

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Proses pengambilan data pada *search engine* yang masih memakai metode konvensional masih menyusahkan user dalam mencari dokumen-dokumen yang diperlukan. Hasil dari pencarian itu masih berisi dokumen yang tidak relevant yang masih harus diseleksi secara manual oleh *user*. Metode Konvensional hanya mengurutkan hasil pencarian berdasarkan *ranking*. Sehingga diharapkan dengan adanya *clustering*, akan memudahkan *user* dalam melakukan pencarian karena *clustering* akan mengelompokkan hasil pencarian.

Keanekaragaman dan pesatnya pertumbuhan jumlah dokumen dapat berpengaruh besar pada saat pencarian suatu dokumen. Suatu teknik Pencarian suatu dokumen yang ada pada saat ini merupakan dengan menampilkan hasil pencarian berurutan berdasarkan kecocokan kata atau kecocokan dokumen (*document ranking*). Pada referensi [1] menyatakan bahwa, terkadang hasil yang ditampilkan dengan pencarian berdasarkan *document ranking* terdiri dari beberapa halaman dan pengguna biasanya hanya melihat hasil pencarian maksimal pada 10 halaman pertama.

Pencarian dikelompokkan sesuai dengan kata yang diinputkan oleh pengguna. Dengan adanya pengelompokan dokumen tersebut sangat menguntungkan bagi pengguna yang mencari lebih dari satu dokumen yang diinginkan selain itu pengguna tidak perlu membuka banyak halaman karena dokumen hasil pencarian dikelompokkan berdasarkan kategori yang

mencerminkan isi dari suatu dokumen tersebut, sehingga pengguna dapat dimudahkan dalam pencarian dokumen yang diinginkan.

*Clustering* merupakan salah satu teknik dalam pengkategorian dokumen. Ide dasarnya merupakan dengan mengelompokkan dokumen-dokumen ke dalam grup-grup atau *cluster* berdasarkan kemiripan (*similarity*) antar dokumen, sehingga dokumen yang berhubungan dengan suatu topik tertentu ditempatkan pada *cluster* yang sama. Saat ini ada beberapa algoritma *clustering* yang dikenal salah satunya algoritma *Fuzzy C-Means Clustering*. [2] *Fuzzy C-Means (FCM)* merupakan salah satu metode *clustering* yang merupakan bagian dari *Hard K-Means*. *FCM* menggunakan konsep pendekatan *fuzzy* sehingga sebuah data dapat menjadi anggota dari semua *cluster* yang ada. Matriks keanggotaan terbentuk dengan tingkat keanggotaan masing-masing data pada setiap *cluster* yang ada yang bernilai antara 0 hingga 1. Tingkat keberadaan data pada suatu *cluster* ditentukan oleh derajat keanggotaannya. Dengan adanya pengelompokkan dokumen serupa terlebih dahulu diharapkan nantinya dapat meningkatkan performa kategorisasi dokumen teks serta lebih menghemat waktu komputasi untuk memilih k-tetangga terdekat dokumen. Hal ini dikarenakan pencarian k-dokumen terdekat hanya cukup dicari pada lingkup anggota kelompok dokumen yang serupa saja dan tidak perlu dibandingkan terhadap keseluruhan koleksi dokumen. [3]

Berdasarkan latar belakang tugas akhir ini maka akan dibangun aplikasi klasterisasi yang ditujukan untuk data hadits shahih. Diharapkan aplikasi yang dibuat akan lebih menghemat waktu dan memudahkan pencarian

informasi yang sesuai (relevan) dengan yang diinginkan oleh pengguna. Selain itu, penerapan algoritma *FCM* ini dapat terbilang baru mengenai kasus pencarian teks dan pencarian tersebut diproses dengan sebuah proses pengelompokan atau *clustering*. sehingga penelitian ini diwujudkan sebagai bahan penyusunan skripsi yang berjudul “*Implementasi Algoritma Fuzzy C-Means Pada Perangkat Lunak Bantu Pengklasteran Hasil Pencarian Hadits Shahih*”.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang ada maka dapat dirumuskan menjadi beberapa masalah diantaranya:

- a. Bagaimana menerapkan metode *FCM* untuk pengelompokan hasil pencarian hadits shahih?
- b. *Cluster* manakah yang memiliki anggota *cluster* (kelompok) yang terbaik pada pengelompokan hasil pencarian hadits shahih dengan *FCM* ?

## 1.3 Tujuan Penelitian

Adapun tujuan penelitian dari tugas akhir ini antara lain:

- a. Untuk menerapkan metode *FCM* dalam pengelompokan hasil pencarian hadits shahih.
- b. Untuk mengetahui kelompok manakah yang memiliki anggota terbaik pada pengelompokan hasil pencarian hadits shahih

## 1.4 Batasan Masalah

Berdasarkan permasalahan yang ada berikut ini merupakan batasan masalah dari tugas akhir ini yaitu:

- a. Data yang digunakan dalam penelitian ini berupa data hadits shahih tentang rukun islam.
- b. Dalam penelitian ini menggunakan domain hadits shahih untuk melakukan klasterisasi data.
- c. Dalam Penelitian ini menggunakan hadits shahih yang di riwayatkan oleh Sunan An-Nasa'i
- d. Klasterisasi data dengan menggunakan algoritma *FCM* dilakukan dengan mengelompokan hasil pencarian hadits shahih.

## 1.5 Metode Penelitian

Metode penelitian yang digunakan dalam penelitian tugas akhir ini yaitu:

### a. Metode Pengumpulan Data

Pada tahap ini pengumpulan data dibagi menjadi dua tahapan yaitu yang pertama, tahap observasi merupakan pengumpulan suatu data dengan pendataan langsung dengan menganalisis dan menelaah data-data yang telah ada sebelumnya. Kemudian tahap yang kedua, studi literatur merupakan suatu teknik atau cara dalam pengumpulan suatu data dengan mempelajari literatur, *internet*, buku-buku, dokumentasi, modul, jurnal dan segala sesuatu yang dirasa perlu untuk memperkuat konsep dan teori yang dapat membantu menyelesaikan permasalahan sedang dibahas.

### b. Metode Analisis Data

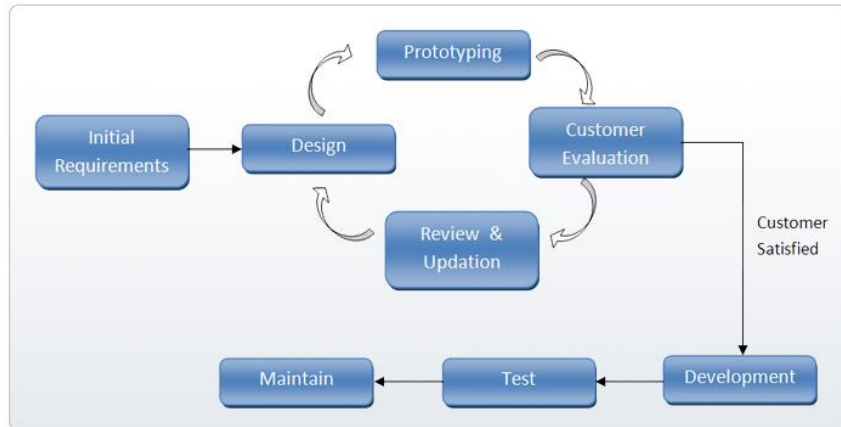
Adapun untuk menganalisis data dalam penerapan data mining ini menggunakan tahapan *Knowledge Discovery in Databases (KDD)* yang terdiri dari beberapa tahapan, yaitu *data selection*, *preprocessing*, *transformation*, *data mining*, dan *evaluation*.

## 1.6 Model Proses Pengembangan Perangkat Lunak

Metode Pengembangan Perangkat Lunak yang diterapkan dalam klasterisasi hasil pencarian menggunakan *Prototype Model*. Metode *Prototype* merupakan suatu paradigma baru dalam metode pengembangan perangkat lunak dimana metode ini tidak hanya sekedar evolusi dalam dunia pengembangan perangkat lunak, tetapi juga merevolusi metode pengembangan perangkat lunak yang lama yaitu sistem sekuensial yang biasa dikenal dengan nama *SDLC* atau *waterfall development model*.

Dalam Model *Prototype*, *prototype* dari perangkat lunak yang dihasilkan kemudian dipresentasikan kepada pelanggan, dan pelanggan tersebut diberikan kesempatan untuk memberikan masukan sehingga perangkat lunak yang dihasilkan nantinya betul-betul sesuai dengan keinginan dan kebutuhan pelanggan. [4].

Metode ini menyediakan suatu konsep yang lengkap dari suatu sistem perangkat lunak, terdiri dari model kertas, model kerja dan program. Pengembang akan menganalisa kebutuhan pengguna, menganalisa sistem, melakukan studi kelayakan dan melakukan studi terhadap kebutuhan pemakai, meliputi model *interface*, teknik prosedural dan teknologi yang akan digunakan.



**Gambar 1. 1** *Prototype Model*

Berikut ini merupakan Tahapan-tahapan Proses Pengembangan Perangkat Lunak dengan *Prototype* yaitu :

**a. Pengumpulan Kebutuhan**

Pelanggan dan pengembang bersama-sama mendefinisikan format seluruh perangkat lunak, mengidentifikasi semua kebutuhan, dan garis besar sistem yang akan dibuat.

**b. Membangun *Prototyping***

Membangun *prototyping* dengan membuat perancangan sementara yang berfokus pada penyajian kepada pelanggan (misalnya dengan membuat input dan format output).

**c. Evaluasi *Prototyping***

Evaluasi ini dilakukan oleh pelanggan, apakah *prototyping* yang sudah dibangun sudah sesuai dengan keinginan pelanggan atau belum. Jika sudah sesuai, maka langkah selanjutnya akan diambil. Namun jika tidak, *prototyping* direvisi dengan mengulang langkah-langkah sebelumnya.

**d. Pengkodean Sistem**

Dalam tahap ini *prototyping* yang sudah di sepakati diterjemahkan ke dalam bahasa pemrograman yang sesuai.

**e. Pengujian Sistem**

Setelah sistem sudah menjadi suatu perangkat lunak yang siap pakai, kemudian dilakukan proses Pengujian. Pengujian ini dilakukan dengan *White Box, Black Box, Basis Path*, pengujian arsitektur, dll.

**f. Evaluasi Sistem**

Pelanggan mengevaluasi apakah perangkat lunak yang sudah jadi sudah sesuai dengan yang diharapkan. Jika ya, maka proses akan dilanjutkan ke tahap selanjutnya, namun jika perangkat lunak yang sudah jadi tidak/belum sesuai dengan apa yang diharapkan, maka tahapan sebelumnya akan diulang.

**g. Penggunaan Sistem**

Perangkat lunak yang telah diuji dan diterima pelanggan siap untuk digunakan.

**1.7 Sistematika Penulisan**

Sistematika penulisan dari tiap bab dalam laporan tugas akhir ini bertujuan untuk mendapatkan keterarahan dalam penulisan sehingga mudah dipahami, adapun sistematika secara umum dari penulisan laporan ini merupakan:

**BAB I : PENDAHULUAN**

Pada BAB I merupakan pengantar yang memberikan gambaran mengenai permasalahan permasalahan yang kemudian akan dibahas pada bab selanjutnya. Terdapat tujuh pokok bahasan dalam bab ini, yaitu latar belakang, perumusan masalah, batasan

masalah, tujuan, metode penelitian, model proses perangkat lunak dan sistematika penulisan.

## **BAB II : STUDI PUSTAKA**

Pada BAB II berisi tentang teori-teori yang digunakan untuk menganalisa pemecahan masalah yang telah dirumuskan serta memaparkan teori-teori yang digunakan dalam perancangan dan implementasi.

## **BAB III : ANALISIS DAN PERANCANGAN**

Pada BAB III menjelaskan tentang analisis permasalahan yang telah dirumuskan dan menganalisis suatu kebutuhan yang diperlukan dalam menyelesaikan permasalahan tersebut. Selain itu pada BAB ini juga dijelaskan tentang pembuatan *desain* dari sistem dengan mengacu pada analisis yang telah dibahas. Desain sistem yang akan dijelaskan terbagi menjadi tiga bagian, meliputi desain *user interface*, desain data dan desain proses.

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

Pada BAB IV berisi tentang detail spesifikasi aplikasi, kebutuhan aplikasi, implementasi aplikasi, dan pengujian yang dilakukan terhadap aplikasi yang dibangun.

## **BAB V : PENUTUP**

Bab V berisi kesimpulan dan saran untuk pengembangan aplikasi lebih lanjut dalam upaya memperbaiki kelemahan pada aplikasi guna untuk mendapatkan hasil kinerja aplikasi yang lebih baik dan pengembangan program selanjutnya.