

BAB I

PENDAHULUAN

1.1 Latar Belakang

Hadits merupakan sumber ajaran islam yang ke dua setelah Al-Qur'an, yang dimana hadits merupakan pedoman serta tuntunan bagi seluruh umat islam sebagai panduan untuk melakukan seluruh aktivitas, baik itu untuk ibadah, budi pekerti, sosialisasi dalam masyarakat, dan lain sebagainya. Hadits adalah sikap serta perilaku Nabi Muhammad SAW dalam kehidupan sehari-hari. Dimana tidak terlepas dari tuntunan Allah SWT yang dijelaskan dalam Al-Qur'an, sudah sepantasnya kita menjadikan Nabi Muhammad SAW sebagai suri tauladan bagi umat manusia. Namun, tidak jarang pemahaman pada suatu hadits sering menyesatkan, padahal fungsi dari hadits sebagai pembenaran hukum yang berfungsi untuk kehidupan seluruh manusia setelah Al-Qur'an [1].

Abu Hanifah meletakkan hadits Mutawatir sebagai hadits tertinggi karena diyakini kebenarannya tanpa sikap skeptis dalam hal melihat validasi dari suatu hadits. Hal ini dipengaruhi dari jumlah kuantitas (*al-kamm*) perawi yang banyak, dan ke '*adalaha-annya*, disertai pada tempat kejadian turunnya (*makanu al-wurud*) hadits tersebut. Sehingga suatu hal yang ditetapkan secara Mutawatir akan menghasilkan suatu ilmu yang pasti (*al-ilmu al-dlarury*) yang sebagaimana seseorang melihat kejadian tersebut secara langsung (*al-mu'ayanah*). Hadits akan mencapai derajat Mutawatir, apabila hadits tersebut telah disepakati serta diterima secara *aklamatif* oleh seluruh umat tentang keabsahannya [2].

Hadits Mutawatir memiliki banyak pembahasan dalam berbagai perkara, misalnya pembahasan tentang shalat, ilmu dan lain sebagainya. Dalam sumber asalnya penggunaan hadits dengan menggunakan bahasa Arab, seiring perkembangan jaman telah banyak terdapat hadits dengan terjemahan terutama terjemahan hadits kedalam bahasa Indonesia. Dari banyaknya pembahasan dalam hadits akan sangat menarik untuk dapat dipadukan dengan teknologi serta teknik data mining terutama dalam teks mining yang berfungsi untuk melakukan pengelompokan hadits berdasarkan isi dari pembahasan hadits yang bersangkutan. Pada referensi penelitian sebelumnya, masih jarang ditemukan penelitian mengenai

pengelompokan hadits Mutawatir berdasarkan Terjemah hadits dalam bahasa Indonesia. Adapun beberapa aplikasi tentang hadits yang mengatur pengelompokan hadits masih dilakukan secara manual melalui relasi *database*, yang dimana aplikasi tersebut blum ter-otomatis menggunakan metode atau algoritma tertentu.

Metode yang sudah digunakan dalam melakukan pengelompokan pada suatu data salah satunya dengan metode *clustering*. Menurut Larose (2015) *clustering* adalah proses pada pengelompokan record, observasi, ataupun pengelompokan pada kelas yang memiliki kesamaan pada objek. Perbedaan *clustering* dengan klasifikasi adalah tidak adanya variable target dalam proses melakukan pengelompokan pada proses *clustering*. Pada *clustering* tidak jarang dilakukan sebagai tahap awal dalam proses data mining. Karena kualitas dari hasil proses *clustering* sangat bergantung dengan metode yang dipakai. Pada *clustering* juga harus dapat mengukur kemampuannya sendiri dalam usaha menentukan suatu pola yang tersembunyi pada suatu data yang sedang diteliti. Terdapat berbagai macam metode yang bisa digunakan dalam mengukur nilai kesamaan antar objek-objek yang dibandingkan, salah satunya dengan metode *Euclidean Distance*, yang dapat menghitung jarak dua buah point dengan mengetahui nilai masing-masing atribut pada kedua point tersebut [3].

Algoritma K-Medoids atau dikenal juga dengan PAM (*Partitioning Around Medoids*) merupakan metode partisi *Clustering* untuk mengelompokkan sekumpulan n objek ke dalam k cluster. Pengelompokan algoritma ini menggunakan *medoid* sebagai pusat dari *cluster*. Medoids yaitu objek yang letaknya terpusat di dalam sebuah *cluster*[4]. Algoritma K-Medoids lebih baik dibandingkan dengan algoritma K-Means karena pada algoritma K-Medoids dapat menemukan k sebagai objek yang representatif dalam meminimalkan jumlah ketidak samaan objek pada data, sedangkan algoritma K-Means menggunakan jumlah jarak *euclidean distance* sebagai objek data[5]. Algoritma K-Medoids muncul sebagai penanggulangan dari kelemahan algoritma K-Means yang sensitif terhadap *outlier* dikarenakan sebuah objek dengan nilai yang besar secara substansi menyimpang dari distribusi data[6]. Rujukan penelitian sebelumnya [3][7][8].

Pada penelitian [3] dilakukan pengujian dengan menggunakan algoritma K-Medoids yang merupakan sebuah metode *partisional clustering* dimana memiliki fungsi untuk melakukan pengelompokan data kedalam satu set *k-cluster* diantara data yang memiliki kemiripan satu dengan yang lain. untuk mengelompokkan penerimaan mahasiswa baru di Universitas Kanjuruhan Malang yang tersebar dari berbagai jurusan. Algoritma K-Medoids yang berfungsi menemukan *k-cluster* diantara data yang mencirikan objek didalam suatu data, yang dimana pada penelitian ini memiliki tiga *cluster*, proses pengujian didapat nilai *Silhouette Coefficient* terbaik yaitu 0,690754 dari jumlah *cluster* sebanyak tiga dengan jumlah data sebanyak 15.

Pada penelitian [7] melakukan perbandingan algoritma *clustering* yaitu K-medoids dengan K-Means dalam mengelompokkan data set dari kasus penyandang cacat di provinsi Riau Klaster yang dihasilkan pada penelitian ini berjumlah tiga *cluster*. Hasil dari penelitian ini menunjukkan nilai validitas yang dihasilkan oleh algoritma K-Medoids ialah sebesar 0.5009. Sedangkan hasil dari nilai validitas dari algoritma K-Means yaitu 0.1443. dari hasil tersebut membuktikan bahwa algoritma K-Medoids lebih baik dalam melakukan proses klasterisasi pada data sebaran Anak Cacat dibandingkan dengan metode K-Means.

Dan pada penelitian [8] dilakukan penelitian dengan menggunakan algoritma K-Medoids dalam pengelompokan Berita Olahraga bertujuan untuk supaya berita tersebut menjadi suatu kelompok yang sesuai dengan kesamaan berita. Hasil dari penelitian ini menunjukkan bahwa nilai akurasi yang didapat sebesar 0.584, sebanyak 5 *cluster* yang diinput. Kemudian dilakukan perbaikan metode dengan cara menambahkan metode *cosine similarity* dapat bisa meningkatkan nilai akurasi yang dari nilai semula memiliki rata-rata 0,44 bertambah secara optimal menjadi 0,55.

Dari rujukan penelitian diatas membuktikan bahwa terdapatnya hasil perbedaan akurasi dari penelitian pengelompokan menggunakan tipe data yang berbeda. Karena cukup berpengaruh pada penentuan pusat *cluster* awal serta derajat keanggotaan dari hasil *cluster* yang pada dasarnya dilakukan dengan cara acak. Pada penelitian ini akan mencoba mengelompokkan hadits Mutawatir secara

otomatis dengan menggunakan algoritma K-Medoids berdasarkan kemiripan matan hadits dengan menggunakan data teks yaitu Hadits terjemahan bahasa Indonesia. Atribut yang akan dipakai yaitu matan atau isi dari hadits. Dimana isi hadits tersebut akan melalui proses *Text Preprocessing*, pembobotan kata, kemudian proses *clustering* pada proses ini akan di kelompokkan berdasarkan kemiripan isi topiknya kedalam beberapa *cluster* kemudian akan dihitung berapa nilai akurasi ketepatan *cluster* nya.

Terdapat beberapa parameter yang dapat difokuskan pada penelitian ini, yaitu parameter yang bisa mempengaruhi hasil dari *cluster* serta akurasi dari algoritma K-Medoids yang diantaranya jumlah *cluster*, awal dari derajat pada keanggotaan, serta pengujian awal pada awal *cluster* serta awal dari derajat keanggotaan. Pada penelitian ini diharapkan adanya pertimbangan pada parameter diatas agar dapat diketahui factor apa saja yang dapat mempengaruhi hasil dari *clustering* data agar menjadi lebih baik.

Maka dari itu, penulis tertarik untuk melakukan sebuah penelitian analisis dengan judul “IMPLEMENTASI ALGORITMA K-MEDOIDS PADA *CLUSTERING* HADITS MUTAWATIR”.

1.2 Rumusan Masalah

Berdasarkan latar belakang di atas, maka didapatkan rumusan masalah sebagai berikut:

1. Bagaimana menerapkan Algoritma K-Medoids pada *clustering* hadits Mutawatir?
2. Bagaimana kinerja Algoritma K-Medoids pada *clustering* hadits Mutawatir?

1.3 Tujuan Tugas Akhir

1. Menerapkan Algoritma K-Medoids pada *clustering* pada hadits Mutawatir
2. Mengetahui kinerja Algoritma K-Medoids pada *clustering* hadits Mutawatir.

1.4 Batasan Masalah

Mengingat luasnya pembahasan dan perkembangan yang ditemukan dalam permasalahan diatas, maka perlu adanya batasan-batasan masalah mengenai apa yang akan dibuat dalam penelitian ini. Batasan-batasan tersebut yaitu:

1. Data yang digunakan berupa teks (data tidak terstruktur) Bahasa Indonesia
2. Data yang akan digunakan untuk pengelompokan yaitu isi (matan) berbahasa Indonesia Bab hadits yang digunakan untuk percobaan yaitu bab Shalat dan Ilmu
3. Metode *Text Preprocessing* menggunakan *Case Folding* (meratakan kata), *Tokenizing* (Pemisah kata), *Stopword* (penghapus kata), *Stemming* (Perubah kata dasar), serta TF-IDF (pembobotan kata)
4. Penelitian ini menggunakan bahasa pemrograman Python untuk mempermudah perhitungan klasterisasi dan implementasi web.
5. Metode pengembangan perangkat lunak menggunakan *Rapid Application Development (RAD)*
6. Pengujian hasil *cluster* menggunakan *Sillhouette Coefficient*

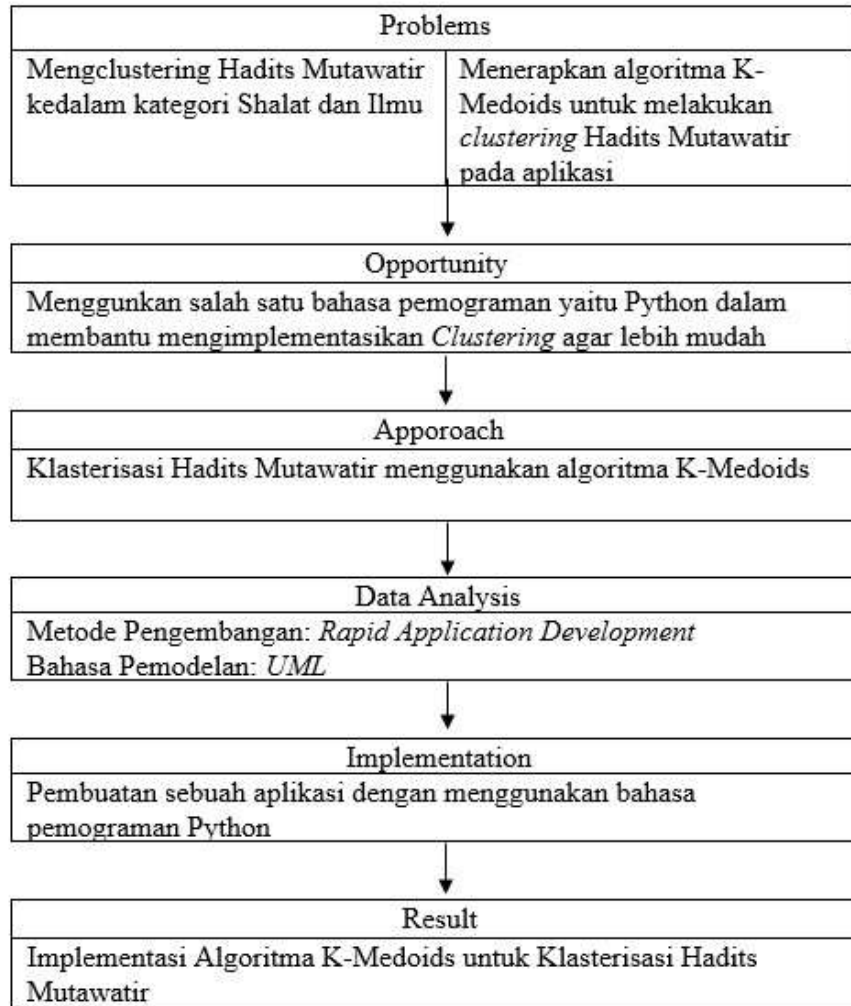
1.5 Manfaat Penelitian

Manfaat yang diharapkan pada penelitian ini adalah:

1. Mengelompokan hadits secara otomatis berdasarkan isi hadits.
2. Menganalisis kinerja K-Medoids terhadap data teks hadits.
3. Memberikan gambaran mengenai metode klasterisasi dengan metode K-Medoids.
4. Menambahkan rekomendasi scenario percobaan untuk dimanfaatkan dalam pengujian lain.
5. Sebagai referensi bagi peneliti lain yang ingin mengembangkan metode klasterisasi.

1.6 Kerangka Pemikiran

Diperlukannya sebuah kerangka pemikiran dalam pembuatan sebuah aplikasi untuk menggambarkan secara global alur logika pada suatu penelitian. Adapun kerangka pemikiran dari penelitian ini digambarkan pada Gambar 1.1.



Gambar 1.1 Kerangka Pemikiran

1.7 Metode Penelitian

Dibutuhkannya metode penelitian karena untuk memperoleh data yang dibutuhkan untuk melakukan sebuah penelitian, metode yang diperlukan sebagai berikut:

1.7.1 Pengumpulan Data

Data yang digunakan pada penelitian ini berupa data teks hadits Mutawatir berbahasa Indonesia, yang dimana diperlukan metode berikut:

1. Metode Keputusan

Yang dimaksud dengan metode keputusan adalah sebuah metode yang digunakan dalam memperoleh konsep teoritis dengan melakukan pencarian data melalui jurnal-jurnal ataupun buku-buku yang bisa membantu proses pemecahan masalah yang sama. didalam penelitian ini menggunakan kitab karangan Imam Jalaluddin as-Syuthi yang berjudul Hadits-hadits Mutawatir.

2. Web Scraping

Yang dimaksud Web Scraping adalah sebuah teknik untuk mendapatkan sebuah informasi yang bersumber dari website secara otomatis tanpa diharuskan menyalin secara manual. Website yang digunakan pada penelitian ini adalah hadits.net yang merupakan website hadits 9 imam dengan terjemah lengkap Bahasa Indonesia.

1.8 Metode Pengembangan Data Mining

Pada proses data mining dalam penelitian ini menggunakan metode CRISP-DM. CRISP-DM adalah standarisasi pada proses data mining yang digunakan untuk strategi dalam pemecahan sebuah masalah secara umum dari bisnis ataupun unit penelitian. Dalam CRISP-DM dalam sebuah proyek dari data mining mempunyai siklus hidup yang terbagi dalam 6 fase yaitu [9] :

1. *Business Understanding Phase* atau Fase Pemahaman Bisnis.
2. *Data Understanding Phase* atau Fase dalam pemahaman data.
3. *Data Prepartion Phase* atau Fase dalam pengolahan dari sebuah data.
4. *Modeling Phase* atau Fase dalam pemodelan.
5. *Evaluation Phase* atau Fase Evaluasi.
6. *Deployment Phase* atau Fase Penyebaran.

1.9 Sistematika Penulisan

Sistematika dalam penulisan sebuah perangkat lunak ini dibagi dalam lima bab, yang disusun sebagai berikut:

BAB I PENDAHULUAN

Dalam bab ini menjelaskan bagaimana tentang Latar Belakang Masalah, Rumusan Masalah, Tujuan dari Penelitian, Batasan Masalah, Manfaat Penelitian, Kerangka Pemikiran, Metode Penelitian, serta Sistematika Penulisan.

BAB II STUDI PUSTAKA

Bab ini menjelaskan bagaimana tahapan dari sebuah proses pengerjaan sebuah perangkat lunak dan dalam bab II juga menjelaskan teori-teori yang relevan yang berguna dan dapat digunakan dalam penelitian ini.

BAB III ANALISIS DAN PERANCANGAN

Bab ini menjelaskan bagaimana analisis serta kebutuhan sistem serta perancangan dari sebuah sistem dimulai, dimulai dari perancangan umum, perancangan prosedur, perancangan kode, serta perancangan basis data.

BAB IV IMPLEMENTASI APLIKASI

Dalam bab ini menjelaskan bagaimana implementasi dari sebuah perangkat lunak yang sudah dianalisa dan sudah dirancang sebelumnya serta hasil dari pengujian sistem yang dimana memperlihatkan berjalan tidaknya program yang telah dibuat.

BAB V KESIMPULAN DAN SARAN

Dalam bab ini berisi kesimpulan yang didapat dari permasalahan dan saran yang berguna untuk pengembangan tugas akhir ini.