

Penemuan jawaban pada dokumen berbahasa Indonesia

Dessy Natalia, author

Deskripsi Lengkap: <https://lib.ui.ac.id/detail?id=124257&lokasi=lokal>

Abstrak

Pada era informasi ini, informasi yang tersedia untuk diakses oleh masyarakat sangat berlimpah. Untuk mencari suatu informasi, seorang pengguna dapat memanfaatkan sistem perolehan informasi untuk memperoleh sekumpulan dokumen yang berisi informasi yang dicarinya. Tetapi seringkali untuk mendapatkan jawaban singkat dari suatu pertanyaan, pengguna perlu mencarinya dari banyak dokumen. Untuk menangani kesulitan tersebut, diperlukan suatu sistem yang dapat menerima pertanyaan dalam bahasa sehari-hari dan mengembalikan jawaban secara tepat sehingga pengguna tidak perlu membaca keseluruhan dokumen untuk mencari jawaban. Sistem yang dimaksud adalah sistem tanya jawab, suatu sistem yang dapat mengembalikan jawaban atas pertanyaan bahasa alami dari sekumpulan teks dokumen.

Penelitian ini mengembangkan suatu Sistem Tanya Jawab Bahasa Indonesia (STJBI), yaitu sistem tanya jawab yang melakukan pencarian jawaban dari koleksi dokumen berbahasa Indonesia. Pertanyaan yang dapat diproses STJBI adalah pertanyaan yang berdasarkan fakta tertentu dan berkaitan dengan nama orang, nama organisasi, nama lokasi, serta waktu kejadian. Pertanyaan tersebut dijadikan masukan untuk mencari dokumen yang relevan. Dokumen yang dianggap relevan ini kemudian dibagi menjadi bagian-bagian yang lebih kecil (passage) untuk mendapatkan kandidat jawaban. STJBI menggunakan aturan statistik untuk memberikan bobot pada tiap kandidat jawaban.

Kandidat jawaban yang memiliki nilai tertinggi akan dipilih sebagai jawaban. Pada penelitian ini digunakan tiga ukuran passage yang berbeda yaitu 2, 5, dan 8 kalimat. Hasil uji coba menunjukkan bahwa STJBI dapat mengembalikan jawaban dengan tingkat ketepatan tertinggi sebesar 88.89%, yaitu dengan menggunakan passage yang berukuran 2 kalimat. Dengan menggunakan passage yang berukuran 5 kalimat, STJBI memiliki tingkat ketepatan 86.1%. Sedangkan dengan menggunakan passage yang berukuran 8 kalimat, STJBI hanya memiliki tingkat ketepatan 80.56%. Ketiga jenis ukuran passage memberikan tingkat ketepatan yang sama untuk pertanyaan WAKTU dan ORANG, yaitu sebesar 100% dan 90%. Untuk pertanyaan ORGANISASI, STJBI memiliki tingkat ketepatan 90% (dengan passage berukuran 2 iv kalimat), 80% (dengan passage berukuran 5 kalimat), dan 70% (dengan passage berukuran 8 kalimat). Kinerja STJBI paling buruk diperoleh untuk pertanyaan LOKASI, yaitu hanya 71.4% (dengan passage berukuran 2 dan 5 kalimat) dan 57.1% (dengan passage berukuran 8 kalimat).