



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

Perkembangan teknologi informasi pada saat ini dalam berbagai aspek kehidupan seperti politik, budaya, pendidikan, bisnis dan lainnya memberikan kita kemudahan untuk mencari informasi. Hal tersebut dapat dilihat dengan banyaknya mesin pencarian informasi yang ada pada saat ini. Dengan mesin pencarian tersebut manusia dapat mencari informasi berupa dokumen, gambar, video, cerita, laporan penelitian dan lainnya dengan mudah.

Salah satu aplikasi digitalisasi ayat Al-Quran adalah Lidwa, aplikasi ini memiliki terjemahan ayat Al-Quran lengkap yang bisa dicari akan tetapi, hasil pencarian hanya menampilkan *query* yang sama persis dari dokumen yang ada di dalam basis data. Penelitian mengenai digitalisasi dan retrieval ayat Al-Quran juga sudah pernah dilakukan sebelumnya. Beberapa penelitian tersebut diantaranya adalah penelitian (Wulandari, 2013) dengan judul penelitian “Sistem Pencarian Ayat Al-Qur’an Berdasarkan Terjemahan Bahasa Indonesia dengan Pemodelan Ruang Vektor”. Berdasarkan penelitian tersebut dengan menggunakan metode ruang vektor pada kasus pencarian ayat Al-Qur’an memberikan hasil dengan rata-rata *precision* 74% dan *recall* 100% dari 12 kali percobaan terhadap *query* dan pada penelitian (Bintana, 2012), “penelitian penerapan model Okapi BM25 pada sistem temu balik informasi” namun penelitian ini tidak menggunakan *stemming*, *Stemming* adalah proses mengembalikan kata kedalam bentuk dasarnya dengan cara memotong imbuhan pada kata tersebut baik itu berupa awalan, akhiran ataupun sisipan (Marsya & Abidin, 2011). Oleh karena itu, terdapat kekurangan pada hasil pencarian yang berimbuhan.



- 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.

Penelitian tentang pengembangan algoritma *stemming* menggunakan algoritma oleh *enhanced confix stripping* Tahitoe dan Purwitasari, (2010), dengan judul “Implementasi Modifikasi *Enhanced Confix Stripping Stemmer* untuk Bahasa Indonesia dengan Metode *Corpus Based stemming* “ penelitian ini menjelaskan tentang memperbaiki kesalahan-kesalahan *stemming* yang dilakukan oleh algoritma nazief dan andriani dengan hasil dapat memperbaiki seluruh kesalahan *stemming* yang dilakukan oleh algoritma ECS Stemmer.

Penelitian Sistem Pencarian Terjemahan Ayat Al-Qur’an Berdasarkan Bahasa Indonesia dengan Pemodelan Ruang Vektor hampir memberikan hasil yang relevan oleh sistem yang dibangun akan tetapi system tidak bisa menampilkan sinonim dari *query* yang dimasukan. Karena di dalam Al-Quran banyak istilah yang memiliki arti yang sama. Kelemahan pada penelitian tersebut dapat ditutupi dengan penggunaan tesaurus, hal ini didasarkan pada penelitian sistem temu kembali Hadits oleh (Rasyidi, Romadhony, & Wibowo, 2013), pada penelitian mereka dengan judul “*Indonesian Hadith Retrieval System using Thesaurus*” didapatkan suatu kesimpulan bahwa penggunaan tesaurus pada sistem temu kembali hadist juga dapat meningkatkan kualitas hasil pencarian dalam menemukan dokumen yang lebih relevan dan ekspansi *query* menggunakan tesaurus dapat meningkatkan kualitas hasil pencarian.

Penelitian tentang sistem temu kembali menggunakan *Algoritma Latent Semantic Indexing (LSI)* telah dilakukan oleh penelitian (Lee, Baker, Song, & Wetherbe, 2010), yang berjudul “*An Empirical Comperison Of Four Text Mining Methods*” yang membandingkan empat buah metode untuk penentuan topik terbaik. Penelitian ini menerangkan bahwa metode LSI adalah metode terbaik dalam penyelesaian kasus relevansi dan mencari kata yang bersemantik

LSI adalah teknik matematika dan statistika untuk menentukan relavansi dan semantik antara kata satu dengan yang lain didalam dokumen, sehingga dapat memperoleh hasil yang relavan dan akurat, karna metode LSI ini adalah metode yang sudah diakui akurasi dan peformanya dalam memberikan informasi kembali



(Landauer, Folt, & Laham, 1998), Penelitian tentang sistem temu kembali menggunakan algoritma LSI juga telah dilakukan oleh (Kumar & Srinivas, 2009)“ *On the Performance of Latent Semantic Indexing-based Information Retrieval*” pada penelitian ini menjelaskan pada model LSI pada IR lebih baik dari model VSM *classic* dengan rata-rata 30% lebih baik dan pada penelitian (Aji, Baisal, & Firdaus, 2011) yang berjudul “*Automatic Essay Grading Menggunakan Latent Semantic analysis*” mampu memberikan hasil yang baik pada permasalahan relevansi dan semantik kata pada dokumen. Maka dari itu penulis mengusulkan penerapan metode LSI tersebut.

Berdasarkan permasalahan dan penelitian terkait yang telah di uraikan di atas, maka dibutuhkan sebuah sistem yang dapat menemukan kembali Al-Quran yang sangat banyak tersebut secara efisien dan lebih cepat dalam pencariannya, berdasarkan *query* yang di masukan oleh pengguna, yang dilakukan melalui isi dari Al-Quran. Hal tersebutlah yang menjadi latar belakang penulis untuk melakukan sebuah penelitian dengan judul “**Penerapan *Latent Semantic Indexing* pada Sistem Temu Kembali Terjemahan Ayat Al-Quran Bahasa Indonesia** “.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, maka dapat dirumuskan masalahnya yaitu bagaimana membangun sistem temu kembali informasi Al-Quran bahasa Indonesia dengan menerapkan algoritma *Latent Semantic Indexing* dan tesaurus

1.3 Batasan Masalah

Agar fokus penelitian ini terjaga, diberikan beberapa batasan sebagai berikut:

1. Korpus yang digunakan adalah korpus terjemahan ayat Al-Quran bahasa Indonesia.
2. Jumlah korpus sebanyak 500 terjemahan ayat Al-Quran bahasa Indonesia dari surah ke-64 (At-Taghabun) ayat ke-13 sampai dengan surah ke-79 (An-nazi’at) ayat ke-24.



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.

Analisa perangkat lunak merupakan langkah pemahaman persoalan sebelum mengambil tindakan atau keputusan penyelesaian hasil utama. Memaparkan analisa sistem temu kembali informasi Al-Quran bahasa Indonesia dengan menerapkan model *Latent Semantic Indexing* perancangan sistem yang mencakup rancangan usecase diagram, DFD, perancangan *interface* sistem, dan mesin IR (*Information Retrieval*).

Bab V Implementasi Dan Pengujian

Bab ini berisi tentang penerapan aplikasi sistem temu kembali informasi Al-Quran bahasa Indonesia. Tahapan ini juga menggambarkan hasil akhir sistem.

Bab VI Penutup

Bab ini berisi tentang kesimpulan hasil dari semua tahap yang telah dilalui selama penelitian tugas akhir beserta saran-saran yang berkaitan dengan tahapan penelitian.