan menghasilkan aplikasi berbasis web.
4. Bahasa pemrograman yang digunakan pada penelitian ini adalah *Python*.

### **1.4 Tujuan Penelitian**

Adapun tujuan yang ingin dicapai dalam penelitian ini adalah :

1. Mempermudah pembaca dalam memperoleh atau mendapatkan informasi dengan cepat dan hemat waktu.
2. Mengimplementasikan algoritma *textrank* dalam peringkasan teks otomatis.
3. Membuat aplikasi berbasis web dengan menerapkan peringkasan teks otomatis pada berita *online*.

### **1.5 Manfaat Penelitian**

Berikut beberapa manfaat dari penelitian ini adalah :

1. Dapat dijadikan sebagai referensi terhadap penelitian selanjutnya untuk pemanfaatan teknik peringkasan teks otomatis.
2. Penggunaan peringkasan teks otomatis dengan algoritma *textrank* dalam meringkas teks.
3. Dapat berkontribusi dalam ilmu *text mining* bahasa Indonesia yang dapat diimplementasikan untuk membuat ringkasan teks otomatis.
4. Dapat menjadikan *highlight* berita menjadi lebih baik.
5. Mendapatkan informasi kepada pembaca menjadi lebih cepat.