

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Dizaman yang semakin maju dengan perkembangan teknologi, manusia mencoba mendigitalisasi Al-Qur'an dengan membuat sebuah perangkat lunak yang dapat mencari terjemahan di dalam Al-Qur'an dengan memasukan kata kunci (*query*) dan menampilkan daftar hasil hanya berupa dokumen terjemahan yang relevan dalam Al-Qur'an dengan menerapkan ilmu *Information Retrieval*.

Pendigitalisasian Al-Qur'an pernah dilakukan oleh Wulandari (2013). Menerapkan pemodelan ruang vektor dalam pencarian ayat Al-Qur'an berdasarkan teks terjemahan dengan algoritma *stemming* Nazief dan Adriani. Menggunakan korpus terjemahan ayat Al-Qur'an bahasa Indonesia. Hasil penelitian tersebut mendapatkan tingkat keakuratan mencapai 74%. Tingkat kemampuan untuk mengambil semua ayat berdasarkan *query* mencapai nilai 100%.

Peningkatan kinerja *Information Retrieval* dapat dilakukan dengan menerapkan teknik *stemming* (Tahitoe, 2010). *Stemming* adalah suatu teknik pencarian bentuk dasar dari suatu *term*. *Term* adalah setiap kata yang berada pada suatu dokumen. *Stemming* dilakukan pada saat pembuatan indeks dari suatu dokumen. Pembuatan indeks dilakukan karena suatu dokumen tidak dapat dikenali langsung oleh suatu *information retrieval (IR)*.

Menurut Tahitoe (2010), penggunaan *stemming* dengan algoritma *Enhanced Confix Stripping (ECS)* memiliki kemampuan presentase keakuratan (*precision*) lebih baik dibanding algoritma *Confix Stripping* dan algoritma *Nazief Adriani* dengan perbedaan mendasar antara algoritma tersebut adalah penambahan aturan-aturan baru terhadap algoritma *Enhanced Confix Stripping (ECS)* yang belum terdapat pada algoritma *stemming Confix Stripping* dan algoritma *Nazief Adriani*.



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.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah di paparkan, maka dirumuskan masalah pada penelitian ini, yaitu “Bagaimana menerapkan Algoritma *Enhanced Confix Stripping (ECS)* untuk melakukan proses *stemming* dan Pemodelan Ruang Vektor untuk mendapatkan dokumen yang terurut berdasarkan *query*, serta melakukan ekspansi *query* dengan daftar kata sinonim dan hiponim dalam pencarian terjemahan Al-Qur’an dalam bahasa Indonesia dengan mengadaptasi cara kerja dari *Information Retrieval*.”

## 1.3 Batasan Masalah

Adapun yang menjadi batasan masalah dalam laporan tugas akhir ini adalah sebagai berikut:

1. Koleksi dokumen (korpus) yang digunakan sebagai bahan penelitian adalah korpus terjemahan Al-Qur’an bahasa Indonesia dari penelitian sebelumnya yang dilakukan oleh Wulandari (2013). Korpus yang digunakan merupakan satu ayat atau gabungan dari beberapa ayat Al-Qur’an. Karena dalam Al-Qur’an ada ayat yang bersambung dengan ayat sesudahnya.
2. Terjemahan yang digunakan mencakup seluruh isi Al-Qur’an yaitu 30 juz.
3. Ekspansi *query* menggunakan kamus kata sinonim dan hiponim dalam bahasa Indonesia.
4. Masukan yang digunakan berupa kata dalam bahasa Indonesia dan keluaran yang dihasilkan adalah urutan terjemahan yang terurut berdasarkan masukan.

## 1.4 Tujuan

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

1. Menerapkan algoritma *Enhanced Confix Stripping* untuk proses *stemming* dalam sistem temu kembali terjemahan Al-Qur’an dalam bahasa Indonesia.

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.

2. Menerapkan Pemodelan Ruang Vektor dalam sistem temu kembali topik terjemahan Al-Qur'an dalam bahasa Indonesia.
3. Menerapkan ekspansi *query* untuk menghasilkan lebih banyak dokumen yang relevan berdasarkan *query*.
4. Menguji tingkat akurasi sistem temu kembali informasi dalam pencarian terjemahan Al-Qur'an dalam bahasa Indonesia.

## 1.5 Sistematika Penulisan

Sistematika penulisan laporan Tugas Akhir ini terdiri dari pokok-pokok permasalahan yang dibahas pada masing-masing yang diuraikan menjadi beberapa bagian :

### Bab I. Pendahuluan

Bab ini membahas tentang gambaran umum isi tugas akhir yang meliputi latar belakang masalah, rumusan masalah, batasan masalah, tujuan pembahasan dan sistematika penulisan.

### Bab II. Landasan Teori

Bab ini menjelaskan tentang teori-teori yang berhubungan dengan penelitian yang terdiri dari penjelasan mengenai konsep sistem temu balik informasi, model ruang vektor, penerapan *stemming*, dan ekspansi *query*.

### Bab III. Metodologi Penelitian

Bab ini membahas langkah-langkah yang dilaksanakan dalam proses penelitian, yaitu identifikasi masalah, pengumpulan data, analisa, perancangan sistem, implementasi dan pengujian sistem.

### Bab IV. Analisa Dan Perancangan

Berisi pembahasan mengenai analisa sistem meliputi analisa data sistem temu balik informasi, analisa komponen sistem dan perancangan pada

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.

sistem temu balik informasi yang dapat mencari terjemahan Al-Qur'an dan dapat menampilkan terjemahan yang relevan secara terurut.

## Bab V. Implementasi dan pengujian

Dalam bab ini membahas tentang implementasi model ruang vektor dalam pembuatan sistem temu balik informasi serta melakukan pengujian *recall* dan *precision* terhadap sistem yang dibuat untuk mengetahui apakah sistem tersebut telah dapat memaksimalkan pengelolaan teks terjemahan Al-Qur'an dengan pencarian yang lebih cepat.

## Bab VI. Penutup

Dalam bab ini akan dijelaskan mengenai beberapa kesimpulan yang didapatkan dari pembahasan pembuatan sistem temu balik informasi dengan menggunakan model ruang vektor dan menerapkan *stemming Enhanced Confix Stripping* disertai beberapa saran sebagai hasil akhir dari penelitian yang telah dilakukan.