

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB I

PENDAHULUAN

1.1. Latar Belakang

Sistem temu kembali (*information retrieval system*) saat ini digunakan untuk mencari informasi data baik berupa data dokumen ataupun citra gambar. Pencarian berbasis *text* berupa sistem temu kembali dapat digunakan untuk mencari informasi pada terjemahan ayat Al-Qur'an. Sistem temu kembali yang diterapkan pada *text* terjemahan Al-Qur'an telah di bangun oleh berbagai peneliti, salah satu penelitian tentang sistem temu kembali terjemahan Al-Qur'an di bangun oleh Imelda Sukma Wulandari (Sukma, 2013).

Pada penelitian tersebut, terdapat kendala berupa hasil dari *query* masukan pengguna dengan maksud tertentu tidak sesuai dengan harapan pengguna karena pengguna tidak mengetahui detil koleksi dokumen sehingga penggunaan *query* masukan pengguna tidak sama dengan *term* pada koleksi dokumen. Kasus yang sering terjadi juga, dokumen yang sebenarnya relevan terhadap *query* yang frekuensi kata *query*nya kecil akan berada pada *ranking* bawah. Sebaliknya suatu dokumen yang sebenarnya tidak relevan terhadap *query* karena kebetulan mengandung kata *query* dengan frekuensi besar akan berada pada *ranking* atas. Oleh karena itu, dibutuhkan *clustering* antar dokumen (Hamzah, Susanto, Soesianto, & Istiyanto, 2007)

Clustering membantu mengelompokan dokumen yang memiliki jarak terdekat menjadi satu *cluster*. Beberapa penelitian tentang *cluster* pada sistem temu kembali sudah di bangun adalah *Clustering* untuk peningkatan efektifitas penyajian informasi dari mesin pencari teks (Hamzah, Susanto, Soesianto, & Istiyanto, 2007), Temu kembali informasi berbasis kluster untuk sistem temu kembali informasi teks bahasa indonesia (Hamzah, 2009), pada penelitian (Hamzah et al., 2007) yang meneliti tentang *clustering* untuk peningkatan efektifitas penyajian informasi dari mesin pencari teks, diperoleh Peningkatan *F-measure* sebagai kinerja ukuran efektifitas sebesar 14,34% untuk koleksi berita

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1.2. Rumusan Masalah

Berdasarkan latar belakang telah diuraikan di atas, maka dapat diambil suatu rumusan masalah yaitu “Bagaimana menerapkan *Automatic Query Expansion* (AQE) dan *Search Result Clustering* (SRClus) pada aplikasi sistem temu kembali (*Information retrieval system*) terjemahan ayat Al-Qur’an.”

1.3. Batasan Masalah

Batasan masalah dalam laporan tugas akhir ini adalah:

1. Koleksi dokumen yang digunakan sebagai bahan penelitian adalah korpus terjemahan ayat Al-Qur’an bahasa Indonesia berdasarkan terjemahan yang dikeluarkan oleh Kementerian Agama dari penelitian Imelda Sukma pada tahun 2013
2. *Clustering* pada tahapan *Search Result Clustering* (SRClu) menggunakan metode *k-means*.

1.4. Tujuan

Tujuan yang ingin dicapai dari penelitian dan penyusunan tugas akhir ini adalah:

1. Menerapkan *Automatic Query Expansion* (AQE) dan *Search Result Clustering* (SRClus) pada sistem temu kembali Ayat Al-Qur’an.
2. Mengukur *Precision* dan *Recall* untuk pencarian ayat Al-Qur’an dengan menerapkan *Search Result Clustering* (SRClus) dan *Automatic Query Expansion* (AQE).

1.5. Sistematika Penulisan

Laporan tugas akhir ini terdiri dari enam bab, dengan sistematika penulisan tersusun atas bab:

BAB I Pendahuluan

Membahas mengenai latar belakang, rumusan masalah, batasan masalah, tujuan pembahasan, serta sistematika penulisan.

BAB II Landasan Teori

Membahas mengenai konsep dasar dan teori dari korpus, *information retrieval system*, arsitektur *information retrieval system*, model dalam

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

information retrieval system, kualitas teks *retrieval*, *Search result clustering*, *Automatic Query Expansion* dan pengujian untuk menilai kemampuan sistem.

BAB III Metodologi Penelitian

Membahas mengenai bagaimana tahapan-tahapan yang akan dilakukan untuk menjawab perumusan masalah penelitian berupa identifikasi masalah, perumusan masalah, studi literatur, analisa sistem, pengembangan perangkat lunak, implementasi sistem, pengujian sistem, serta kesimpulan dan saran.

BAB IV Analisa dan Perancangan

Membahas mengenai analisa permasalahan, analisa sistem, analisa tahapan *information retrieval system*, perancangan sistem, perancangan *database*, serta perancangan tampilan menu sistem.

BAB V Implementasi dan Pengujian

Membahas mengenai tahapan implementasi terhadap sistem yaitu kebutuhan perangkat keras dan perangkat lunak, hasil implementasi *interface* sistem. Serta tahapan implementasi yaitu rencana pengujian, hasil pengujian dan kesimpulan pengujian untuk kerja sistem.

BAB VI Penutup

Membahas mengenai uraian beberapa kesimpulan dari hasil yang didapatkan selama penelitian dan saran yang dapat digunakan pada penelitian selanjutnya.