

BAB I

PENDAHULUAN

1.1.Latar Belakang Masalah

Perkembangan *aplikasi* web yang semakin pesat sejak munculnya teknologi internet sangat membantu dalam kemudahan serta kecepatan pengiriman, penyampaian dan penerimaan informasi. Mulai dari perusahaan perusahaan, sekolah-sekolah, perguruan tinggi dan lembaga atau organisasi lainnya telah banyak memanfaatkan aplikasi web dalam kegiatan penjualan, promosi, belajar dan kegiatan lainnya dimana dibutuhkan pengiriman penyebaran dan penerimaan informasi sehingga memberikan kemudahan bagi pengguna (*user*) yang membutuhkan.

Dengan banyaknya pemakaian komputer serta penggunaan *internet* yang sangat diperlukan, masyarakat bisa mendapatkan informasi dan saling berkomunikasi dengan cepat dan tepat. Permasalahan yang ada pada saat ini bagi para masyarakat umum khususnya para pelajar kesulitan dalam mencari informasi *ensiklopedia* hadist dari media berupa buku atau kitab.

Banyak hadist pada kitab tertentu dibutuhkan sebuah media yang mampu menghimpun hadits tersebut. Media tersebut salah satunya dapat berupa ensiklopedi. Menurut Kamus Besar Bahasa Indonesia Ensiklopedi adalah karya *universal* yang menghimpun uraian tentang berbagai cabang ilmu atau bidang ilmu tertentu dalam artikel- artikel terpisah dan tersusun menurut abjad atau kategori tertentu. *Ensiklopedi* bertujuan untuk memberikan informasi dan pengetahuan berupa data yang telah diolah sedemikian rupa sehingga mampu mencakup sebuah bidang ilmu secara mendalam.

Banyak *ensiklopedia* yang dihimpun berupa bentuk buku yang didalamnya terdapat kumpulan-kumpulan hadist bahkan sampai dengan ribuan hadist. Ketebalan buku ensiklopedi tersebut menyulitkan dalam pengaksesannya. Misalnya, orang harus merasakan berat jika ingin dibawa buku tersebut, dan jika membutuhkan informasi perlu membeli atau mencari buku ke toko buku atau perpustakaan.

Dengan permasalahan tersebut, diperlukan media yang digunakan untuk memudahkan dalam proses pengeksesan. Dengan berkembangnya teknologi web sekarang ini, dapat memberikan kemudahan dalam pengaksesan dan pencarian konten yang berkaitan tentang isi dari *ensiklopedia* tersebut.

Dengan adanya jumlah hadist yang banyak pada *ensiklopedia* tersebut menyulitkan pengguna dalam menemukan konten tertentu. Dalam hal ini, diperlukan suatu metode untuk mendukung pencarian sesuai yang dikehendaki. Pengolahan kata atau pengolahan teks dapat disebut dengan *Text Mining*, pada *Text Mining* ini, kalimat atau teks akan diproses menjadi kata perkata dan diseleksi termasuk kata penting atau tidak. Kata – kata tersebut akan dioleh untuk mendapatkan kata dasar (*stem*) atau disebut dengan *stemming*. Kemudian, hasil *stem* tersebut akan dianalisis dengan mengitung bobot kata dan keseusian dengan *keyword*.

Dalam penelitian ini menggunakan metode *TF-IDF* untuk melakukan perhitungan bobot kata yang sering muncul dalam dokumen. Metode *Term Frequency-invers Document Frequency (TF-IDF)* merupakan suatu metode yang digunakan untuk menghitung bobot hubungan kata (*term*) terhadap dokumet.

Hasil perhitungan bobot dengan *TF-IDF* akan dilakukan perhitungan bobot kemiripan document dengan kata kunci. Dalam mengukur tingkat kemiripan (*similarity*) antara *keyword* terhadap documen, dalam penelitian ini menggunakan algoritma *Vector Space Model (VSM)* dari

informasi *retrieval system* untuk mengukur relevansi dan menemukan kembali informasi dari dokumen yang telah diproses sebelumnya.

Berdasarkan uraian di atas, maka untuk mendapatkan suatu informasi dan menyadari pentingnya penggunaan pencarian *ensiklopedia* hadits sebagai program aplikasi yang ditujukan untuk mempermudah, pencarian data dan tidak kesulitan lagi dalam mencari data dari ensiklopedi hadits tersebut, maka penulis mencoba meneliti dan menuangkan dalam bentuk Tugas Akhir dengan judul” **Rancang Bangun Aplikasi *Search Engine* Ensiklopedia Hadits Menggunakan Teknik *Text Mining* dengan Algoritma VSM (*Vector Space Model*)**”

1.2.Rumusan Masalah

Berdasarkan latar belakang di atas dapat dirumuskan permasalahan yang akan diselesaikan dalam penelitian ini yaitu: bagaimana mengimplementasikan *Text Mining* menggunakan algoritma *vector space model* untuk mendukung aplikasi *search engine* ensiklopedia hadits Shahih Bukhari?

1.3.Tujuan Penelitian

Adapun tujuan peneliti ini adalah mengimplementasikan *Text Mining* menggunakan algoritma *vector space model* untuk mendukung pencarian pada ensiklopedia hadits

1.4.Batasan Masalah

Batasan masalah ini bertujuan untuk memudahkan pekerjaan dan menghindari adanya kegiatan di luar sasaran, sehingga dalam pembuatan laporan ini perlu ditentukan suatu batasan masalah. Batasan masalah tersebut adalah sebagai berikut.

1. Pencarian dengan konten teks bahasa Indonesia dalam bentuk kata- kata

2. Data- data yang berhubungan dengan ensiklopedia hadits Shahih Bukhari.
3. Data yang dimasukan berjumlah 7008 hadits Shahih Bukhari.
4. Implementasi yang akan dilaksanakan untuk membangun *website* dinamis dengan menggunakan *PHP (Personal Home Page)* serta untuk media penyimpanannya menggunakan *Mysql*, dan untuk tampilan antar muka menggunakan *HTML, CSS* dan *JQuery* yang berguna untuk mendapatkan tampilan yang dinamis.
5. Menggunakan metode *waterfall* sebagai tahapan pengembangan perangkat lunak.

1.5. Manfaat Penelitian

Manfaat dari penelitian ini adalah terbentuknya suatu aplikasi pencarian *ensiklopedia* hadits dengan mengimplementasikan *Text Mining* untuk memudahkan dalam proses pencarian dan mendapatkan informasi hadits sesuai yang dikehendaki berdasarkan tingkat kemiripan dokumen terhadap kata kunci sehingga dapat dijadikan referensi dan solusi sesuai sunnah Rasulullah SAW.

1.6. Metodologi Penulisan

Pembuatan skripsi ini terbagi menjadi beberapa tahap pengerjaan yang tertera sebagai berikut :

1. Pengumpulan data-data yang diperlukan

Beberapa metode yang akan dipakai dalam pengumpulan data :

- a. Studi Literatur

Pada metode ini akan melakukan pencarian dari berbagai macam literature dan dokumen yang menunjang pengerjaan tugas akhir ini khususnya yang berkaitan dengan aplikasi *search engine* ensiklopedi hadits menggunakan teknik *Text Mining*

Text Mining merupakan proses otomatis atau sebagian proses otomatis untuk teks. Ini melibatkan pembentukan text yang lebih terstruktur dan penggalian informasi yang relevan dari teks. (Miller, 2005;104).

Butuh mekanisme untuk menambang teks – teks yang ada dalam koleksi dokumen sehingga didapatkan informasi –informasi yang lebih

bernilai dan terstruktur. Mekanisme tersebut dibagi dalam beberapa tahapan (*fase preprocessing*) yang dilakukan secara umum dalam *Text Mining*, yaitu: *Tokenizing*, *Filtering*, *Stemming*, *Tagging*, dan *Analyzing*. dengan algoritma *vsm* (*vector space model*)



Gambar 1. Tahapan Text Mining

b. Browsing

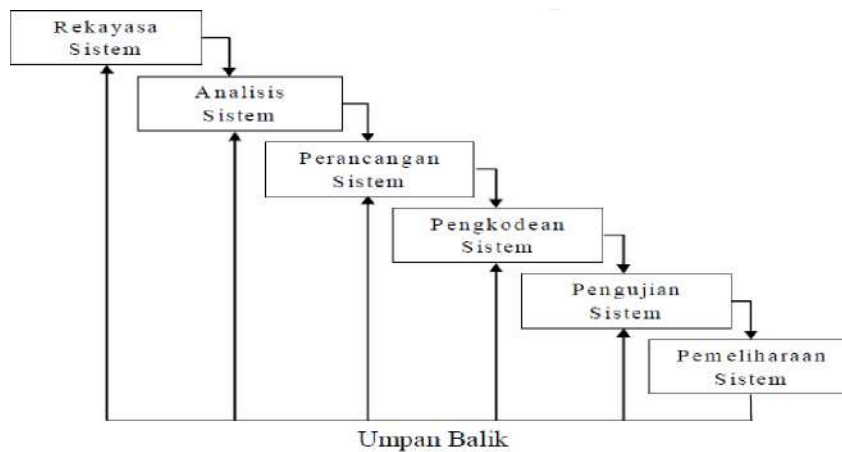
Melakukan pengamatan ke berbagai macam website di *internet* yang menyediakan informasi yang relevan dengan permasalahan dalam pembuatan sistem ini.

2. Perancangan dan Desain Sistem

Memahami rancangan kerja *search engine* sesuai data yang ada dan mengimplementasikan model yang diinginkan oleh pengguna, serta perancangan database dengan di dukung pembuatan *Use Case*, *Activity Diagram*, *Sequence Diagram*, dan *Class Diagram* guna mempermudah dalam proses selanjutnya.

3. Tahap pengembangan perangkat lunak.

Metode pembangunan perangkat lunak yang digunakan yaitu model *Waterfall*, yang meliputi beberapa proses seperti yang ada pada gambar 1.2



Gambar 1.2. Pemodelan Waterfall

Model ini telah lama digunakan untuk pengembangan perangkat lunak yang disebut sebagai model atau paradigma siklus hidup klasik. Model ini sangat terstruktur dan bersifat linier. Model ini memerlukan pendekatan yang sistematis dan sekuensial di dalam

pengembangan sistem perangkat lunaknya. Setiap tahap harus terjadi interaksi dan kerjasama yang harmonis antara pengembang perangkat lunak dengan pemesannya. Proyek akhir yang diterima oleh pengguna merupakan hasil satu siklus pengembangan (mulai dari tahap analisis dan perancangan kebutuhan sistem hingga integrasi dan pengujiannya) yang terdiri dari satu versi perangkat lunak. Adapun uraian lebih jelasnya mengenai pengembangan perangkat lunak diatas adalah sebagai berikut.

1. Rekayasa Sistem

Tahap ini merupakan kegiatan pengumpulan data sebagai pendukung pembangunan sistem serta menentukan ke arah mana aplikasi ini akan dibangun.

2. Analisis Sistem

Mengumpulkan kebutuhan secara lengkap kemudian dianalisis dan didefinisikan kebutuhan yang harus dipenuhi oleh aplikasi yang akan dibangun. Tahap ini harus dikerjakan secara lengkap untuk bisa menghasilkan desain yang lengkap.

3. Perancangan Sistem

Perancangan antarmuka dari hasil analisis kebutuhan yang telah selesai dikumpulkan secara lengkap.

4. Pengkodean Sistem

Hasil perancangan sistem diterjemahkan ke dalam kode-kode dengan menggunakan bahasa pemrograman yang sudah ditentukan. Aplikasi yang dibangun langsung diuji baik secara unit.

5. Pengujian Sistem

Penyatuan unit-unit program kemudian diuji secara keseluruhan.

6. Pemeliharaan Sistem

Mengoperasikan aplikasi dilingkungannya dan melakukan pemeliharaan, seperti penyesuaian atau perubahan karena adaptasi dengan situasi yang sebenarnya.

7. Umpan Balik

Merupakan respon dari pengguna sistem yang bisa digunakan untuk mengetahui sejauh mana aplikasi yang dibangun diterima oleh penggunanya.

4. Pembuatan Aplikasi

Tahap ini merupakan tahap pembuatan dan pengembangan aplikasi sesuai dengan desain sistem yang ditetapkan pada tahap sebelumnya. aplikasi *search engine* ensiklopedi hadits menggunakan teknik *Text Mining* dengan *algoritma vsm (vector space model)* ini dibangun menggunakan *PHP* dan *MySQL*.

5. Ujicoba Sistem

Menguji seluruh spesifikasi terstruktur dan sistem secara keseluruhan. Proses uji coba ini diperlukan untuk memastikan bahwa sistem yang telah dibuat sudah benar, sesuai dengan karakteristik yang ditetapkan dan tidak ada kesalahan-kesalahan yang terkandung di dalamnya.

Uji coba sistem dan program `dalam pembuatan aplikasi *search engine* ensiklopedi hadits menggunakan teknik *Text Mining* dengan algoritma *VSM (vector space model)* ini akan dilakukan dengan pengetesan fungsi dan logika yang telah di berikan pada masing- masing program atau modul, apakah masih ada kesalahan atau tidak. Dengan menjalankan program dapat dideteksi kesalahan yang ada dan fungsi-fungsi yang tidak sesuai.

6. Penyusunan Buku Skripsi

Tahap terakhir ini merupakan dokumentasi pelaksanaan skripsi. Diharapkan, buku tugas akhir ini bermanfaat bagi pembaca yang ingin mengembangkan sistem ini lebih lanjut maupun pada lain kasus.

1.7.State of The Art

Penelitian pertama yang dilakukan oleh **Sumartini (2011)**. Pada penelitian ini dibahas klasifikasi opini sebagai opini positif dan opini negatif pada data berbahasa Inggris dan data berbahasa Indonesia menggunakan metode *Naïve Bayes Classifier (NBC)* dan *Support Vector Machine (SVM)*. Baik metode *NBC* maupun metode *SVM* memberikan unjuk kerja yang baik dalam *sentiment analysis* pengklasifikasian opini berbahasa Inggris dan berbahasa Indonesia pada penelitian ini. Hasil percobaan menunjukkan bahwa metode *SVM* memberikan unjuk kerja yang lebih baik daripada metode *NBC* untuk mengklasifikasikan opini berbahasa Inggris dan opini positif berbahasa Indonesia. Sedangkan *NBC* memberikan unjuk kerja yang lebih baik dalam mengklasifikasikan data uji opini negatif berbahasa Indonesia.

Penelitian kedua yang dilakukan oleh **Firnas (2006)**. Pada penelitian ini yang berupa pengembangan sistem temu-kembali informasi yang dibuat dapat mencari informasi dari isi file

dokumen yang disimpan di dalam sistem. Proses pengindeksan dokumen didalam sistem temu-kembali informasi yang dikembangkan melalui beberapa tahapan pemrosesan teks, yaitu parsing, penghilangan *stopwords* dan penghitungan nilai bobot setiap kata yang akan dijadikan indeks. Sedangkan untuk proses pencariannya juga melalui beberapa tahapan proses yang yang hampir sama dengan proses pengindeksan, yaitu parsing, penghilangan *stopwords*, *cek frasa* dan yang terakhir adalah penghitungan fungsi kesamaan untuk mendapatkan nilai bobot setiap dokumen yang akan dicari.

Penelitian ketiga yang dilakukan oleh **Abdul dan Rendy Hardi (2010)** tentang. Pada aplikasi ini terdapat fitur pencarian kata dimana penulis menggunakan *vector space model* yang melakukan perhitungan kemiripan kata berdasarkan kata yang diinput dan kata yang ada di database sebagai metode untuk melakukan pencarian kata pada terjemahan surat-surat yang terdapat pada Juz Amma. Aplikasi Juz Amma ini dikembangkan dengan metodologi *RUP*, metodologi ini terdapat empat tahapan, yaitu: *inception*, *elaboration*, *construction* dan *transition*. Sedangkan bahasa pemrograman yang digunakan adalah bahasa *Java ME*. Dalam menganalisis kebutuhan digunakan UML diantaranya adalah diagram *use case*, *activity diagram*, *sequence diagram*, dan *class diagram*. Hasil yang dicapai adalah suatu aplikasi bergerak (*mobile*) yang memudahkan pengguna dalam mempelajari *Juz Amma* dengan hasil pencarian kata yang tepat dan cepat sesuai kata yang diinput.

1.8.Sistematika Penulisan

Sistematika penulisan laporan ini disusun dalam beberapa bab yang masing-masing bab menguraikan beberapa pokok pembahasan. Adapun sistematika penulisan laporan ini adalah sebagai berikut :

BAB I PENDAHULUAN

Bab ini berisikan tentang latar belakang permasalahan, perumusan masalah yang dihadapi, batasan masalah, tujuan, metodologi, serta bagaimana penulisan sistematika penulisan

BAB II LANDASAN TEORI

Bab ini menjelaskan tentang teori-teori apa saja yang berkaitan dengan topik yang telah dibuat berdasarkan hasil penelitian dan hal-hal yang berguna dalam proses penyusunan tugas akhir ini.

BAB III ANALISIS KEBUTUHAN

Memuat gambaran analisis yang dibutuhkan oleh sistem, diantaranya proses bisnis sistem, kebutuhan perangkat lunak dan skenario untuk pembuatan proses pembuatan aplikasi.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Menerangkan pengimplementasian dari sistem yang telah dibangun baik itu *software* yang diperlukan, *hardware* yang mendukung, implementasi *user interface* termasuk pengujian sistem yang telah dibangun.

BAB V PENUTUP

Bab ini berisikan tentang kesimpulan dan saran yang diperoleh dari hasil penulisan laporan tugas akhir.



uin

UNIVERSITAS ISLAM NEGERI
SUNAN GUNUNG DJATI
BANDUNG