

# 1. PENDAHULUAN

## 1.1 Latar Belakang Masalah

Dunia Internet sangat diminati oleh banyak kalangan masyarakat. Banyak diantaranya yang mencari solusi, tutorial, pembahasan, atau diskusi lewat browsing di Internet. Latar belakang penulis mengambil judul tugas akhir ini adalah sulitnya mengelompokkan terjemahan ayat-ayat Al - Qur'an pada kehidupan sehari – hari, dikarenakan bentuk Al – Qur'an yang konvensional sulit untuk dipelajari. Ayat – ayat dalam Al – Qur'an disusun tidak berdasarkan sebuah permasalahan dan sifatnya yang berpecah, menyebabkan pencarian ayat – ayat berdasarkan kelompok tertentu membutuhkan waktu lama. Dengan begitu, penulis berencana membuat sebuah sistem untuk mengelompokkan kata-kata yang ada pada terjemahan ayat-ayat Al-Qur'an yang berhubungan dengan akidah. Sehingga sistem tersebut dapat mengelompokkan dan menganalisa pengaruh stemming pada clustering terjemahan ayat – ayat Al – Qur'an sebagai referensi dan solusi.

Dengan masalah diatas, maka digunakan sebuah proses yang disebut *text mining*. Salah satu proses *text mining* adalah *stemming*. *Stemming* digunakan untuk mengganti bentuk dari suatu kata menjadi kata dasar dari kata tersebut yang sesuai dengan struktur morfologi bahasa Indonesia yang baik dan benar. Pada algoritma *stemming* teks bahasa Indonesia yang ada selama ini mempunyai beberapa kendala, salah satunya adalah pengenalan affix serapan (Imbuan Asing). Oleh karena itu, penulis mencoba mengembangkan algoritma Porter, sehingga algoritma affix removal *stemming* ini mampu menyelesaikan beberapa permasalahan tersebut.

Pengelompokan informasi yang berkaitan dengan suatu kejadian tentu sulit dilakukan, bila hanya mengandalkan query biasa. Sebab pemilihan query yang kurang spesifik akan berakibat membanjirnya dokumen-dokumen yang tidak relevan. Oleh karena itu, penulis menggunakan proses *clustering* dalam pengelompokannya, dengan menggunakan *Algoritma Hierarchical Clustering* karena mudah diimplementasikan dan menghasilkan hirarki tree sehingga mempermudah pada proses analisa pengaruh *stemming* terhadap *clustering*. Pada tugas akhir ini, penulis melakukan analisa terhadap proses yang menggunakan *stemming* dan tidak menggunakan *stemming*, sehingga diharapkan menghasilkan kesimpulan analisa pengaruh *stemming* pada *clustering* dalam hal kualitas *cluster (Purity)* dan akurasi *stemmingnya*.

## 1.2 Perumusan Masalah

Berdasarkan uraian diatas, maka permasalahan yang muncul dan yang menjadi objek penelitian pada Tugas Akhir ini ialah:

- a. Bagaimana menerapkan algoritma *affix removal stemming* pada *clustering* yang menggunakan *Algoritma Hierarchical Clustering*?

- b. Merancang dan pembangun perangkat lunak untuk penerapan algoritma *affix removal stemming* pada *clustering*.
- c. Bagaimana menganalisis akurasi dan kualitas cluster (*Purity*) pengaruh algoritma *affix removal stemming* pada *clustering*?

Batasan masalah agar tidak meluasnya materi pembahasan dalam tugas akhir ini ialah:

- a. Terjemahan ayat-ayat Al-Qur'an yang diinputkan pada ayat yang berkaitan dengan akidah dalam teks bahasa Indonesia yang baku.
- b. Bahasa pemrograman yang digunakan adalah Visual Basic 6.0 (VB6) sebagai design *interface* dan *database* yang digunakan adalah Microsoft Office Access (MOA).
- c. Pada proses *clusteringnya*, menggunakan *Algoritma Hierarchical Clustering*, dengan kondisi berhenti proses *clusteringnya* ketika iterasi memiliki satu cluster.

### 1.3 Tujuan

Tujuan yang hendak dicapai dalam pembuatan Tugas Akhir ini yaitu : mengimplementasikan algoritma *affix removal stemming* dan *Algoritma Hierarchical Clustering* pada suatu aplikasi simulasi *Clustering and Stemming Sistem (CSS)* pada terjemahan ayat-ayat Al-Quran tentang permasalahan akidah.

### 1.4 Metode Penyelesaian Masalah

Metodologi yang digunakan untuk menyelesaikan masalah dalam Tugas Akhir ini ialah:

#### 1. Studi Literatur

Mempelajari sumber-sumber pustaka yang ada, yang dapat dijadikan referensi mengenai *text mining* khususnya algoritma *stemming*, proses *clustering* dan proses perbandingan kualitas cluster dan akurasi serta sumber-sumber lain yang terkait untuk menunjang penyelesaian tugas akhir ini. Sumber-sumber pustaka dapat berupa paper, buku, Al-Quran digital, maupun halaman web.

#### 2. Analisis dan Desain

Tahap ini meliputi analisis kebutuhan serta penyelesaian masalah untuk merancang perangkat lunak yang mempunyai fungsionalitas proses *stemming* dan *clustering* terjemahan ayat-ayat Al-Quran.

#### 3. Implementasi Sistem

Tahap ini meliputi pembangunan perangkat lunak yang telah dirancang pada tahap sebelumnya. Pembangunan perangkat lunak berbasiskan aplikasi desktop yaitu menggunakan aplikasi desktop dengan bahasa pemrograman Visual Basic (VB6).

4. Analisis dan Pengujian  
Melakukan pengujian perangkat lunak yang telah dikembangkan, dan kemudian menganalisis hasil performansi yang didapatkan. Pengukuran performansi adalah kualitas cluster (*Purity*) yang dihasilkan. Tujuan pengujian adalah untuk mengetahui hasil penerapan algoritma *affix removal stemming* pada *clustering*, maka juga dilakukan analisa akurasi *stemming*nya.
5. Penyusunan Laporan  
Hasil penelitian akan disusun menjadi suatu laporan yang meliputi aspek-aspek dalam penelitian yaitu teori, perancangan dan implementasinya, serta membuat kesimpulan dari hasil penelitian tersebut.

## 1.5 Sistematika Penulisan

Sistematika dari Penulisan Tugas Akhir ini adalah sebagai berikut :

### **BAB I      PENDAHULUAN**

Bab ini membahas mengenai latar belakang pembuatan tugas akhir ini, rumusan masalah yang akan di analisis, ruang lingkup masalah yang ada pada tugas akhir ini, tujuan dari pembuatan tugas akhir ini, metodologi pemecahan masalah serta sistematika penulisan dokumentasi.

### **BAB II     LANDASAN TEORI**

Bab ini terdiri dari teori-teori yang digunakan dalam mendukung dalam penyelesaian tugas akhir ini, dalam hal ini adalah penerapan *text mining* dan *information retrieval* yaitu algoritma *stemming* pada bahasa Indonesia dan *Algoritma Hierarchical Clustering* pada *clustering* terjemahan ayat – ayat Al Qur'an tentang permasalahan akidah.

### **BAB III    PERANCANGAN PERANGKAT LUNAK**

Bab ini berisi pengumpulan data analisis dan perancangan perangkat lunak yang terdiri dari perancangan struktur data, perancangan modul dan *interface*.

### **BAB IV    IMPLEMENTASI DAN PENGUJIAN**

Berisi tentang implementasi detil dari implementasi dan analisa pengaruh *affix removal stemming* terhadap *clustering* dengan studi kasus *clustering* terjemahan ayat – ayat Al-Qur'an. Pengujian terhadap sistem juga dibahas pada bab ini.

## **BAB V    PENUTUP**

Berisi mengenai kesimpulan dan saran-saran yang dapat diambil oleh penulis dari keseluruhan sistem yang telah dibuat untuk pengembangan tugas akhir ini.