

Peringkasan otomatis untuk dokumen berbahasa indonesia dengan metode frase penunjuk dan metode TF-IDF

Haryani Diah Sitawati, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=124131&lokasi=lokal>

Abstrak

Peringkasan Otomatis merupakan pembuatan ringkasan dengan bantuan komputer. Peringkasan Otomatis bertujuan untuk menemukan dokumen yang relevan atau mengetahui isi suatu dokumen secara mudah dan cepat. Penelitian ini mengembangkan Peringkasan Otomatis untuk dokumen ilmiah bahasa Indonesia menggunakan dua metode yaitu metode Tf-Idf dan metode Frase Penunjuk. Metode Tf-Idf dan metode Frase Penunjuk memiliki kesamaan yaitu sama-sama menghitung bobot untuk kata-kata tertentu dalam dokumen untuk selanjutnya dijumlahkan untuk tiap kalimat dan diambil yang bobot kalimatnya tertinggi. Khusus untuk metode Frase Penunjuk dilakukan serangkaian pelatihan lebih dahulu untuk menghasilkan kamus Frase Penunjuk untuk dokumen ilmiah berbahasa Indonesia.

Pengujian hasil ringkasan dilakukan dengan menggunakan rasio kompresi 10% dan 20%. Berdasarkan pengujian, didapat bahwa nilai precision dan recall dari ringkasan dengan metode Tf-Idf lebih baik daripada menggunakan metode Frase Penunjuk. Untuk rasio kompresi 10%, nilai precision metode Tf-Idf adalah sebesar 45,1%, sedangkan nilai precision metode Frase Penunjuk adalah sebesar 43%. Namun untuk rasio kompresi 20%, nilai precision metode Tf-Idf adalah sebesar 60,8%, sedangkan nilai precision metode Frase Penunjuk adalah sebesar 60,1%. Nilai recall untuk metode Tf-Idf dengan rasio kompresi 10% adalah 21,6% sedangkan untuk metode Frase Penunjuk adalah sebesar 20,6%. Untuk rasio kompresi 20%, nilai recall dari metode Tf-Idf adalah sebesar 14,5% dan nilai recall dari metode Frase Penunjuk adalah sebesar 13,3%. Jika diukur berdasarkan waktu eksekusi, ringkasan dengan metode Frase Penunjuk membutuhkan waktu yang lebih singkat daripada waktu eksekusi metode Tf-Idf yaitu 0,6 menit dibanding 5,4 menit. Sedangkan untuk panjang kata, ringkasan dengan metode Tf-Idf menghasilkan ringkasan dengan jumlah kata yang lebih banyak dari pada metode Frase Penunjuk. Untuk rasio kompresi 10%, jumlah kata metode Tf-Idf adalah 484 kata, sedangkan untuk metode Frase Penunjuk adalah sebanyak 357 kata. Untuk rasio kompresi 20%, jumlah kata dari metode Tf-Idf adalah sebanyak 845 kata dan untuk metode Frase Penunjuk adalah sebanyak 670 kata.