

BAB I

PENDAHULUAN

1.1 Latar Belakang

Hadits secara etimologi memiliki arti al-jadid (sesuatu yang baru), kebalikan dari al-qadim. Kata hadits juga bermakna al-khbar (berita), yaitu sesuatu yang dibicarakan dan dipindahkan dari satu orang ke orang lain[1].

Adapun terminologi, merujuk pada pendapat ulama hadits sendiri terdapat beberapa terjemahan yang berbeda. Perbedaannya adalah tentang hal-hal atau ciri-ciri Nabi sebagai hadits dan ada yang mengatakan tidak termasuk hadits. Ada yang menyebut *Taqrir* Rasulullah secara eksplisit sebagai bagian dari jenis hadits dan ada yang secara implisit memasukkannya ke dalam *aqwal* atau perbuatannya. Dari berbagai perbedaan tersebut, hadits merupakan sesuatu yang berdasarkan Nabi Muhammad Shallallahu 'alaihi wasallam, para sahabat, dan *tabiin* yang dapat dijadikan hukum *syara'*[1].

Hadits merupakan sumber hukum agama Islam kedua setelahnya Al-Qur'an, yaitu seperti yang sudah dijelaskan sebelumnya. Dalam hadits, ada banyak bab yang membahas beberapa kasus. Misalnya bab tentang iman, ilmu, salat, wudu, haji, puasa, dan lebih banyak. Adapun dalam tugas akhir ini digunakan hadits yang bertemakan tentang salat, wudu, iman dan akhlak yang diambil dari *kutubut tis'ah* karangan 9 imam dimana untuk total hadits dari imam Muslim dan Abu Daud saja terdapat 12,075 hadits[2].

Hadits yang didapat dari *kutubut tis'ah* dapat dijadikan sebagai dasar untuk dilakukannya *data mining* klusterisasi. *Data Mining* adalah proses ekstraksi

informasi yang bermanfaat, pola, pengetahuan, atau wawasan yang tersembunyi dan berharga dari kumpulan data yang besar dan kompleks. Tujuan dari data mining adalah untuk mengidentifikasi hubungan yang tidak terlihat secara langsung atau untuk mendapatkan pemahaman yang lebih dalam tentang data dengan memanfaatkan teknik dan algoritma analisis yang canggih[3]. Sedangkan Klasterisasi adalah salah satu teknik dalam data mining yang bertujuan untuk mengelompokkan data menjadi kelompok-kelompok atau klaster-klaster yang memiliki kemiripan dalam karakteristik tertentu. Tujuan utama klasterisasi adalah untuk mengidentifikasi pola atau struktur dalam data yang mungkin tidak terlihat secara langsung. Dengan adanya *data mining* terhadap data hadits, hadits dapat dikategorikan atau di kelompokkan.

Untuk memaksimalkan proses klastering sebuah sistem menggunakan algoritma sebagai pemecah masalah yang akan menjadikan sebuah sistem yang lebih sederhana. Algoritma ialah urutan sistematis dan valid dalam memecahkan masalah atau mencapai tujuan[4]. Sementara itu pemilihan metode *K-Means* dikarenakan metode ini harus menggunakan data fisik tidak abstrak dan bersifat jelas[5], hal ini sesuai dengan data yang akan digunakan pada permasalahan di dalam pengelompokan hadits dengan menggunakan data hadits dari *kutubut tis'ah* sebagai acuan datanya. Algoritma *K-Means* adalah salah satu algoritma klasterisasi yang paling umum digunakan untuk mengelompokkan data menjadi kelompok-kelompok berdasarkan kesamaan atribut atau fitur tertentu. Algoritma ini bekerja dengan cara mencari titik pusat (*centroid*) dari setiap kelompok dan mengelompokkan data ke kelompok yang memiliki *centroid* terdekat[6].

Berdasarkan penelitian sebelumnya oleh Miftachur Robani, Achmad Widodo dengan judul *Algoritma K-Means Clustering Untuk Pengelompokan Ayat Al Quran Pada Terjemahan Bahasa Indonesia* penelitian ini menggunakan beberapa percobaan dan nilai *Silhouette Coefficient* untuk surah Al-baqarah sebesar 0,3777. Akurasi *Cluster* tidak maksimal karena pemilihan pusat awal kluster yang acak[6]. Penelitian lainnya oleh Nurhayati Buslim, Rayi Pradono Iswara dan Fajar Agustian tahun 2020[7] melakukan pemodelan dari data *mustahiq* dengan algoritma k means dan *big data analysis* dengan mendapatkan data *Mustahiq* dengan sampel acak dari survei online dan offline, menghasilkan simulasi dari total 1109 koresponden, 300 koresponden termasuk dalam kluster *Mustahiq* dan 809 koresponden. koresponden termasuk dalam cluster *Non-Mustahiq* dan memiliki tingkat akurasi sebesar 83,40%.

Berdasarkan referensi di atas, penelitian ini akan secara otomatis mengelompokkan hadits berdasarkan tema isi hadits yang sudah ditentukan yaitu, hadits yang bertemakan salat, wudu, iman dan akhlak yang dari semua itu terdapat 159 hadits yang digunakan. Selain itu, penelitian ini juga berfokus pada pembelajaran sistem menggunakan algoritma *K-Means* dan menggunakan data teks (terjemahan hadits bahasa indonesia). Ada beberapa parameter yang akan difokuskan pada penelitian ini, diantaranya jumlah cluster, nomor iterasi, jumlah *centroid*, jumlah hadits yang digunakan dan kombinasi tingkat pengujian *cluster*.

Namun Algoritma *K-Means* juga mempunyai kelemahan seperti model Algoritma Kluster yang lain, yaitu jumlah kluster harus diketahui: Anda harus menentukan jumlah kluster sebelum menjalankan algoritma, yang mungkin sulit jika Anda tidak memiliki pengetahuan sebelumnya tentang data. Jumlah kluster

yang salah dapat menghasilkan hasil yang tidak bermakna[8]; Sensitif terhadap *Outlier*: Data outlier atau data yang sangat berbeda dapat mempengaruhi posisi centroid dan mengganggu pembentukan kluster yang akurat[9]; Non-*Deterministik*: *K-Means* tidak menjamin solusi yang sama setiap kali dijalankan karena sensitivitas terhadap posisi awal. Ini dapat membuat hasil klasterisasi tidak konsisten pada percobaan yang berbeda. Meskipun memiliki kelemahan ini, *K-Means* tetap menjadi algoritma klasterisasi yang populer dan efektif dalam banyak kasus ketika asumsi dan batasan yang terkait dengannya sesuai dengan karakteristik data.

Ada variasi dan modifikasi *K-Means* yang telah dikembangkan untuk mengatasi beberapa dari kelemahan di atas dan memilih algoritma yang paling cocok tergantung pada konteks dan karakteristik data. Seperti algoritma *K-Means++* dimana perlunya ditentukan iterasi dan centroid sehingga dapat menghasilkan data *cluster* yang tidak berubah ubah dan stabil ketika di jalankan berulang kali. Oleh karena itu, dengan adanya sistem ini diharapkan sistem dapat mempelajari data hadits terjemahan Indonesia dengan *cluster* dan tema hadits yang sudah ditentukan. Maka dilakukanlah penelitian dengan judul **Interpretasi Hasil Klasterisasi Hadits Shahih Menggunakan Variasi Algoritma *K-Means***.

1.2 Perumusan Masalah

Perumusan masalah dibuat agar penelitian menjadi lebih mudah dan terarah. Berdasarkan isi dari latar belakang di atas, penulis merumuskan rumusan masalah

1. Bagaimana mengimplementasikan metode Variasi Algoritma *K Means* untuk mengelompokkan hadits shahih?
2. Bagaimana kinerja atau performa Algoritma *K-Means* dalam klasterisasi Hadits?

1.3 Tujuan Penelitian

Selanjutnya, secara keseluruhan tugas akhir ini bertujuan untuk:

1. Untuk menerapkan algoritma *K-Means* dalam pengelompokkan hadits shahih.
2. Untuk mengetahui bagaimana kinerja Algoritma *K-Means* dalam klasterisasi hadits.

1.4 Batasan Masalah

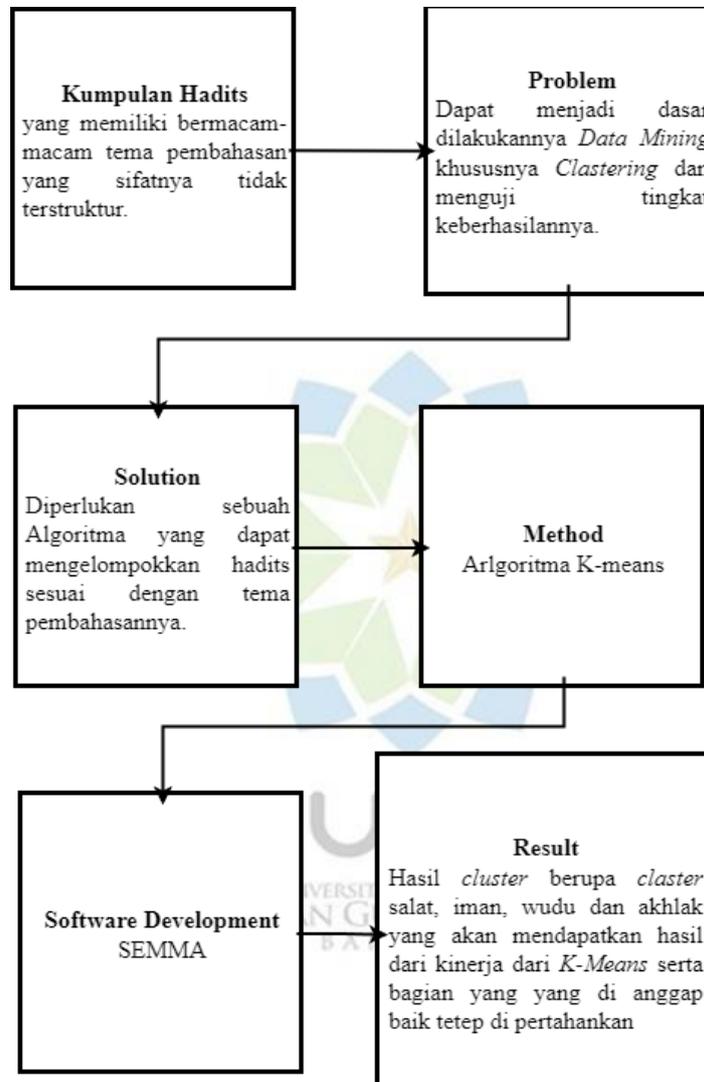
Merujuk pada luasnya pembahasan dan pengembangan yang dapat ditemukan pada perumusan masalah di atas, oleh karena itu harus adanya batasan masalah mengenai fitur apa saja yang akan dibuat dan diselesaikan dalam penelitian ini.

Keterbatasan ini yaitu:

1. Hadis yang digunakan dalam penelitian ini hanya berupa hadits Shahih yang bersumber dari Ensiklopedia Hadits 9 Imam karangan sembilan imam[10]. Dimana dalam pelaksanaan pengumpulan hadits juga menggunakan hadits *hasan* dikarenakan harus memeriksa satu per satu hadits dan banyak diantaranya hadits *hasan* yang masih bisa digunakan sebagai sumber ajaran.
2. Menggunakan hadits yang bertemakan, iman, akhlak, sholat dan wudu yang digunakan pada penelitian kali ini sebanyak 159 hadits.
3. Mengelompokkan menjadi 4 clusters dengan menggunakan variasi algoritma *k means*,
4. Hadits yang digunakan menggunakan terjemahan Bahasa Indonesia.
5. Menggunakan Bahasa baku yang sesuai dengan KBBI (Kamus Besar Bahasa Indonesia).

1.5 Kerangka Pemikiran

Kerangka pemikiran pada penelitian ini divisualisasikan dan diuraikan pada gambar 1.1 di bawah ini:



Gambar 1.1 Kerangka Pemikiran Penelitian

1.6 Metodologi Pengerjaan Tugas Akhir

Untuk mengatasi permasalahan yang merujuk pada tujuan pembuatan penelitian ini, maka metodologi pengerjaan yang digunakan adalah sebagai berikut:

1.6.1 Tahap Pengumpulan Data

1. Studi Literatur

Mengumpulkan data dengan cara mengumpulkan literatur, jurnal, paper dan bacaan yang berhubungan dengan judul tugas akhir.

2. Observasi

Metode mengumpulkan data dilakukan dengan melakukan penelitian dan observasi langsung terhadap permasalahan yang diambil.

3. Wawancara

Pada tahap wawancara, mengumpulkan data dengan orang yang berkompeten dalam bidangnya atau narasumber untuk memperoleh data yang sesuai dengan tujuan atau dengan pengguna yang memiliki keluhan yang sama.

1.6.2 Metode Pengembangan Perangkat Lunak

Penelitian kali ini menggunakan metode SEMMA dalam mengolah data yang alurnya terdiri dari *sampel* (pengumpulan data), *explore* (deskripsi data), *modify* (perubahan data), *model* (pemodelan data) dan *assess* (evaluasi data).

1.7 Sistematika Penulisan

Sistematika penulisan laporan tugas akhir dibawah ini merupakan gambaran umum dari penelitian yang akan dilakukan, oleh karena itu penulisan laporan akan disusun sebagai berikut:

BAB I PENDAHULUAN

Menjelaskan secara umum tentang latar belakang masalah, rumusan masalah, tujuan penelitian, batasan masalah, metodologi penelitian, dan sistematika penulisan.

BAB II KAJIAN LITERATUR

Bab ini membahas landasan teori dan teori pendukung peneliti terkait dengan penelitian baik perancangan dan pengembangan serta implementasi sistem pada tugas akhir ini.

BAB III METODOLOGI PENELITIAN

Pada bab ini menjelaskan tahapan-tahapan pemodelan data SEMMA dari “Interpretasi Hasil Klasterisasi Hadits Shahih Menggunakan Variasi Algoritma *K-Means*”.

BAB IV HASIL DAN PEMBAHASAN

Bab empat tentang implementasi sistem dan pengujian sistem berisi hasil sistem yang kemudian dievaluasi (*assess*) dan disebarluaskan.

BAB V SIMPULAN DAN SARAN

Bab ini berisi kesimpulan perihal pengembangan sistem yang dibuat, dan saran yang diajukan untuk perbaikan perangkat lunak.

DAFTAR PUSTAKA

Daftar pustaka memuat sumber-sumber tertulis yang digunakan dalam penelitian ini.