

BAB 1

PENDAHULUAN

1.1 Latar belakang

Sejak terjadinya revolusi digital pada tahun 1980, internet seakan-akan menjadi kebutuhan primer orang-orang baik dari generasi “old” atau lama, dan generasi jaman “now” atau sekarang. Media *online* di era saat ini telah menjadi kebutuhan masyarakat. Apalagi, generasi milenial tidak bisa lepas dari teknologi dan juga perangkat telepon pintar mereka dalam kesehariannya. (<https://kominfo.go.id>) Informasi berbentuk elektronik lebih dibutuhkan dan dicari dibandingkan dengan informasi berbentuk padat misalnya koran dan majalah. Tetapi, tidak semua informasi berbentuk elektronik adalah informasi yang benar-benar dibutuhkan. Terkadang, informasi yang dicari dari search engine tidak akurat dan bahkan mengandung informasi palsu atau HOAX.

Dibutuhkan waktu yang lama untuk mencermati informasi yang kita dapatkan. Apakah informasi yang dicari adalah akurat atau palsu. Belum lagi, terkadang kata kunci yang dimasukan kedalam *search engine* kurang tepat. Hal ini menimbulkan rasa kesal yang cukup mendalam bila sesuatu yang dicari-cari tidak cepat ditemukan. Rasa lelah dan putus asa dapat diterima bila sesuatu yang dicari adalah sebuah informasi penting yang dibutuhkan dalam pekerjaan atau tugas.

Peringkasan Teks Otomatis (Automatic Text Summarization) merupakan pembuatan rangkuman dari sebuah sumber teks secara otomatis dengan menggunakan serta memanfaatkan sistem peringkasan teks yang dijalankan pada komputer. Sebuah sistem peringkasan teks diberi masukan berupa teks, kemudian sistem melakukan proses peringkasan, dan system menghasilkan keluaran berupa teks yang lebih singkat dari sumber teks aslinya. Hasil peringkasan mengandung poin-poin penting atau informasi utama dari teks sumber asli. (<https://repository.widyatama.ac.id>)

Metode *Term Frequency-Inverse Document Frequency (TF-IDF)* ini merupakan suatu cara untuk memberikan bobot hubungan suatu kata terhadap dokumen. Metode ini menggabungkan dua konsep untuk perhitungan bobot, yaitu *Term frequency (TF)* merupakan frekuensi kemunculan kata (t) pada kalimat (d). *Document frequency (DF)* adalah banyaknya kalimat dimana suatu kata (t) muncul. Hasil vektorisasi dari metode TF-IDF lebih akurat dari vektorisasi TF dikarenakan nilai yang didapatkan tidak berdasarkan hanya pada vektorisasi TF, tetapi digabungkan dengan hasil vektorisasi IDF. (<https://journal.uinjkt.ac.id>)

Berdasarkan atas alasan tersebut, maka implementasi TF-IDF kedalam aplikasi peringkasan teks otomatis menjadi tema dari penelitian ini dirasa tepat sasaran dalam masalah tersebut..

1.2 Rumusan masalah

Berdasarkan latar belakang tersebut, rumusan masalah dalam penelitian ini adalah bagaimana membangun aplikasi yang dapat membantu pengguna untuk meringkas dokumen teks dengan baik menggunakan metode TF-IDF.

1.3 Ruang lingkup

Agar dapat mencapai sasaran dan tujuan yang diharapkan maka diberikan batasan masalah sebagai berikut:

1. Aplikasi peringkas menggunakan metode *Term Frequency-Inverse Document Frequency (TF-IDF)* yang dikhususkan untuk metode peringkas.
2. Teks di-*input*-kan secara manual kedalam sebuah *text area*.
3. Proses peringkasan teks bersifat *online* dengan menggunakan flask sebagai dasarnya.
4. Sistem dikhususkan untuk teks Bahasa Indonesia.
5. Data yang digunakan dalam penelitian ini adalah teks berita dengan jumlah paragraf hanya tiga.

1.4 Tujuan penelitian

Adapun tujuan penelitian ini adalah membuat aplikasi peringkas dokumen teks yang dapat merepresentasikan dokumen teks yang aslinya.

1.5 Manfaat penelitian

Manfaat penelitian ini adalah mempermudah pembacaan teks dengan ukuran besar dan mempersingkat waktu dalam membaca teks berukuran besar.

1.6 Sistematika penulisan

Untuk memahami lebih jelas penelitian ini, maka materi yang tertera pada penelitian ini dikelompokkan menjadi beberapa bab dan sub bab dengan sistematika penulisan sebagai berikut:

Bab 1 Pendahuluan

Berisi tentang alasan pentingnya penelitian ini dibuat beserta dengan tujuan dan manfaat penelitian.

Bab 2 Tinjauan Pustaka dan Dasar Teori

Berisi tentang landasan dari pengetahuan yang dikumpulkan dalam rangka menyusun dan meneliti bahan yang akan diteliti.

Bab 3 Metode Penelitian

Berisi tentang pembahasan dari gambaran dan sejarah dari penelitian serupa, struktur dari sistem yang akan diteliti, permasalahan yang terjadi dalam penelitian, alternatif pemecahan masalah, analisa proses, desain sistem yang diteliti, serta tahapan proses penelitian.

Bab 4 Implementasi dan Pembahasan

Berisi tentang hasil dari penelitian, identifikasi hasil, serta penjabaran dari hasil penelitian beserta dengan alasan keberhasilan atau kegagalan dari penelitian ini.

Bab 5 Penutup

Berisi tentang kesimpulan dari hasil penelitian beserta saran dari kekurangan yang didapatkan dalam proses penelitian.