BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi dibidang komputasi dan telekomunikasi telah memberi kemudahan kepada setiap orang untuk menyampaikan dan mendapatkan informasi. Kemudahan ini menyebabkan informasi menjadi semakin banyak dan beragam. Informasi dapat berupa dokumen, artikel, surat, cerita, laporan penelitian, data keuangan dan lain-lain.

Ketersedian informasi yang semakin banyak menjadikan ringkasan sebagai kebutuhan yang sangat penting. Dengan adanya ringkasan, pembaca dengan cepat dan mudah memahami makna sebuah teks tanpa harus membaca keseluruhan teks. Hal ini dapat menghemat waktu pembaca karena dapat menghindari pembacaan teks yang tidak relevan dengan informasi yang diharapkan oleh pembaca, terutama jika informasi yang diharapkan oleh pembaca, terutama jika informasi tersedia diinternet cukup banyak.

Berbagai metode telah banyak diterapkan dan masih terus dikembangkan oleh para peneliti tentang ringkasan. Salah satunya adalah ringkasan dokumen yang kemuduian akan menghasilkan informasi yang penting atau inti kalimat dari dokumen yang dengan menggunakan komputer. Tujuannya adalah mengambil sumber informasi dengan mengutip sebagian besar isi yang penting untuk mencari kalimat inti dari sebuah dokumn artikel yang menampikan kepada pembaca dalam bentuk yang ringkas dengan kebutuhan pembaca. Dengan demikina teknologi ini dapat membantu pembaca untuk menyerap informasi yang ada yang ada dalam artikel melalui ringkasan tanpa membaca seluruh dokumen.

Pencarian otomatis kalimat ini dibutuhkan pembobotan dalam masing-masing term sehingga metode *term frequecy – invers dokument frequensy (tf-idf)* digunakan sebagai penentu bobot masing masing term tersebut. Seingga penulis mengajukan judul tugas akhir "APLIKASI TEKS MINING UNTUK AUTOMASI PENCARIAN KALIMAT INTI DALAM DOKUMEN TUNGGAL BERBAHASA INDONESIA DENGAN METODE *Term Frecuency-Invers Document Frecuenfy* (TF-IDF)".

1.2 RUMUSAN MASALAH

Dari latar belakang yang sudah dijelaskan dia atas diperoleh rumusan masalah sebagai berikut :

a. Bagaimana mengelola dokumen dengan teks mining.

b. Bagaimana menghitung bobot yang dimiliki dalam suatu dokumen dalam keseluruhan teks menggunakan metode TF-IDF.
c. Bagaimana mengurutkan bobot yang diperoleh dari metode TF-IDF.

1.3 Batasan Masalah

a. Input hanya dari satu dokumen atau satu artikel saja.
b. Dokumen berupa artikel dari webside/situs olahraga.
c. Dokumen yang digunakan sebagai data set mengambil dari satus goal.com.
d. Database untuk kata-kata yang tidak relevan mengambil dari kamus bahasa indonesia.
e. Input dan ouput dari bahas indonesia.
f. Tidak menangani kesalahan penulisan kata.
g. Singkatan dianggap satu kata.
h. Hanya melayani list dokumen.

1.4 Tujuan Penelitian.

a. Mengelola dokumen dengan teks mining dan melakukan pembobotan dengan TF-IDF.
b. Melakukan pencarian otomatis kalimat ini dalam sebuah dokumen tunggal berbahasa indonesia.
c. Mengambil sumber informasi dengan mengutip sebagian besar isi yang penting dan menapilkan kepada pembaca.

1.5 Manfaat Penelitian.

a. Dapat menghemat waktu pembaca karena dapat menghindari pembacaan teks yang tidak relevan dengan informasi yang diharapkan oleh pembaca.
b. Dapat membantu pembaca untuk menyerap informasi yang ada dalam artikel melalui ringkasan tanpa harus membaca seluruh isi dokumen.