

# BAB 1

## PENDAHULUAN

### 1.1. Latar Belakang

Al-Quran merupakan kitab suci bagi umat Islam yang ditulis menggunakan bahasa Arab. Dengan jumlah ayat yang cukup banyak, pencarian kata pada teks Al-Quran secara manual sulit dilakukan. Oleh karena itu dibutuhkan sistem pada komputer untuk membantu melakukan pencarian kata pada ayat-ayat Al-Quran.

Bagi pengguna yang tidak bisa berbahasa Arab maupun menulis aksara Arab akan terasa sangat sulit untuk melakukan pencarian ayat Al-Quran. Kendala lain saat melakukan pencarian adalah kesalahan pada pengetikan kata yang dimaksud sehingga solusi yang diharapkan tidak dapat ditemukan. Pada referensi [1] disebutkan bahwa umumnya perangkat lunak pencarian ayat menggunakan teknik *exact string matching* yang merupakan teknik pencarian ayat yang sesuai dengan *string* inputan secara tepat dengan susunan karakter dalam *string* yang dicocokkan memiliki jumlah maupun urutan karakter yang sama. Sebagai contoh jika potongan ayat yang ingin dicari adalah “*robana*” maka teknik *exact string matching* tidak dapat menemukan pada Al-Quran transliterasi Latin karena tidak terdapat ayat yang mengandung potongan *string* tersebut. Potongan ayat Al-Quran yang ada hanyalah “*rabbanaa*”. Kesalahan tersebut terjadi karena perbedaan penulisan *string* pencarian dengan teks ayat Al-Quran, namun memiliki kesamaan pengucapan. Oleh karena itu perlu dikembangkan suatu sistem pencarian ayat Al-Quran berdasarkan kata yang terkandung dalam Al-Quran maupun berdasarkan kemiripan kata dalam pengucapan (fonetis) untuk memperkecil pengaruh kesalahan dalam pengetikan dengan menggunakan teknik *inexact string matching* [1].

Pencocokan fonetis dapat dilakukan dengan kata kunci berupa pelafalan kata dalam aksara Latin Indonesia dan bersifat toleran terhadap perbedaan penulisan *string* ayat Al-Quran berdasarkan cara pelafalannya yang mungkin terjadi. Pada penelitian sebelumnya telah dilakukan pencocokan *string* berdasarkan kemiripan pengucapan menggunakan algoritma *metaphone* yang menunjukkan nilai akurasi pencarian ayat Al-Quran berupa *precision* sebesar 54% dan nilai *recall* sebesar 100% [2]. Pada penelitian Mokhammad Syaroni menunjukkan hasil pengujian kemampuan *retrieval* dalam pencarian kata yang paling baik ditunjukkan oleh algoritma *Soundex* [1]. Hal ini disebabkan algoritma *soundex* menggunakan panjang kode fonetis lebih pendek, yaitu empat karakter.

Dengan permasalahan di atas, pada penelitian ini akan dibangun sistem pencarian ayat Al-Quran dengan metode pencocokan *string* berdasarkan kemiripan pengucapan (*Phonetic String Matching*) menggunakan algoritma *Soundex*.

Algoritma *soundex* melakukan prosedur pencocokan string berdasarkan suara dari pelafalan kata [3]. Dan untuk mengukur tingkat kemiripan dari kata yang dicari menggunakan algoritma *jaro-winkler distance*. Algoritma ini dapat melakukan perhitungan tingkat kemiripan antara dua dokumen, teks atau kata.

## 1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan di atas, maka dapat diambil rumusan permasalahan pada tugas akhir ini adalah:

1. Bagaimana implementasi algoritma *soundex* dan *jaro-winkler distance* dalam sistem pencarian ayat Al-Quran dengan tulisan Arab yang dilatinkan untuk pencocokan *string* berdasarkan pengucapan dan pemeringkatan dokumen pada pencarian.
2. Bagaimana perhitungan akurasi pencarian kemiripan ayat Al-Quran yang dilatinkan berdasarkan algoritma *soundex* dan *jaro-winkler distance*.

## 1.3. Tujuan

Tujuan pembahasan yang ingin dicapai dari penelitian pada tugas akhir ini adalah:

1. Mengimplementasikan algoritma *Soundex* dan *jaro-winkler distance* dalam pembuatan sistem pencarian ayat Al-Quran dengan tulisan Arab yang dilatinkan untuk pencocokan *String* berdasarkan pengucapan dan pemeringkatan dokumen pada pencarian.
2. Menghitung akurasi pencarian kemiripan ayat Al-Quran yang dilatinkan berdasarkan algoritma *soundex* dan *jaro-winkler distance*.

## 1.4. Batasan Masalah

1. Menggunakan korpus teks Al-Quran dengan aksara Latin atau Al-Quran yang telah dilatinkan, begitu juga dengan *query* yang digunakan.
2. Menggunakan aturan alih aksara kritis Indonesia yg diatur dalam Surat Keputusan Bersama Menteri Agama dan Menteri P dan K Nomor 158 tahun 1987 – Nomor: 0543 b/u/1987.
3. Dataset yang digunakan adalah Al-Quran Latin Juz 1, Juz 2, Juz 3 dan Juz 30.
4. *Query* masukan sistem adalah *string* per kata berdasarkan variasi pengucapan bahasa Indonesia yang tidak dipengaruhi bahasa daerah.

## 1.5. Sistematika Penulisan

Sistematika penulisan laporan tugas akhir ini disusun sesuai rencana berikut:

### 1. BAB 1 PENDAHULUAN

Pada bab pertama ini menjelaskan gambaran secara umum tentang tugas akhir yang dikerjakan mengenai latar belakang masalah, batasan masalah, perumusan masalah, tujuan penelitian, manfaat penelitian, metodologi penelitian dan sistematika penulisan.

### 2. BAB 2 KAJIAN PUSTAKA

Pada bab kedua ini akan dijelaskan dasar-dasar teori yang mendukung penyelesaian tugas akhir, khususnya teori yang berkaitan dengan masalah penelitian sehingga dapat mendukung pengimplementasian algoritma serta pengujiannya terutama pada algoritma *Soundex* dalam pencocokan kata berdasarkan pengucapan.

### **3. BAB 3 PERANCANGAN SISTEM**

Pada bab ketiga ini dijelaskan perancangan yang diperlukan berdasarkan mekanisme dan batasan yang digunakan, sehingga dapat diimplementasikan serta direalisasikan sistem perangkat lunak yang akan dibuat sesuai dengan yang dibutuhkan.

### **4. BAB 4 PENGUJIAN DAN ANALISIS SISTEM**

Pada bab keempat ini akan dijelaskan hasil pengujian dari perancangan dan implementasinya mengenai parameter-parameter yang dibutuhkan sesuai metode yang digunakan dalam sistem yang telah dibangun pada bab sebelumnya.

### **5. BAB 5 PENUTUP**

Pada bab terakhir ini menyajikan kesimpulan terhadap hasil penelitian yang telah dilakukan berikut saran yang perlu dikemukakan dari hasil penelitian.