

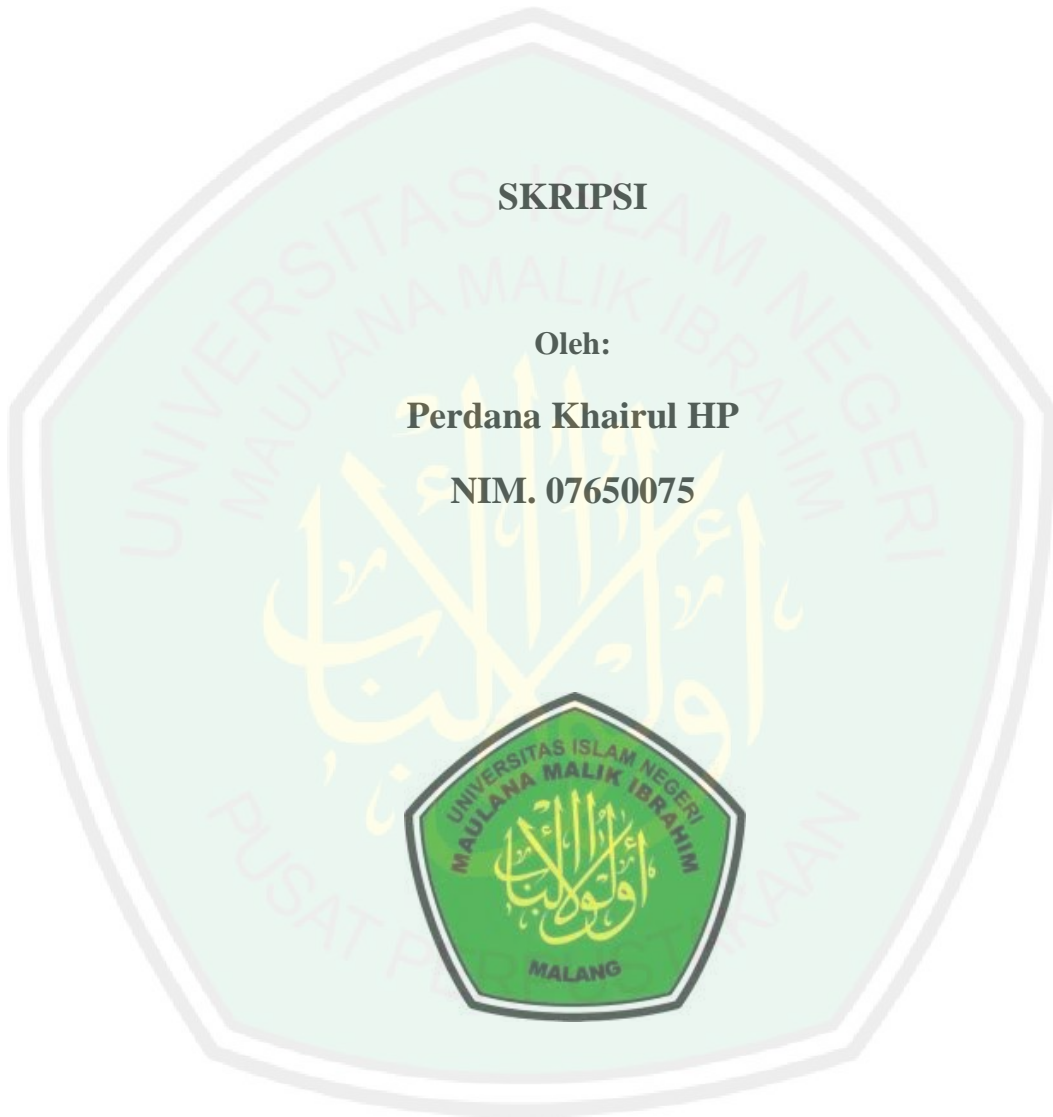
**PENCARIAN DAN PERANKINGAN OBAT TRADISIONAL
BERDASARKAN GEJALA PENYAKIT MENGGUNAKAN
METODE COSINE SIMILARITY**

SKRIPSI

Oleh:

Perdana Khairul HP

NIM. 07650075



**JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI (UIN)
MAULANA MALIK IBRAHIM MALANG
2014**

**PENCARIAN DAN PERANKINGAN OBAT TRADISIONAL
BERDASARKAN GEJALA PENYAKIT MENGGUNAKAN
METODE COSINE SIMILARITY**

SKRIPSI

**Diajukan Kepada :
Dekan Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim (UIN Maliki) Malang
untuk Memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)**

Oleh:

Perdana Khairul HP

NIM. 07650075

**JURUSAN TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI (UIN)
MAULANA MALIK IBRAHIM MALANG
2014**

**PENCARIAN DAN PERANKINGAN OBAT TRADISIONAL
BERDASARKAN GEJALA PENYAKIT MENGGUNAKAN
METODE COSINE SIMILARITY**

SKRIPSI

Oleh:
Perdana Khairul HP
NIM. 07650075

Telah Diperiksa dan Disetujui untuk Diuji
Tanggal : 8 Juli 2014

Pembimbing I

Pembimbing II

Zainal Abidin, M.Kom
NIP. 19760613 200501 1 004

A'la Syauqi, M.Kom
NIP. 19771201 200801 1 007

Mengetahui,
Ketua Jurusan Teknik Informatika

Dr. Cahyo Crysdiان
NIP. 19740424 200901 1 008

**PENCARIAN DAN PERANKINGAN OBAT TRADISIONAL
BERDASARKAN GEJALA PENYAKIT MENGGUNAKAN
METODE COSINE SIMILARITY**

SKRIPSI

Oleh:

**Perdana Khairul HP
NIM. 07650075**

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal, 10 Juli 2014

Susunan Dewan Penguji :

Tanda Tangan

- | | | |
|-----------------------|---|-----|
| 1. Penguji Utama | : Irwan Budi Santoso, M.Kom
NIP. 19770103 201101 1 004 | () |
| 2. Ketua Penguji | : Fatchurrochman, M.Kom
NIP. 19700731 200501 1 002 | () |
| 3. Sekretaris Penguji | : Zainal Abidin, M.Kom
NIP. 19760613 200501 1 004 | () |
| 4. Anggota Penguji | : A'la Syauqi, M.Kom
NIP. 19771201 200801 1 007 | () |

Mengesahkan,
Ketua Jurusan Teknik Informatika

Dr. Cahyo Crysdian
NIP. 19740424 200901 1 008

LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini :

Nama : Perdana Khairul Hima Pandalu

NIM : 07650075

Fakultas/Jurusan : Sains dan Teknologi/Teknik Informatika

Judul Penelitian : Pencarian dan Perankingan Obat Tradisional Berdasarkan Gejala Penyakit Menggunakan Metode Cosine Similarity

Menyatakan dengan sebenarnya bahwa hasil penelitian saya ini tidak terdapat unsur-unsur penjiplakan karya penelitian atau karya ilmiah yang pernah dilakukan atau dibuat orang lain, kecuali yang secara tertulis dikutip dalam naskah ini atau disebutkan dalam sumber kutipan dan daftar pustaka. Apabila ternyata hasil penelitian ini terbukti terdapat unsur-unsur jiplakan maka saya bersedia untuk mempertanggung jawabkan, serta diproses sesuai peraturan yang berlaku.

Malang, 8 Juli 2014

Yang membuat pernyataan,

Perdana Khairul HP

NIM 07650075

MOTTO

فَإِنَّ مَعَ الْعُسْرِ يُسْرًا ﴿٥﴾ إِنَّ مَعَ الْعُسْرِ يُسْرًا ﴿٦﴾

**“karena Sesungguhnya sesudah kesulitan itu ada kemudahan.
Sesungguhnya sesudah kesulitan itu ada kemudahan.”**

(QS ALAM NASYRAH 5-6)



Lembar Persembahan

Alhamdulillah, Segala puji bagi-Mu ya Allah

Telah memberikan kesempatan hamba berada dalam keluarga yang penuh kasih sayang. Terima kasih atas doa dukungan nasehat Bunda Tercinta, Yeti Martiningsih. Terima kasih untuk Ayahku, Hadi Sucipto pemimpin terbaik dalam keluargaku yang tidak pernah bosan memberikan arahan untuk kehidupan yang lebih baik. Terima kasih adik-adikku, Galang dan Agil kalian merupakan penyemangatku. Terima kasih keluargaku

Terima kasih juga untuk,

Teman-teman TI kelas B '07, IOC, IMM REVIVALIS dukungan kalian adalah dorongan semangat yang luar biasa

KATA PENGANTAR

Bismillahirrahmanirrahim.....

Assalamualaikum Wr. Wb.

Alhamdulillahillobbilalamin, Segala puji bagi Tuhan Semesta Alam Allah Azza Wajalla karena atas rahmat, taufik dan Hidayah-Nya, sehingga penulis dapat menyelesaikan penulisan tugas akhir ini dengan judul “Pencarian dan Perankingan Obat Tradisional Berdasarkan Gejala Penyakit Menggunakan Metode Cosine Similarity”. Shalawat serta salam semoga tetap tercurahkan kepada junjungan kita Nabi Muhammad SAW beserta keluarga dan sahabatnya sampai hari akhir nanti.

Penulis menyadari bahwa banyak pihak yang telah membantu dalam menyelesaikan penulisan tugas akhir ini. Untuk itu, iringan doa’ dan ucapan terima kasih penulis sampaikan kepada:

1. Prof. Dr. Mudjia Raharjo, M,Si selaku Rektor Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
2. Dr. drh. Bayyinatul Muchtaromah, M.Si. selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
3. Dr. Cahyo Crysdiyan, selaku Ketua Jurusan Teknik Informatika Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.
4. Zainal Abidin, M.Kom selaku dosen pembimbing utama, karena atas bimbingan, pengarahan dan kesabaran beliau, penulisan tugas akhir dapat terselesaikan.
5. A’La Syauqi, M.Kom selaku dosen pembimbing agama, karena atas bimbingan, pengarahan dan kesabaran beliau penulisan tugas akhir dapat terselesaikan.
6. Segenap Dosen Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri (UIN) Maulana Malik Ibrahim Malang.

7. Teman-teman TI '07, Laboran dan Staff administrasi Jurusan Teknik Informatika yang banyak membantu selama penelitian.

Semoga Allah memberikan balasan atas jerih payah keringat yang telah diberikan kepada penulis. Akhir kata, penulis berharap skripsi ini bermanfaat dan dapat menjadi inspirasi bagi peneliti lain serta menambah khasanah ilmu pengetahuan. Dengan kerendahan hati penulis mohon maaf apabila ada kesalahan dalam penulisan skripsi ini.

Malang, 17 Juli 2014

Penulis



DAFTAR ISI

HALAMAN JUDUL	
HALAMAN PENGAJUAN	i
HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN	iv
MOTTO	v
LEMBAR PERSEMBAHAN	vi
KATA PENGANTAR	vii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
ABSTRAK	xiv
ABSTRACT	xv
BAB I PENDAHULUAN	
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian	5
1.4 Batasan Masalah	5
1.5 Manfaat Penelitian	6
1.6 Metode Penelitian	6
1.7 Sistematika Penulisan	7
BAB II KAJIAN PUSTAKA	
2.1 Obat	9
2.1.1 Golongan Obat Fisiologis dan Biokimia	11
2.1.2 Golongan Obat Berdasarkan Ketersediaan	12
2.1.3 Golongan Obat Berdasarkan Sumber	12
2.1.4 Golongan Obat Berdasarkan Undang-undang	12
2.1.5 Golongan Obat Berdasarkan Cara Kerja	14

2.1.6	Golongan Obat Menurut Penggunaan	14
2.1.7	Golongan Obat Menurut Kegunaan	14
2.2	Obat Tradisional	15
2.2.1	Ketepatan Tanaman Obat yang Digunakan	16
2.2.2	Ketepatan Waktu Penggunaan	16
2.2.3	Ketepatan Penggunaan Tanaman Obat	16
2.3	Text Mining	17
2.3.1	Case Folding	17
2.3.2	Tokenizing	18
2.3.3	Filtering	18
2.3.4	Stemming	18
2.3.5	Pembobotan tf/idf	19
2.3.6	Vector Space Model	19
2.4	Pemrograman PHP	20
2.4.1	Kelebihan PHP	21
2.4.2	Kekurangan PHP	22
2.5	MySQL	22
2.6	Webserver	23
2.7	Apache	25
BAB III DESAIN DAN PERANCANGAN SISTEM		
3.1	Perancangan Sistem	26
3.1.1	Data	26
3.1.2	Spesifikasi Sistem	26
3.1.3	Spesifikasi Pengguna	28
3.1.4	Lingkungan Operasi	29
3.1.5	Deskripsi Sistem	29
3.2	Desain Sistem	30
3.2.1	Ekstraksi Dokumen	30
3.2.2	Analisa	33
3.3	Contoh Perhitungan Manual	35
3.4	Perancangan Aplikasi	32

BAB IV HASIL DAN PEMBAHASAN

4.1 Implementasi	46
4.2 Pembahasan	50
4.3 Uji Coba	58
4.4 Hubungan Penelitian dengan Kaidah Islam	62

BAB V KESIMPULAN

5.1 Kesimpulan	65
5.2 Saran	65

DAFTAR PUSTAKA	66
-----------------------------	-----------



DAFTAR TABEL

Tabel 3.1 Hasil Tokenizing	29
Tabel 3.2 Hasil Filtering	29
Tabel 3.3 Hasil Perhitungan tf	33
Tabel 3.4 Hasil Perhitungan idf	34
Tabel 3.5 Rancangan Tabel Stoplist	32
Tabel 3.6 Rancangan Tabel Data Obat	33
Tabel 3.7 Rancangan Tabel Data User	33

DAFTAR GAMBAR

Gambar 3.1 Case Folding	21
Gambar 3.2 Tokenizing	22
Gambar 3.3 Filtering	22
Gambar 3.4 Cosine Similarity	24
Gambar 3.5 Sitemap	34
Gambar 3.7 Perancangan Login	34
Gambar 3.8 Perancangan Halaman	35
Gambar 3.9 Perancangan Halaman Tambah Obat	35
Gambar 4.1 Halaman Login	37
Gambar 4.2 Halaman Admin	38
Gambar 4.3 Halaman Tambah Obat 1	38
Gambar 4.4 Halaman Tambah Obat 2	39
Gambar 4.5 Halaman Lihat Obat	39

Abstrak

Khairul, Perdana HP. 2014. **Pencarian dan Perankingan Obat Tradisional Berdasarkan Gejala Penyakit Menggunakan Metode Cosine Similarity**. Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (1) Zainal Abidin, M.Kom. (2) A'la Syauqi, M.Kom.

Kata kunci: Obat Tradisional, Pencarian, *Text Mining*, *Vector Space Model*

Obat-obatan yang beredar sekarang banyak yang mengandung bahan kimia, sehingga bersifat anorganik sedangkan tubuh kita bersifat organik. Hal ini sangat berbahaya jika jumlah zat anorganik berlebihan di dalam tubuh. Sejumlah pasien memiliki beberapa penyakit maupun keterbatasan fisik yang kompleks. Beberapa orang bahkan memiliki lebih dari satu penyakit dan penggunaan varian obat kimia tidak terelakan. Padahal jika ada beberapa reaksi kimia, memungkinkan penurunan manfaat dalam obat kimia tersebut dan tentunya menambah resiko penggunaan obat kimia. Salah satu alternatifnya adalah menggunakan tanaman-tanaman obat tradisional yang terbukti berkhasiat untuk pengobatan dari pusing hingga penyakit kanker.

Penelitian ini menggunakan metode text mining yang mengimplementasikan algoritma cosine similarity untuk peringkatan jenis obat yang paling mirip. Hal ini diperlukan karena terdapat berbagai macam jenis obat untuk jenis penyakit yang sama. Pengguna memasukkan gejala penyakit yang diderita, kemudian system akan mengukur kemiripan gejala yang diberikan dengan manfaat obat tradisional. Dalam text mining ada beberapa proses yang penting, yaitu case folding, tokenizing, filtering dan analisa. Analisa dalam pembobotan dokumen digunakan algoritma TF/IDF dan cosine similarity (Vector Space Model). Data uji coba diperoleh dari buku-buku yang berhubungan dengan obat tradisional.

Abstract

Khairul, Perdana HP. 2014. **Searching and Ranking Traditional Medicine base Disease Symptom Use Cosine Similarity Method**. Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (1) Zainal Abidin, M.Kom.(2) A'la Syauqi, M.Kom.

Drugs are now circulating that contain a lot of chemicals, so it is inorganic organik. So our bodies are very dangerous if too much of inorganic substances in the body. Some patients have some illness or physical limitations that complex. Some people even have more than one variant of the disease and the use of chemical drugs are not inevitable. And if there are some chemical reactions, allowing a decrease in the benefits of the medicinal chemistry and certainly adds to the risk of the use of chemical drugs. One alternative is the use of traditional medicinal plants that are proven efficacious for the treatment of dizziness to cancer.

This study uses text mining methods that implement the cosine similarity algorithm to ranking the most similar type of drug. This is necessary because there are many kinds of drugs for the same type of disease. Users enter the symptoms of the illness, then the system will measure the similarity of symptoms given with the benefits of traditional medicine. In text mining there are several important processes, namely the folding case, tokenizing, filtering and analysis. Analysis of documents used in the weighting algorithm TF / IDF and cosine similarity (Vector Space Model). Test data obtained from books related to traditional medicine.

Keyword : Traditional Medicine, Searching, *Text Mining*, *Vector Space Model*



BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Kesehatan merupakan kebutuhan yang sangat penting. Menurut MUI dalam Musyawarah Nasional Ulama tahun 1983 merumuskan kesehatan sebagai “Ketahanan jasmaniah, rohaniah, dan sosial yang dimiliki manusia, sebagai karunia Allah yang wajib disyukuri dengan mengamalkan (tuntunanNya) dan memelihara serta mengembangkannya”. Dari pengertian tersebut bisa disimpulkan bahwa penyakit dapat dibagi menjadi dua yaitu penyakit hati yang akan berpengaruh terhadap kehidupan rohani dan sosial seseorang dan penyakit jasmani yang berwujud fisik.

Dalam kaitannya terhadap penyakit hati, Allah berfirman dalam surat Al Isra ayat 82 :

خَسَارًا إِلَّا الظَّالِمِينَ يَزِيدُ وَلَا لِلْمُؤْمِنِينَ وَرَحْمَةً شِفَاءً هُوَ مَا الْقُرْءَانِ مِنْ وَنُنَزَّلُ



Artinya : “ dan Kami turunkan dari Al Quran suatu yang menjadi penawar dan rahmat bagi orang-orang yang beriman dan Al Quran itu tidaklah menambah kepada orang-orang yang zalim selain kerugian.”

Dari ayat diatas dapat disimpulkan penyembuh dari penyakit hati adalah Al Quran.Sedangkan dalam kaitannya dengan penyakit fisik, rosululloh menjelaskan melalui beberapa hadist yang diriwayatkan ke beberapa sahabatnya.

Dari riwayat Imam Muslim dari Jabir bin Abdillah dia berkata bahwa Nabi bersabda,

Artinya : “ *Setiap penyakit pasti memiliki obat. Bila sebuah obat sesuai dengan penyakitnya maka dia akan sembuh dengan seizin Allah Subhanahu wa Ta’ala.*” (HR. Muslim)

Diriwayatkan pula dari musnad Imam Ahmad dari shahabat Usamah bin Suraik , bahwasanya Nabi bersabda,

كُنْتُ عِنْدَ النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ، وَجَاءَتِ الْأَعْرَابُ، فَقَا : يَا رَسُولَ اللَّهِ، أَنْتَدَاوَى؟ فَقَالَ: نَعَمْ يَا عِبَادَ اللَّهِ، تَدَاوَوْا، فَإِنَّ اللَّهَ عَزَّ وَجَلَّ لَمْ يَضَعْ دَاءً إِلَّا وَضَعَ لَهُ شِفَاءً غَيْرَ دَاءٍ وَاجِدِ. : مَا هُوَ؟ قَالَ: الْهَرَمُ

Artinya : “ *Aku pernah berada di samping Rasulullah b. Lalu datanglah serombongan Arab dusun. Mereka bertanya, “Wahai Rasulullah, bolehkah kami berobat?” Beliau menjawab: “Iya, wahai para hamba Allah, berobatlah. Sebab Allah I tidaklah meletakkan sebuah penyakit melainkan meletakkan pula obatnya, kecuali satu penyakit.” Mereka bertanya: “Penyakit apa itu?” Beliau menjawab: “Penyakit tua.”* (HR. Ahmad, Al-Bukhari dalam Al-Adabul Mufrad, Abu Dawud, Ibnu Majah, dan At-Tirmidzi, beliau berkata bahwa hadits ini hasan shahih. Syaikhuna Muqbil bin Hadi Al-Wadi’i menshahihkan hadits ini dalam kitabnya Al-Jami’ Ash-Shahih mimma Laisa fish Shahihain, 4/486)

Hadits-hadits di atas memberikan pengertian kepada kita bahwa semua penyakit yang menimpa manusia maka Allah turunkan obatnya.

Obat-obatan yang beredar sekarang banyak yang mengandung bahan kimia, sehingga bersifat anorganik sedangkan tubuh kita bersifat organik. Hal ini sangat berbahaya jika jumlah zat anorganik berlebihan di dalam tubuh. Terlebih sejumlah pasien memiliki beberapa penyakit maupun keterbatasan fisik yang kompleks. Beberapa orang bahkan memiliki lebih dari satu penyakit dan penggunaan varian obat kimia tidak terelakan. Padahal jika ada beberapa reaksi kimia, memungkinkan penurunan manfaat dalam obat kimia tersebut dan tentunya menambah resiko penggunaan obat kimia.

Salah satu alternatifnya adalah menggunakan tanaman-tanaman obat tradisional yang terbukti berkhasiat untuk pengobatan dari pusing hingga penyakit kanker. Tanaman obat tradisional memiliki beberapa kelebihan, antara lain : efek sampingnya relatif rendah, dalam suatu ramuan dengan komponen berbeda memiliki efek saling mendukung, pada satu tanaman memiliki lebih dari satu efek obat serta lebih sesuai untuk penyakit-penyakit metabolic (penyakit yang berkaitan dengan produksi energi sel manusia) dan degeneratif (penyakit yang diakibatkan faktor usia) (katno,S Pramono.Tingkat Keamanan dan Manfaat tanaman obat tradisional).

Berkaitan dengan bahaya yang ditimbulkan dari konsumsi obat kimia perlu kita berpikir ulang sebelum kita memutuskan untuk menggunakan obat-obat kimia. Namun, ketika kita mulai berpikir untuk menggunakan obat-obat tradisional muncul sebuah masalah yaitu jumlah tanaman obat di Indonesia yang

mencapai 30.000 mengingat kita merupakan Negara Tropis. Dengan jumlah tanaman obat yang banyak itu diperlukan metode untuk mempermudah dalam pencarian data obat tradisional sehingga dapat menekan efisiensi waktu pencarian.

Metode Cosine Similarity merupakan metode pencarian data berdasarkan kata kunci tertentu dimana metode ini menghitung bobot setiap dokumen. Dokumen yang memiliki bobot tertinggi merupakan dokumen yang termirip terhadap kata kunci. Dengan adanya bobot ini akan dapat dilihat sebagai rank data-data yang memiliki hubungan dengan kata kunci. Dengan menggunakan metode Cosine Similarity diharapkan dapat mempermudah pencarian data-data obat-obat tradisional berdasarkan kemiripan gejala-gejala penyakit yang dirasakan.

Mengingat harga penggunaan internet yang semakin terjangkau, penggunaan internet sebagai media informasi menjadi solusi yang efektif untuk penyebaran informasi keseluruhan lapisan masyarakat. Oleh karena itu penggunaan internet sebagai media untuk berjalannya sistem ini akan membuat sistem ini lebih bermanfaat bagi masyarakat.

1.2 Rumusan Masalah

Beberapa lingkup permasalahan yang akan dibahas dalam penelitian proyek akhir ini diantaranya :

1. Tidak ada aplikasi pencarian dan perankingan obat tradisional Indonesia secara khusus.
2. Bagaimana menampilkan data-data obat tradisional indonesia berdasarkan gejala-gejala penyakit menggunakan metode Cosine Similarity.

1.3 Tujuan Penelitian

Tujuan dari pelaksanaan tugas akhir ini adalah membuat aplikasi yang dapat mencari dan meranking jenis obat tradisional berdasarkan gejala penyakit menggunakan metode Cosine Similarity

1.4 Batasan Masalah

Pada aplikasi sistem berbasis pengetahuan untuk pemilihan obat tradisional indonesia ini diberikan pembatasan masalah agar tidak keluar dari pokok permasalahan yang dirumuskan, maka ruang lingkup pembahasan dibatasi pada :

- a. Gejala-gejala penyakit yang disediakan berjumlah 36 gejala penyakit.
- b. Masukkan berupa gejala penyakit yang diberikan pengguna saat melakukan pencarian obat.
- c. Proses pencarian jenis tanaman obat beserta ramuannya yang relevan dengan input gejala penyakit menggunakan metode cosine similarity untuk mencari bobot hubungan antara query dan term.

- d. Penyakit-penyakit yang dapat terdeteksi berjumlah 10 jenis penyakit.
- e. Ramuan-ramuan tanaman obat yang disediakan untuk mengobati penyakit sebanyak 150 ramuan tanaman obat.

1.5 Manfaat Penelitian

Manfaat yang dapat dihasilkan dari penyusunan tugas akhir ini adalah :

- a. Sebagai bahan acuan serta pembuka wawasan untuk masyarakat maupun akademisi mengenai penyakit-penyakit umum dengan macam tanaman obat yang berkaitan.
- b. Sebagai salah satu acuan alternatif dalam hal terkait pengobatan penyakit.
- c. Sebuah langkah awal sebelum pengguna memutuskan untuk mengonsumsi obat kimia mengingat akibat dari obat kimia yang berbahaya.

1.6 Metode Penelitian

Untuk mencapai tujuan yang telah dirumuskan sebelumnya, maka metodologi penelitian yang dilakukan dalam penulisan skripsi ini adalah :

1. Data yang berupa jenis obat tradisional dan manfaatnya dimasukkan dalam proses preprocessing awal yaitu (merubah ke huruf kecil dan hanya menerima karakter a- z).
2. Hasil case folding kemudian diproses menjadi kumpulan kata dalam proses tokenizing.
3. Hasil tokenizing kemudian dihilangkan kata-kata tidak penting dalam proses filtering
4. Hasil proses filtering kemudian dianalisa untuk menentukan bobot tiap kata menggunakan TF/IDF dan Cosine Similarity

1.7 Sistematika Penulisan

Adapun sistematika pada Tugas Akhir ini adalah sebagai berikut :

Bab I : Pendahuluan

Bab ini berisi Latar Belakang, Rumusan Masalah, Tujuan Penelitian, Batasan Masalah, Manfaat Penelitian, Metode Penelitian, dan Sistematika Penulisan

Bab II : Landasan Teori

Bab ini berisi Landasan Teori mengenai permasalahan yang berhubungan dengan penelitian. Bab ini menjelaskan teori-teori yang mendukung penelitian.

Bab III : Desain dan Perancangan Sistem

Bab ini menjelaskan tentang analisa yang dilakukan dalam merancang dan membuat system. Bab ini berisi Perencanaan desain system yang akan diimplementasikan. Dan perencanaan penggunaan algoritma TF/IDF dan penggunaan metode Cosine Similarity.

Bab IV : Hasil dan Pembahasan

Bab ini membahas tentang implementasi dari aplikasi yang dibuat secara keseluruhan. Serta melakukan pengujian terhadap aplikasi yang dibuat untuk mengetahui aplikasi tersebut telah dapat menyelesaikan permasalahan yang dihadapi sesuai dengan yang diharapkan.

Bab IV : Penutup

Bab ini berisi tentang kesimpulan dan saran yang diharapkan dapat bermanfaat untuk pengembangan pembuatan program aplikasi selanjutnya



BAB II

LANDASAN TEORI

2.1 Obat

Pengertian obat adalah semua bahan tunggal/campuran yang dipergunakan oleh semua makhluk untuk bagian dalam dan luar tubuh guna mencegah, meringankan, dan menyembuhkan penyakit. Sedangkan, menurut undang-undang, pengertian obat adalah suatu bahan atau campuran bahan untuk dipergunakan dalam menentukan diagnosis, mencegah, mengurangi, menghilangkan, menyembuhkan penyakit atau gejala penyakit, luka atau kelainan badaniah atau rohaniah pada manusia atau hewan termasuk untuk memperelok tubuh atau bagian tubuh manusia.

Selain pengertian obat secara umum di atas, ada juga pengertian obat secara khusus. Berikut ini beberapa pengertian obat secara khusus:

1. Obat baru: Obat baru adalah obat yang berisi zat (berkhasiat/tidak berkhasiat), seperti pembantu, pelarut, pengisi, lapisan atau komponen lain yang belum dikenal sehingga tidak diketahui khasiat dan kegunaannya.
2. Obat esensial: Obat esensial adalah obat yang paling banyak dibutuhkan untuk layanan kesehatan masyarakat dan tercantum dalam daftar Obat Esensial Nasional (DOEN) yang ditetapkan oleh Menteri Kesehatan RI.

3. Obat generik: Obat generik adalah obat dengan nama resmi yang ditetapkan dalam FI untuk zat berkhasiat yang dikandungnya.
4. Obat jadi: Obat jadi adalah obat dalam keadaan murni atau campuran dalam bentuk salep, cairan, supositoria, kapsul, pil, tablet, serbuk atau bentuk lainnya yang secara teknis sesuai dengan FI atau buku resmi lain yang ditetapkan pemerintah.
5. Obat paten: Obat paten adalah obat jadi dengan nama dagang yang terdaftar atas nama pembuat yang telah diberi kuasa dan obat itu dijual dalam kemasan asli dari perusahaan yang memproduksinya.
6. Obat asli: Obat asli adalah obat yang diperoleh langsung dari bahan-bahan alamiah, diolah secara sederhana berdasarkan pengalaman dan digunakan dalam pengobatan tradisional.
7. Obat tradisional: Obat tradisional adalah obat yang didapat dari bahan alam, diolah secara sederhana berdasarkan pengalaman dan digunakan dalam pengobatan tradisional.

Obat dapat digolongkan berdasarkan beberapa kriteria penggolongan.

Kriteria penggolongan obat yaitu berdasarkan proses fisiologis dan biokimia dalam tubuh, bentuk sediaan obat, sumber obat, undang-undang, cara kerja obat, cara penggunaan obat, serta kegunaan obat.

2.1.1 Golongan Obat Fisiologis dan Biokimia

Menurut proses fisiologis dan biokimia dalam tubuh, obat digolongkan menjadi:

1. Obat diagnostik: Obat diagnostik adalah obat yang membantu dalam mendiagnosis (mengenali penyakit), misalnya barium sulfat untuk membantu diagnosis pada saluran lambung-usus, serta natriummiopanoat dan asam iod organik lainnya untuk membantu diagnosis pada saluran empedu.
2. Obat kemoterapeutik: Obat kemoterapeutik adalah obat yang dapat membunuh parasit dan kuman di dalam tubuh inang. Obat ini hendaknya memiliki kegiatan farmakodinamik yang sekecil-kecilnya terhadap organisme inang dan berkhasiat untuk melawan sebanyak mungkin parasit (cacing protozoa) dan mikroorganisme (bakteri, virus). Obat-obat neoplasma (onkolitika, sitostika, atau obat kanker) juga dianggap termasuk golongan ini.
3. Obat farmakodinamik: Obat farmakodinamik adalah obat yang bekerja terhadap inang dengan jalan mempercepat atau memperlambat proses fisiologis atau fungsi biokimia dalam tubuh contohnya hormon, diuretik, hipnotik, dan obat otonom.

2.1.2 Golongan Obat Berdasarkan Ketersediaan

Penggolongan obat berdasarkan bentuk sediaan obat dikelompokkan menjadi:

- a. Bentuk gas; contohnya, inhalasi, spray aerosol.
- b. Bentuk cair atau larutan; contohnya, lotio, dauche, infus intravena, injeksi, epithema, clysm, gargarisma, obat tetes, eliksir, sirup dan potio.
- c. Bentuk setengah padat; misalnya salep mata (occulenta), gel, cerata, pasta, krim, salep (unguetum).
- d. Bentuk padat; contohnya, suppositoria, kapsul, pil, tablet, dan serbuk.

2.1.3 Golongan Obat Berdasarkan Sumber

Penggolongan obat berdasarkan sumbernya, dikelompokkan menjadi:

- a. Mikroba dan jamur/fungi; misalnya, antibiotik penisilin.
- b. Sintesis (tiruan); contohnya, vitamin C dan kamper sintesis.
- c. Mineral (pertambangan); contohnya, sulfur, vaselin, parafin, garam dapur, iodkali.
- d. Hewan (fauna); contohnya, cera, adeps lanae, dan minyak ikan.
- e. Tumbuhan (flora); contohnya, minyak jarak, kina, dan digitalis.

2.1.4 Golongan Obat Berdasarkan Undang-undang

Penggolongan obat menurut undang-undang dikelompokkan menjadi:

1. Obat bebas: Obat bebas adalah obat yang dapat dibeli secara bebas dan tidak membahayakan si pemakai dalam batas dosis yang dianjurkan; diberi tanda lingkaran bulat berwarna hijau dengan garis tepi hitam.

2. Obat bebas terbatas (daftar W = waarschuwing = peringatan): Obat bebas terbatas adalah obat keras yang dapat diserahkan tanpa resep dokter dalam bungkus aslinya dari produsen atau pabrik obat itu, kemudian diberi tanda lingkaran bulat berwarna biru dengan garis tepi hitam serta diberi tanda peringatan (P No.1 sampai P No.6).
3. Obat keras (daftar G = gevelijk = berbahaya): Obat keras adalah semua obat yang memiliki takaran dosis minimum (DM), diberi tanda khusus lingkaran bulat merah garis tepi hitam dan huruf K menyentuh garis tepinya, semua obat baru kecuali ada ketetapan pemerintah bahwa obat itu tidak membahayakan, dan semua sediaan parenteral/injeksi/infus intravena.
4. Psikotropika: Psikotropika adalah obat yang memengaruhi proses mental, merangsang atau menenangkan, mengubah pikiran/perasaan/kelakuan seseorang; contohnya golongan barbital/luminal, diazepam, dan ekstasi.

Narkotik: Narkotik adalah obat yang diperlukan dalam bidang pengobatan dan IPTEK serta dapat menimbulkan ketergantungan dan ketagihan/adiksi yang sangat merugikan individu apabila digunakan tanpa pembatasan dan pengawasan dokter; contohnya kodein, metadon, petidin, morfin, dan opium.

2.1.5 Golongan Obat Berdasarkan Cara Kerja

Penggolongan obat berdasarkan cara kerjanya dalam tubuh dikelompokkan menjadi:

1. Sistemik: obat yang didistribusikan ke seluruh tubuh; contohnya obat analgetik.
2. Lokal: obat yang bekerja pada jaringan setempat, seperti pemakaian topikal.

2.1.6 Golongan Obat Menurut Penggunaan

Penggolongan obat menurut cara penggunaannya, obat digolongkan menjadi:

1. *Medicamentum ad usum externum* (pemakaian luar) melalui implantasi, injeksi, membran mukosa, rektal, vaginal, nasal, ophthalmic, aurical, collutio/gargarisma/gargle, diberi tiket biru.
2. *Medicamentum ad usum internum* (pemakaian dalam) melalui oral, diberi tiket putih.

2.1.7 Golongan Obat Menurut Kegunaan

Penggolongan obat berdasarkan kegunaan dalam tubuh digolongkan ke dalam:

1. Untuk diagnosis (diagnostic).
2. Untuk mencegah (prophylactic).
3. Untuk menyembuhkan (terapeutic).

2.2 Obat Tradisional

Sesuai Pasal 1 Peraturan Kepala Badan POM No. HK.00.05.4.1384 Tahun 2005 tentang Kriteria dan Tata Laksana Pendaftaran Obat Tradisional, Obat Herbal Terstandar dan Fitofarmaka, ditetapkan bahwa : Obat tradisional adalah bahan atau ramuan bahan yang berupa bahan tumbuhan, bahan hewan, bahan mineral, sediaan sarian (galenik) atau campuran dari bahan tersebut, yang secara turun temurun telah digunakan untuk pengobatan berdasarkan pengalaman.

Menurut Badan Kesehatan Dunia mendefinisikan obat tradisional sebagai obat asli suatu Negara yang digunakan secara turun-temurun di Negara itu atau Negara lain. Obat tradisional harus memenuhi kriteria sudah digunakan minimal 3 generasi karena dengan penggunaan yang sudah lama maka telah terbukti aman dan bermanfaat. Sifat obat tradisional cenderung tertutup dalam hal ramuan dan kasiatnya. Sehingga mengakibatkan sulit berkembang. Hal ini juga yang membuat sulitnya mendapatkan hasil yang sama jika pengobatan diulang. Namun dengan semakin tua umur pengobatan makin banyak bukti empiris yang diungkap.

Sementara ini banyak orang beranggapan bahwa penggunaan tanaman obat atau obat tradisional relatif lebih aman dibandingkan obat sintesis. Namun demikian bukan berarti tanaman obat atau obat tradisional tidak memiliki efek samping yang berbahaya, bila penggunaannya tidak tepat. Agar penggunaannya optimal, perlu diketahui informasi yang memadai tentang kelebihan dan kelemahan serta kemungkinan penyalahgunaan obat

tradisional. Dengan informasi yang cukup diharapkan masyarakat lebih cermat untuk memilih dan menggunakan suatu produk obat tradisional atau tumbuhan obat dalam upaya kesehatan.

Efek samping obat tradisional relatif kecil jika digunakan secara tepat, yang meliputi :

2.2.1 Ketepatan Tanaman Obat yang Digunakan

Jumlah tanaman obat tradisional di Indonesia sangat banyak. Setiap tanaman obat memiliki manfaat yang berbeda-beda. Ketepatan bahan menentukan tercapai atau tidaknya manfaat yang diinginkan.

2.2.2 Ketepatan Waktu Penggunaan

Ketepatan waktu penggunaan Tanaman Obat sangat berpengaruh terhadap efek yang dihasilkan. Misalnya saja buah belimbing sangat baik untuk menjaga stamina, namun jika badan terasa gejala batuk maka buah belimbing justru akan memacu untuk timbulnya penyakit batuk. Hal ini menunjukkan bahwa ketepatan waktu penggunaan obat tradisional menentukan tercapai atau tidaknya efek yang diharapkan.

2.2.3 Ketepatan Penggunaan Tanaman Obat

Satu tanaman obat dapat memiliki beberapa zat yang berkhasiat terhadap penyakit. Setiap zat kemungkinan membutuhkan perlakuan yang berbeda dalam penggunaannya. Salah satu contohnya adalah daun kecubungakan menjadi obat asma jika dihisap seperti rokok bersifat bronkodilator. Namun jika diseduh dan diminum dapat menyebabkan keracunan.

2.3 Text Mining

Menurut buku *The Text Mining Handbook*, text mining dapat diartikan sebagai suatu proses menggali informasi dimana seorang user berinteraksi dengan sekumpulan dokumen menggunakan tools analisis yang merupakan komponen-komponen dalam data mining yang salah satunya adalah peringkatan dokumen. Penggunaan text mining bertujuan untuk mendapatkan informasi yang bermanfaat dari sekumpulan dokumen. Sumber data yang digunakan pada proses Text Mining adalah kumpulan text yang tidak memiliki arti khusus.

Dengan adanya proses Text Mining ini maka dapat diambil informasi yang bermanfaat bagi pengguna. Adapun tugas khusus dari Text mining antara lain yaitu pengkategorisasian teks dan pengelompokan teks. Sebenarnya dilihat dari fungsinya Text Mining merupakan bagian dari Data Mining. Hanya saja Text Mining mendapatkan data dari kumpulan text yang tidak terstruktur. Text Mining terdiri dari beberapa proses yaitu Case Folding dan Tokenizing, Filtering, Stemming, Pembobotan TF/IDF, Vector Space Mode.

Berikut beberapa penjelasan mengenai fungsi-fungsi dari tahap-tahap Text Mining :

2.3.1 Case Folding

Proses ini melakukan penghilangan kata. Perubahan huruf kapital menjadi huruf kecil. Dan penghilangan symbol-symbol yang dianggap tidak penting. Jika proses ini tidak dilakukan kata "batuk" dan "Batuk" itu

dianggap tidak sama. Hasilnya berupa teks yang memanjang seperti paragraph. Paragraf hasil proses ini akan di pisah-pisah menjadi kata per kata.

2.3.2 Tokenizing

Untuk mengukur kecocokan suatu kata kunci dengan sebuah teks, kita perlu memisahkan perkata dalam teks. Sehingga setelah mejadi per kata kita bisa menganalisa antar kata. Oleh karena itu kita perlu proses Tokenizing. Tokenizing merupakan proses pemotongan data input berupa string menjadi kata-kata tunggal berdasarkan spasi.

2.3.3 Filtering

Setelah proses tokenizing, data yang dihasilkan berupa kata-kata tunggal. Namun kata-kata tunggal itu masih terdiri dari kata-kata penting dan kata-kata tidak penting. Kata penting yang dimaksud adalah kata-kata yang nantinya akan dianalisa, sedangkan kata tidak penting merupakan kata-kata berupa kata depan dan kata sambung.

2.3.4 Stemming

Ini adalah proses merubah dari kata berimbuhan menjadi kata dasar. Proses diperlukan untuk mengurangi proses perhitungan. Kata makan, dimakan, memakan akan memiliki kata dasar yang sama. Sehingga pemrosesan tidak perlu sampai tiga kali namun hanya sekali.

2.3.5 Pembobotan TF/IDF

Untuk mengalisa keterhubungan antar kata-kata dari document yang ada maka digunakan algoritma TF/IDF. Algoritma ini memiliki formula yang digunakan untuk menghitung bobot (w) masing-masing dokumen terhadap kata kunci yaitu :

$$W_{d,t} = tf_{d,t} * IDF_t \dots\dots\dots (2.1)$$

$$IDF = \log (D/df) \dots\dots\dots (2.2)$$

Dimana :

d = dokumen ke $-d$

t = kata ke $-t$ dari kata kunci

W = bobot dokumen ke $-d$ terhadap kata ke $-t$

D = total dokumen

df = banyak dokumen yang mengandung kata yang dicari

tf = banyak kata yang dicari pada sebuah dokumen

Setelah bobot (W) masing-masing dokumen diketahui, maka dilakukan proses sorting / pengurutan dimana semakin besar nilai W , semakin besar tingkat similaritas dokumen tersebut terhadap kata yang dicari, demikian sebaliknya.

2.3.6 Vector Space Model

Vector space model (Model Ruang Vektor) adalah suatu model yang digunakan untuk mengukur kemiripan antara suatu dokumen dengan suatu query. Setiap sel dalam matriks bersesuaian dengan bobot yang diberikan dari

suatu term dalam dokumen yang ditentukan. Nilai nol berarti bahwa term tersebut tidak hadir di dalam dokumen.

Pembobotan lokal hanya berpedoman pada frekuensi munculnya term dalam suatu dokumen dan tidak melihat frekuensi kemunculan term tersebut di dalam dokumen lainnya.

$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

Pada algoritma vector space model gunakan rumus untuk mencari nilai cosinus sudut antara dua vector dari setiap bobot dokumen dan bobot dari kata kunci.

Setelah mendapatkan cosine tiap-tiap dokumen, maka hasil bobot dari kata kunci diurutkan. Bobot yang besar menjadi prioritas sebagai document yang memiliki hubungan dengan kata kunci. Penghitungan jumlah kata per artikel yang bersangkutan.

2.4 Pemrograman PHP

PHP merupakan program aplikasi yang bersifat server side, artinya berjalan pada sisi server saja dan tidak dapat berfungsi tanpa adanya sebuah server di dalamnya. *PHP* juga bukan sebuah bahasa pemrograman yang lengkap. Maksudnya program ini tidak menyertakan sebuah *compiler* tersendiri yang membuat program hasilnya menjadi program .exe yang dapat dijalankan tersendiri. Program ini akan selalu membutuhkan sebuah

server pendukung yang disebut *web server* dan program *PHP* itu sendiri untuk menjalankan *script* program.

2.4.1 Kelebihan PHP

Banyak sekali kelebihan yang dimiliki PHP dibandingkan dengan bahasa pemrograman yang lain, Diantaranya :

- a. Web menjadi Dinamis.
- b. PHP bersifat Open Source yang berarti dapat digunakan oleh siapa saja secara bebas.
- c. Program yang dibuat dengan PHP bisa dijalankan oleh Semua Sistem Operasi karena PHP berjalan secara Web Base yang artinya semua Sistem Operasi bahkan HP terbaru yang mempunyai Web Browser dapat menggunakan program PHP.
- d. Aplikasi PHP lebih ringan dibandingkan dengan ASP maupun Java sehingga berjalan lebih cepat.
- e. Mendukung banyak paket Database seperti MySQL, Oracle, PostgreSQL, dan lain-lain.
- f. Bahasa pemrograman PHP tidak memerlukan Kompilasi / Compile dalam penggunaannya.
- g. Banyak Web Server yang mendukung PHP seperti Apache, Lighttpd, IIS dan lain-lain.
- h. Pengembangan Aplikasi PHP mudah karena banyak Dokumentasi, Refrensi & Developer yang membantu dalam pengembangannya.

- i. Banyak bertebaran Aplikasi & Program PHP yang Gratis & Siap pakai seperti WordPress, PrestaShop, dan lain-lain.

2.4.2 Kekurangan PHP

Selain kelebihan PHP, PHP juga mempunyai kekurangan. Namun masalah kekurangannya sangat sedikit. Diantaranya :

- a. PHP Tidak mengenal Package.
- b. Jika tidak di encoding, maka kode PHP dapat dibaca semua orang & untuk meng encodingnya dibutuhkan tool dari Zend yang mahal sekali biayanya.
- c. PHP memiliki kelemahan keamanan. Jadi Programmer harus jeli & berhati-hati dalam melakukan pemrograman & Konfigurasi PHP.

2.5 Mysql

Mysql adalah sebuah server database open source yang terkenal yang digunakan berbagai aplikasi terutama untuk server atau membuat WEB. Mysql berfungsi sebagai SQL (Structured Query Language) yang dimiliki sendiri dan sudah diperluas oleh Mysql umumnya digunakan bersamaan dengan PHP untuk membuat aplikasi server yang dinamis dan powerfull.

MySQL adalah sebuah implementasi dari sistem manajemen basisdata relasional (RDBMS) yang didistribusikan secara gratis dibawah lisensi GPL (General Public License). MySQL sebenarnya merupakan turunan salah satu konsep utama dalam basisdata yang telah ada sebelumnya; SQL(Structured Query Language).

2.6 Webservice

Webservice atau server adalah komputer yang dikhususkan untuk menaruh data website. Semua komputer logikanya bisa dijadikan server, namun server yang khusus untuk website punya spesifikasi khusus. Hal ini untuk menjaga agar kecepatan server stabil dan memenuhi kriteria yang dihadapkan. Selain bergantung pada spesifikasi hardware dan software, kecepatan webservice juga bergantung pada akses internet di tempat server berada atau juga tergantung pada kecepatan Internet Service Provider (ISP) yang digunakan untuk online, dan terlebih lagi juga bergantung pada jalur lokasi negara yang bersangkutan.

Web server menjadi tulang belakang dari world wide web (www). Web server menunggu permintaan dari client yang menggunakan browser seperti Netscape Navigator, Internet Explorer, Mozilla, dan program browser lainnya. Jika ada permintaan dari browser, maka web server akan memproses permintaan itu kemudian memberikan hasil prosesnya berupa data yang diinginkan kembali ke browser. Data ini mempunyai format yang standar, disebut dengan format SGML (standar general markup language). Data yang berupa format ini kemudian akan ditampilkan oleh browser sesuai dengan kemampuan browser tersebut. Contohnya, bila data yang dikirim berupa gambar, browser yang hanya mampu menampilkan teks (misalnya lynx) tidak akan mampu menampilkan gambar tersebut, dan jika ada akan menampilkan alternatifnya saja.

Web server, untuk berkomunikasi dengan client-nya (web browser) mempunyai protokol sendiri, yaitu HTTP (hypertext transfer protocol). Dengan protokol ini, komunikasi antar web server dengan client-nya dapat saling dimengerti dan lebih mudah. Seperti telah dijelaskan diatas, format data pada world wide web adalah SGML. Tapi para pengguna internet saat ini lebih banyak menggunakan format HTML (hypertext markup language) karena penggunaannya lebih sederhana dan mudah dipelajari.

Kata HyperText mempunyai arti bahwa seorang pengguna internet dengan web browsernya dapat membuka dan membaca dokumen-dokumen yang ada dalam komputernya atau bahkan jauh tempatnya sekalipun. Hal ini memberikan cita rasa dari suatu proses yang tridimensional, artinya pengguna internet dapat membaca dari satu dokumen ke dokumen yang lain hanya dengan mengklik beberapa bagian dari halamanhalaman dokumen (web) itu. Proses yang dimulai dari permintaan webclient (browser), diterima web server, diproses, dan dikembalikan hasil prosesnya oleh web server ke web client lagi dilakukan secara transparan. Setiap orang dapat dengan mudah mengetahui apa yang terjadi pada tiap-tiap proses. Secara garis besarnya web server hanya memproses semua masukan yang diperolehnya dari web clientnya.

2.7 Apache

Server HTTP Apache atau Server Web/WWW Apache adalah server web yang dapat dijalankan di banyak sistem operasi (Unix, BSD, Linux, Microsoft Windows dan Novell Netware serta platform lainnya) yang berguna untuk melayani dan memfungsikan situs web. Protokol yang digunakan untuk melayani fasilitas web/www ini menggunakan HTTP.





BAB III

DESAIN DAN PERANCANGAN SISTEM

3.1 Perancangan Sistem

Dalam pembahasan analisa sistem ini bertujuan mengidentifikasi sistem yang akan dibuat, yang meliputi perangkat lunak serta perangkat keras. Tahapan-tahapan yang menyusun analisis sistem ini adalah analisis data yang dipakai, spesifikasi kebutuhan sistem, spesifikasi pengguna, perancangan basis data (ERD) dan perancangan antar muka.

3.1.1 Data

Data pengujian yang digunakan sebagai data penelitian adalah data penggunaan obat dalam situs ensiklopedia obat herbal dan buku tanaman tradisional trubus.

3.1.2 Spesifikasi Sistem

Analisis kebutuhan merupakan analisis terhadap komponen-komponen yang digunakan untuk pembuatan sistem. Dalam hal ini, komponen yang dibutuhkan terbagi menjadi dua macam, yaitu komponen perangkat lunak dan perangkat keras.

3.1.2.1 Perangkat Lunak

Perangkat lunak adalah program atau aplikasi komputer lain yang dibutuhkan untuk membangun sistem. Perangkat lunak yang dibutuhkan dalam pembuatan sistem ini adalah :

1) Sistem Operasi

Sistem Operasi yang digunakan penulis adalah Windows Seven. Windows Seven, merupakan sistem operasi yang menjembatani antara computer dan user. Sistem operasi ini digunakan karena dukungan dan kecocokan dengan software yang digunakan dalam pembuatan sistem ini.

2) Aplikasi Server

Aplikasi server yang peneliti gunakan adalah AppServ.AppServ merupakan seperangkat paket yang memungkinkan pembuatan server.

3) Web Server

Web Server yang digunakan oleh peneliti adalah Apache. Apache sebagai web server, berfungsi untuk menampilkan halaman web yang benar sesuai request kepada pengguna.

4) Database

MySql merupakan database server untuk pengelolaan database.

5) Tool Uploader

FileZilla sebagai perangkat untuk memindahkan data dari pengguna ke server agar lebih mudah.

6) Editor PHP

Merupakan aplikasi yang memudahkan untuk pengetikan code dan mempermudah peneliti dalam menemukan kesalahan pengetikan. Editor yang digunakan adalah netbean.

7) Browser

Browser merupakan perangkat yang berfungsi untuk menampilkan html dari server.

3.1.2.2 Perangkat Keras

Perangkat keras adalah perangkat fisik yang digunakan untuk menjalankan sistem. Dalam pembuatan sistem ini, perangkat keras yang digunakan yaitu laptop dengan spesifikasi sebagai berikut :

1. Processor core i5
2. RAM DDR3 4 GB
3. Hard Disk 500 GB
4. Monitor 14 inchi

3.1.3 Spesifikasi Pengguna

Aplikasi ini dapat digunakan oleh semua pihak yang ingin mencari obat herbal berdasarkan kata kunci berupa gejala penyakit yang diderita dengan lebih mudah. Sehingga penggunaanya adalah user yang dapat menambah data obat jika dikemudian hari ada penemuan obat herbal baru.

3.1.4 Lingkungan Operasi

Perangkat lunak yang dibangun ini membutuhkan perangkat lunak lain sebagai penunjang agar sistem ini berjalan dengan sesuai dengan fungsinya. Kebutuhan tersebut diantaranya:

a. Antar muka pemakai

Sebagai penunjang antarmuka pemakai dari perangkat lunak, diperlukan web browser, seperti: Internet Explorer 6, Mozilla Firefox 1.5, atau di atasnya.

b. Antar muka komunikasi

Untuk komunikasi antar server dan client, maka diperlukan jaringan internet karena sistem yang dibangun berbasis web.

c. Aplikasi server

Dibutuhkan dua jenis server untuk memusatkan proses dari perangkat lunak, yaitu web server dan database server. Untuk web server-nya dapat menggunakan aplikasi Apache dan untuk database server-nya menggunakan MySQL. Aplikasi server ini akan mengatur request ke server dan juga respon terhadap request dari server ke client.

3.1.5 Deskripsi Sistem

Dalam subbab ini akan dibahas sistem yang akan dibangun dalam skripsi ini. Sistem yang dibangun adalah pengembangan algoritma pencarian dan perankingan obat tradisional berdasarkan gejala penyakit menggunakan metode cosine similarity. Tujuan sistem ini adalah mengukur bobot dan keterkaitan kata kunci dengan gejala obat herbal. Diharapkan dengan penggunaan metode cosine

similarity ini sistem ini dapat mempermudah pencarian obat herbal berdasarkan kata kunci yang diinginkan.

Teknik yang digunakan pada sistem ini adalah text mining. Tahapan dalam sistem ini terdiri dari tahapan text preprocessing, text transformation, dan pembobotan dokumen yang dilakukan oleh user. Tahapan pencarian gejala dimulai dengan proses text preprocessing, text transformation, dan pattern discovery yang terdiri dari pembobotan terhadap kata kunci. Pembobotan menggunakan term-frequency – inversed document frequency (TF-IDF) dan algoritma vector space model (VSM) dengan metode cosine similarity.

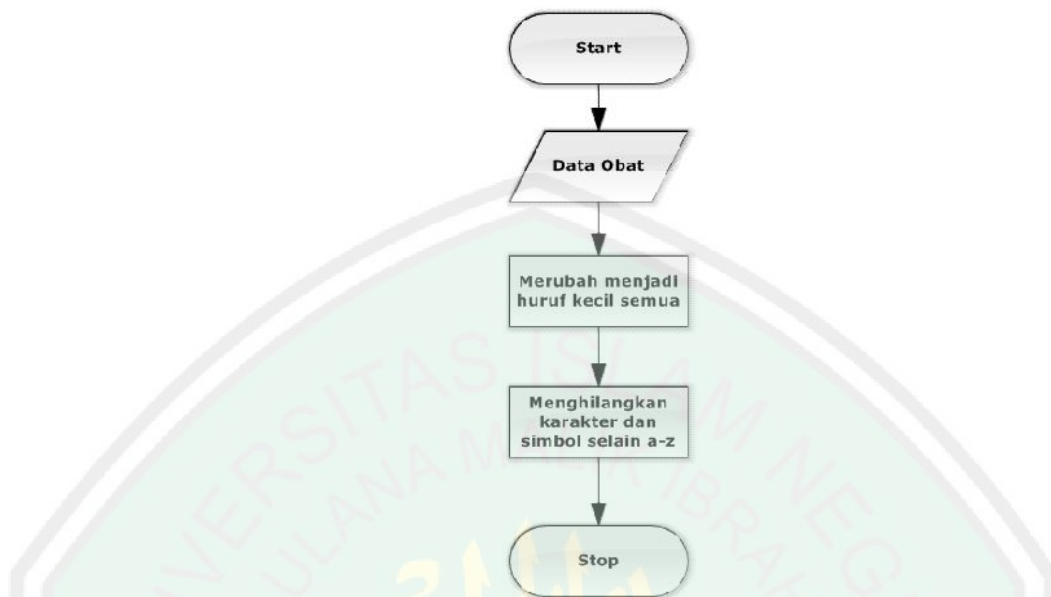
3.2 Desain Sistem

3.2.1 Ekstraksi Dokumen

Proses ini merupakan proses awal sebelum data dapat dianalisa. Proses ini terdiri dari Case Folding dan Tokenizing.

3.2.1.1 Case Folding

Proses ini melakukan penghilangan kata. Perubahan huruf kapital menjadi huruf kecil. Dan penghilangan symbol-symbol yang dianggap tidak penting, seperti pada gambar 3.1.



Gambar 3.1 Case Folding

3.2.1.2 Tokenizing

Tokenizing merupakan proses pemotongan data input berupa string menjadi kata-kata tunggal berdasarkan spasi. Seperti pada gambar 3.2



Gambar 3.2 Tokenizing

3.2.1.3 Filtering

Setelah proses tokenizing, data yang dihasilkan berupa kata-kata tunggal. Namun kata-kata tunggal itu masih terdiri dari kata-kata penting dan kata-kata tidak penting. Kata penting yang dimaksud adalah kata-kata yang nantinya akan dianalisa, sedangkan kata tidak penting merupakan kata-kata berupa kata depan dan kata sambung. Proses filtering ini mengacu pada data kata tidak penting yang sudah disediakan oleh peneliti.



Gambar 3.3 Filtering

3.2.2 Analisa

Tahap analisa merupakan tahap yang menentukan keterkaitan obat dengan kata kunci berupa gejala. Dalam tahap analisa ini peneliti menggunakan algoritma TF/IDF dan metode Vector Space.

a. TF/IDF

TF/IDF merupakan pengukuran statistik yang dimaksudkan untuk mengukur keterkaitan sebuah kata dengan sekumpulan dokumen. Proses ini digunakan sebagai factor bobot untuk mencari informasi pada dokumen-dokumen.

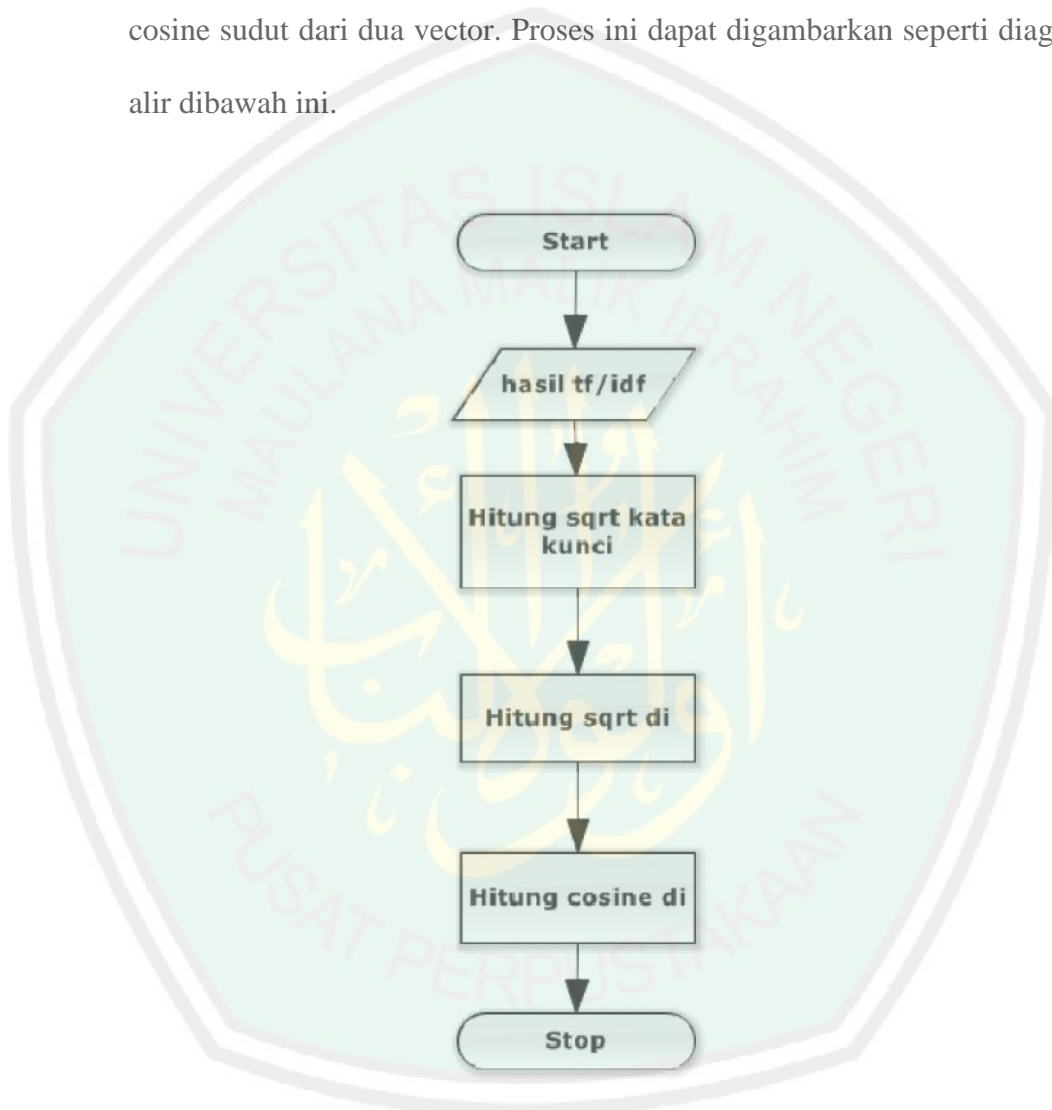
b. Vector Space Model

Model aljabar untuk merepresentasikan teks berupa indek yang dibentuk sebagai vector. Proses ini digunakan untuk mengukur keterkaitan kata terhadap dokumen.

Setelah bobot setiap dokumen telah diketahui hasilnya, maka dilakukan proses pengurutan berdasarkan bobot yang didapat. Sehingga akan menghasilkan urutan dokumen berdasarkan bobot. Semakin besar bobot semakin besar nilai kemiripan terhadap kata kunci, demikian pula sebaliknya. Sehingga menghasilkan urutan / ranking dokumen yang mirip. Jika terdapat hasil bobot yang sama maka perlu dilakukan proses vector space model

c. Cosine Similarity

Setelah ditemukan hasilnya, maka pengukuran akan menghitung cosine sudut dari dua vector. Proses ini dapat digambarkan seperti diagram alir dibawah ini.



Gambar 3.4 Cosine Similarity

3.3 Contoh Perhitungan Manual

Obat 1 :

Nama Obat : Mengkudu

Nama Ilmiah : *Bancudus latifolia*, Rumph.

Nama Lokal : Mengkudu (Indonesia), Pace, Kemudu, Kudu (Jawa);
Cengkudu (Sunda), Kodhuk (Madura), Wengkudu (Bali)

Keterangan : Mengkudu (*MORINDA CITRIFOLIA*) termasuk jenis kopi-kopian. Mengkudu dapat tumbuh di dataran rendah sampai pada ketinggian tanah 1500 meter diatas permukaan laut. Mengkudu merupakan tumbuhan asli dari Indonesia. Tumbuhan ini mempunyai batang tidak terlalu besar dengan tinggi pohon 3-8 m. Daunnya bersusun berhadapan, panjang daun 20-40 cm dan lebar 7-15 cm. Bunganya berbentuk bungan bongkol yang kecil-kecil dan berwarna putih. Buahnya berwarna hijau mengkilap dan berwujud buah buni berbentuk lonjong dengan variasi trotol-trotol. Bijinya banyak dan kecil-kecil terdapat dalam daging buah. Pada umumnya tumbuhan mengkudu berkembang biak secara liar di hutan-hutan atau dipelihara orang pinggiran-pinggiran kebun rumah

Manfaat : Hipertensi, Sakit kuning, Demam, Influenza, Batuk, Sakit perut; Menghilangkan sisik pada kaki

Kandungan : kandungan zat tersebut antara lain morinda diol, morindone, morindin, damnacanthal, metil asetil, asam ka

Cara Pemakaian : 1. Hipertensi

Bahan : 2 buah Mengkudu yang telah masak di pohon dan 1 sendok makan madu.

Cara Membuat : buah mengkudu diperas untuk diambil airnya, kemudian dicampur dengan madu sampai merata dan disaring.

Cara menggunakan: diminum dan diulangi 2 hari sekali.

Obat 2 :

Nama Obat : Tapak Dara

Nama Ilmiah : *Catharantus roseus* (L.) G. Don.

Nama Lokal : Perwinkle (Inggris), Chang Chun Hua (Cina); Keminting Cina, Rumput Jalang (Malaysia); Tapak Dara (Indonesia), Kembang Sari Cina (Jawa); Kembang Tembaga Beureum (Sunda);

Keterangan : tapakdara sering dibedakan menurut jenis bunganya, yaitu putih dan merah. Tumbuhan semak tegak yang dapat mencapai ketinggian batang sampai 100 cm ini, sebenarnya merupakan tumbuhan liar yang biasa tumbuh subur di padang atau dipedesaan beriklim tropis. Ciri-ciri tumbuhan Tapakdara : memiliki batang yang berbentuk bulat dengan diameter berukuran kecil, berkayu, beruas dan bercabang serta berambut. Daunnya berbentuk bulat telur, berwarna hijau dan diklasifikasikan berdaun tunggal. Bunganya yang indah menyerupai terompet dengan permukaan berbulu halus. Tapakdara juga memiliki rumah biji yang berbentuk silindris menggantung pada batang. Penyebaran tumbuhan ini melalui biji.

Manfaat : Diabetes, Hipertensi, Leukimia, Asma, Bronkhitis, Demam; Radang Perut, Disentri, Kurang darah, Gondong, Bisul, Borok; Luka Bakar, Luka baru, Bengkak;

Kandungan : vinkristin, vinrosidin, vinblastin dan vinleurosin merupakan kandungan komposisi zat alkaloid dari tapakdara.

Cara Pemakaian : 1. Hipertensi (tekanan darah tinggi)

Bahan: 15 - 20 gram daun tapakdara kering, 10 gram bunga krisan

Cara membuat: direbus dengan 2,5 gelas air sampai mendidih dan disaring.

Cara menggunakan: diminum tiap sore.

Bahan: 7 lembar daun atau bunga tapakdara.

Cara membuat: diseduh dengan 1 gelas air dan dibiarkan beberapa saat dan disaring

Cara menggunakan: diminum menjelang tidur.

Kata Kunci : Hipertensi, batuk

3.3.1 Preprocessing

1) Case Folding

Mengambil manfaat setiap data obat,

Hasil Case Folding

Data Obat 1 :hipertensi, sakit kuning, demam, influenza, batuk, sakit perut;
menghilangkan sisik pada kaki

Data Obat 2 :diabetes, hipertensi, leukimia, asma, bronkhitis, demam;
radang perut, disentri, kurang darah, gondong, bisul, borok; luka bakar,
luka baru, bengkak;

2) Tokenizing

Hasil dari proses tokenizing terlihat pada table 3.1

1	Hipertensi	16	asma
2	Sakit	17	bronkhitis
3	Kuning	18	demam
4	Demam	19	radang
5	Influenza	20	disentri
6	Batuk	21	kurang
7	Sakit	22	darah
8	Perut	23	gondong
9	menghilangkan	24	bisul
10	Sisik	25	borok
11	Pada	26	luka
12	Kaki	27	bakar
13	Diabetes	28	luka
14	Hipertensi	29	baru
15	Leukemia	30	bengkak

Table 3.1 Hasil Tokenizing

3) Filtering

Hasil proses filtering terlihat pada tabel 3.2.

1	Hipertensi	16	bronkhitis
2	Sakit	17	demam
3	Kuning	18	radang
4	Demam	19	disentri
5	Influenza	20	kurang
6	Batuk	21	darah
7	Sakit	22	gondong
8	Perut	23	bisul
9	menghilangkan	24	borok
10	Sisik	25	luka
11	Kaki	26	bakar
12	Diabetes	27	luka
13	Hipertensi	28	baru
14	Leukemia	29	bengkak
15	Asma		

Tabel 3.2 Hasil Filtering

3.3.2 Analisa

Dokumen / data obat yang digunakan berjumlah 2. Dengan kata kunci / gejala penyakit yaitu hipertensi dan batuk.

3.3.2.1 TF/IDF

Tf = term frekuensi (jumlah kata tertentu yang ada dalam dokumen tertentu)

Df=document frekuensi (jumlah kata tertentu dalam satu dokumen)

Idf=log(jumlah dokumen / df)

	Q	D1	D2	df	n/df	idf=log(n/df)
Hipertensi	1	1	1	3	1	0
Sakit		2		2	1.5	0.176091259
Kuning		1		1	3	0.477121255
Demam		1	1	2	1.5	0.176091259
Influenza		1		1	3	0.477121255
Batuk		1		1	3	0.477121255
Perut		1	1	2	1.5	0.176091259
Hilang		1		1	3	0.477121255
Sisik		1		1	3	0.477121255
Kaki		1		1	3	0.477121255
Diabetes	1			2	1.5	0.176091259
Leukemia				1	3	0.477121255
Asma				1	3	0.477121255
Bronchitis				1	3	0.477121255
Radang				1	3	0.477121255
Disentri				1	3	0.477121255
Kurang				1	3	0.477121255
Darah				1	3	0.477121255
Gondong				1	3	0.477121255
Bisul				1	3	0.477121255
Borok				1	3	0.477121255
Luka				1	3	0.477121255
Bakar				1	3	0.477121255
Bengkak				1	3	0.477121255
Paru				1	3	0.477121255

Tabel 3.3 Hasil Perhitungan tf

Menghitung idf (inverse document frequency)

	idf=log(n/df)	d1	w	d2	d3
Hipertensi	0	0	0	0	0
Sakit	0.176091259	0	0.352183	0	0
Kuning	0.477121255	0	0.477121	0	0
Demam	0.176091259	0	0.176091	0.176091	0.176091
Influenza	0.477121255	0	0.477121	0	0
Batuk	0.477121255	0	0.477121	0	0
Perut	0.176091259	0	0.176091	0.176091	0.176091
Hilang	0.477121255	0	0.477121	0	0
Sisik	0.477121255	0	0.477121	0	0
Kaki	0.477121255	0	0.477121	0	0
Diabetes	0.176091259	0.176091	0	0.176091	0.176091
Leukemia	0.477121255	0	0	0.477121	0.477121
Asma	0.477121255	0	0	0.477121	0.477121
Bronchitis	0.477121255	0	0	0.477121	0.477121
Radang	0.477121255	0	0	0.477121	0.477121
Disentri	0.477121255	0	0	0.477121	0.477121
Kurang	0.477121255	0	0	0.477121	0.477121
Darah	0.477121255	0	0	0.477121	0.477121
Gondong	0.477121255	0	0	0.477121	0.477121
Bisul	0.477121255	0	0	0.477121	0.477121
Borok	0.477121255	0	0	0.477121	0.477121
Luka	0.477121255	0	0	0.477121	0.477121
Bakar	0.477121255	0	0	0.477121	0.477121
Bengkak	0.477121255	0	0	0.477121	0.477121
Paru	0.477121255	0	0	0.477121	0.477121

Tabel 3.4 Hasil Perhitungan IDF

3.3.2.2 Cosine Similarity

Pada tahap ini akan dihitung, cosine similarity antar dokumen

Penghitungan :

D_i = dokumen ke i

$$\text{Cosine}(D_i) = \frac{\sum(kk \times D_i)}{\sqrt{kk^2} \times \sqrt{D_i^2}}$$

Perhitungan bobot tiap dokumen :

$$\text{Cosine D1} = 0 / (0.176091 \times 1.24576)$$

$$= 0$$

$$\text{Cosine D2} = 0.031008 / (0.176091 \times 1.811091)$$

$$= 0.318917$$

3.4 Perancangan Aplikasi

Pada subbab ini akan dijelaskan perancangan sistem yang akan dibangun. Sistem yang akan dibangun berbasis web dengan menggunakan php sebagai bahasa pemrogramannya.

3.4.1 Desain Database

1) Tabel Stoplist

Tabel ini berfungsi menyimpan kata-kata yang tidak penting. Tabel ini digunakan saat proses filtering.

Nama Field	Tipe Data
Katatidakpenting	Varchar

Tabel 3.5 Rancangan Tabel Stoplist

2) Tabel Data Obat

Tabel ini berfungsi untuk menyimpan data obat.

Nama Field	Type Data
Id_obat	Int
Nama_obat	Varchar
Ilmiah_obat	Varchar
Lokal_obat	Varchar
Uraian_obat	Varchar
Gejala_obat	Varchar
Manfaat_obat	Varchar
Cara_obat	Varchar
Gambar_obat	Varchar
Tanggal_obat	Date
Penulis_obat	Varchar
Status	Varchar

Tabel 3.6 Rancangan Tabel Data Obat

3) Tabel User

Tabel ini digunakan untuk manajemen pengguna

Nama Field	Tipe Data
Id_user	Int
Username	Varchar
Password	Varchar
nama_lengkap	Varchar

Tabel 3.7 Rancangan Tabel Data User

3.4.2 Desain Menu



Gambar 3.6 Sitemap

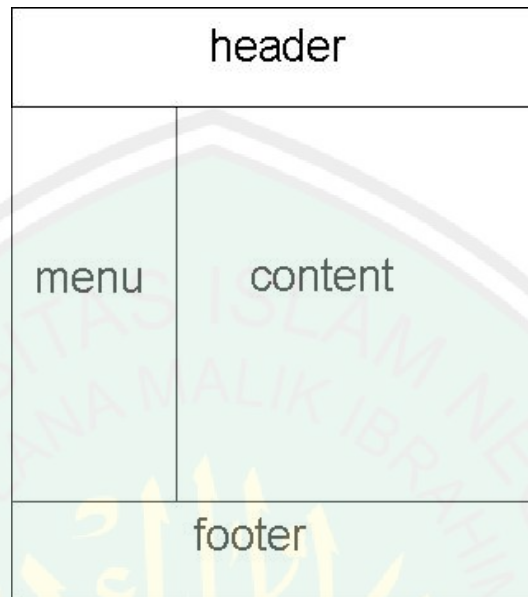
3.4.3 Perancangan Antarmuka

1. Login

header	
username : <input type="text"/> password : <input type="password"/> <input type="button" value="login"/>	
footer	

Gambar 3.7 Perancangan Login

2. Halaman



Gambar 3.8 Perancangan Halaman

3. Form Tambah Obat

The form for adding a drug contains the following fields and a button:

Nama Obat	:	<input type="text"/>
Nama Ilmiah	:	<input type="text"/>
Nama Lokal	:	<input type="text"/>
Uraian Obat	:	<input type="text"/>
Gejala Obat	:	<input type="text"/>
Kandungan Obat	:	<input type="text"/>
Penyajian Obat	:	<input type="text"/>
Literatur Obat	:	<input type="text"/>
Gambar Obat	:	<input type="text"/>
Kategori Obat	:	<input type="text"/>
Status Obat	:	<input type="text"/>
		<input type="button" value="Tambah"/>

Gambar 3.9 Perancangan Halaman Tambah Obat



BAB IV

HASIL DAN PEMBAHASAN

4.1 Implementasi

Merupakan proses pembangunan komponen-komponen pokok sebuah sistem berdasarkan desain yang sudah di buat. Implementasi sistem juga merupakan sebuah proses pembuatan dan penerapan sistem secara utuh baik dari sisi perangkat keras maupun perangkat lunaknya.

Implementasi yang akan dijelaskan disini meliputi lingkungan perangkat keras dan lingkungan perangkat lunak.

4.1.1 Ruang Lingkup Perangkat Keras

Perangkat keras yang digunakan untuk menjalankan aplikasi yang berbasis php ini adalah :

1. Laptop core i5
2. Memory 4 GBytes
3. Hardisk 500 GBytes

4.1.2 Ruang Lingkup Perangkat Lunak

Perangkat lunak yang dibutuhkan agar mampu menjalankan aplikasi adalah sebagai berikut:

1. Instalasi Microsoft Windows Seven
2. Instalasi Appserv 5.2.9
3. Instalasi Adobe Dreamweaver

4. Instalasi Mozilla Firefox 3.6

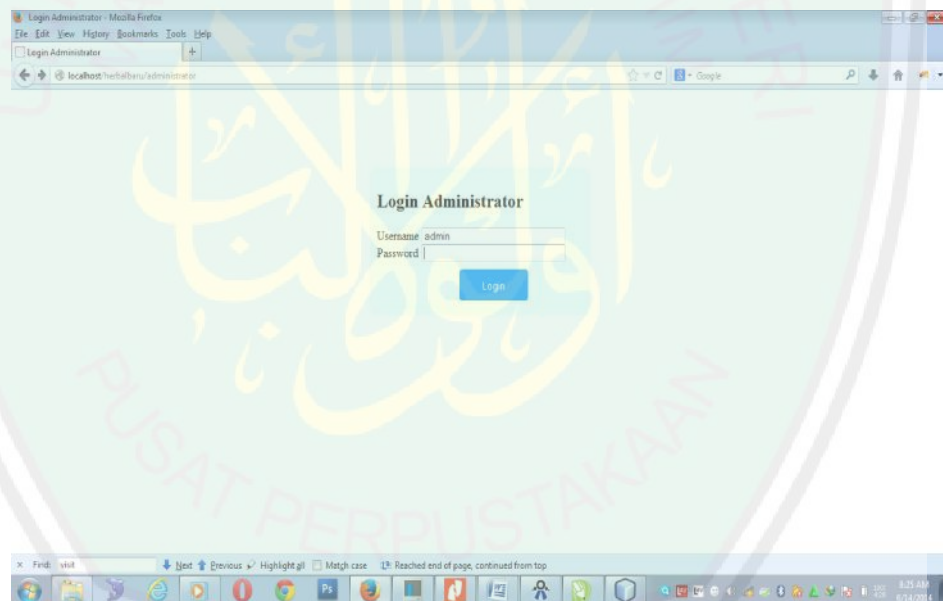
4.1.3 Implementasi Desain Interface

a. Tampilan Utama

Tampilan utama ini merupakan tampilan awal ketika aplikasi dijalankan pertama kali. Dalam tampilan utama ini, diberikan form untuk memasukkan jenis gejala penyakit yang dirasakan

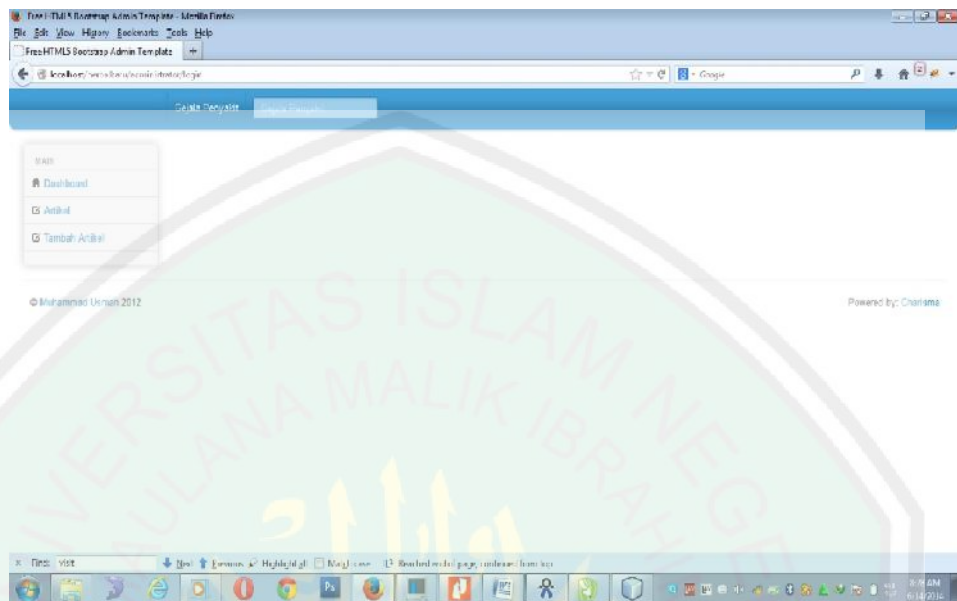
Implementasi Desain Interface

1. Halaman Login



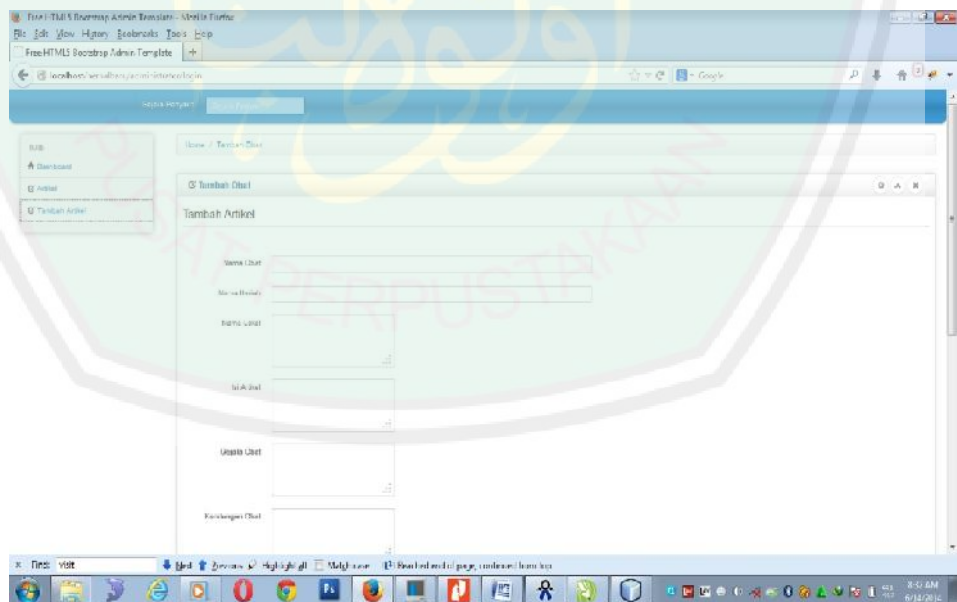
Gambar 4.1 Halaman Login

2. Halaman Admin

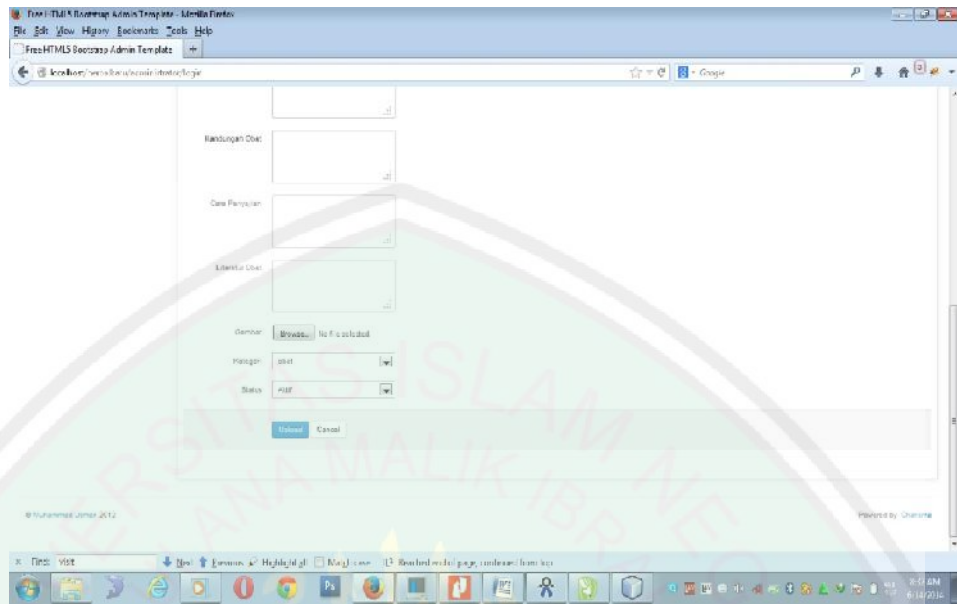


Gambar 4.2 Halaman Admin

3. Halaman Tambah Obat

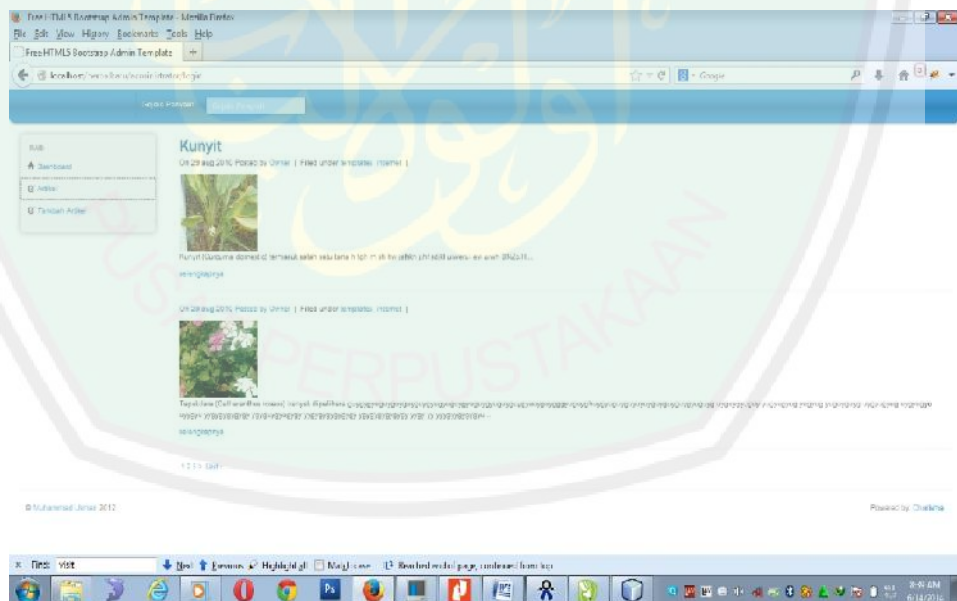


Gambar 4.3 Halaman Tambah Obat 1



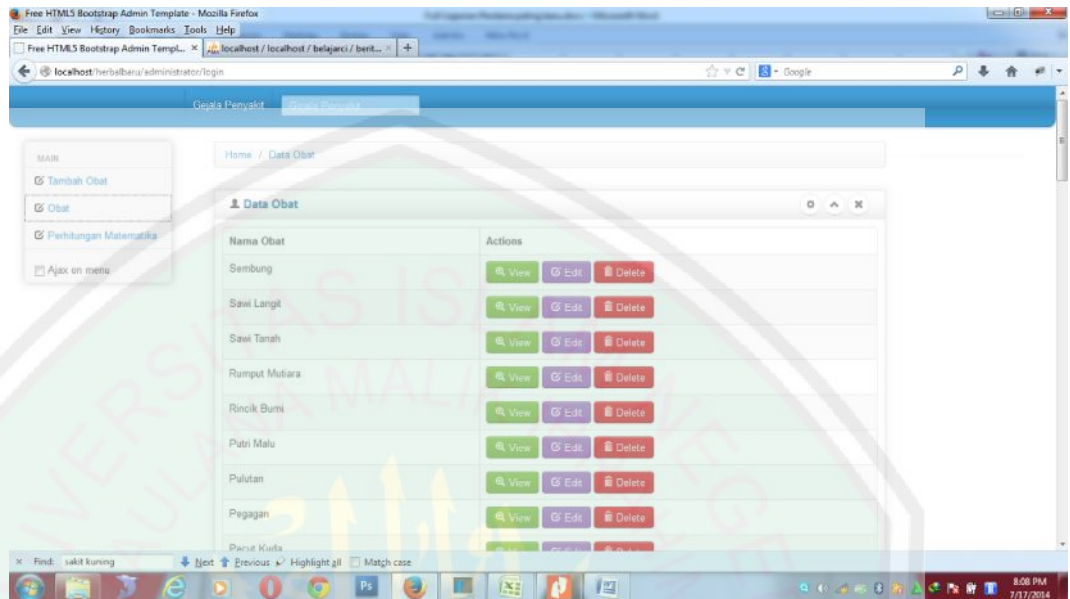
Gambar 4.4 Halaman Tambah Obat 2

4. Lihat Obat



Gambar 4.5 Halaman Lihat Obat

5. Lihat Edit Delete Obat



4.2 Pembahasan

Pada tahap implementasi ini, akan membahas algoritma yang telah dirancang pada bab 3 di atas, yaitu metode text mining yang terdiri dari preprocessing dan analisa. Preprocessing terdiri dari case folding, tokenizing dan filtering. Kemudian untuk menghitung bobot digunakan metode analisa TF/IDF dan Vector Space Model.

4.2.1 Preprocessing

Dalam tahap preprocessing atau disebut sebagai ekstraksi dokumen, tahap - tahap yang dilakukan adalah case folding, tokenizing, filtering.

4.2.1.1 Case Folding

Pada tahap ini dokumen sebagai masukan akan dirubah menjadi huruf kecil dan karakter yang diterima adalah huruf „a“ sampai dengan „z“. Sedangkan karakter lain seperti symbol dan angka akan dihilangkan.

```

function case_folding($dok){
    $dok=" ".$dok." ";
    //menghilangkan karakter enter
    $dok = str_replace(array("\r","\n")," ",$dok);
    if(substr($dok,strlen($dok)-1,strlen($dok))=="/./"){
        $dok1=preg_replace("/./"," ",$dok);
        $dok=$dok1;
    }
    $lower=strtolower($dok);
    $pattern = "/([a-zA-Z\s][-])/i";
    preg_match_all($pattern, $lower,$matches);
    if (count($matches)>=2){
        $string = implode("",$matches[1]);
    }
    else {$string= "";}      return $string; }

```

Pada kode di atas, pada awal proses algoritma akan menghilangkan karakter enter sehingga apabila dokumen terdiri dari beberapa gejala obat maka akan menjadi satu string panjang satu paragraf. Selanjutnya dokumen akan diubah menjadi huruf kecil dengan fungsi `strtolower($string)` dan menfilter seluruh karakter kecuali „a“ sampai dengan „z“ dengan menggunakan regex (regular expression) dengan menggunakan pattern `"([a-zA-Z\s][-])/i`. Dan jika kode diatas di-implementasikan pada dokumen yang berisi huruf, angka, dan tanda baca, maka hasilnya adalah hanya huruf a-z saja.

4.2.1.2 Tokenizing

Pada tahap ini, hasil dari case folding kemudian dipecah menjadi kata dengan menggunakan delimiter spasi.

```
function tokenizing($dok){
    $string=case_folding($dok);
    $token=explode(" ",$string);
    return $token;
}
```

Untuk pemecahan menjadi kata, php telah memberikan fungsi yang mudah yaitu `explode($delimiter,$string)`.

4.2.1.3 Filtering

Tahap filtering adalah membuang kata-kata yang dianggap tidak penting. Sebelum filtering dilakukan maka, diperlukan fungsi untuk mengambil stoplist dari database. Kumpulan kata-kata yang masuk dalam stoplist diambil dari http://fpmipa.upi.edu/staff/yudi/stop_words_list.txt. Kemudian dibuat fungsi lagi untuk filtering.

```
function stopword(){
    $i=0;
    $query=mysql_query("SELECT * from stoplist");
    while($word=mysql_fetch_array($query)){
        $stoplist[$i]="/ ".$word[kata]."/";
        $i++;}
}
```



```
return $stoplist;}

function filtering($dok){
    $stoplist=stopword();
    $tokenizing=tokenizing($dok);
    $dok=explode(" ",$tokenizing);
    $string=preg_replace($stoplist," ",$dok);
    $token=tokenizing($string);return $token;}
}
```

Fungsi `stopword()` di atas digunakan untuk pengambilan kata-kata yang tidak penting dari database. Nilai dari `stoplist` dikembalikan ke fungsi `stopword()` dengan fungsi `return`. Fungsi dibawahnya adalah `filtering()`. Parameter untuk fungsi ini adalah data yang berupa array. Dalam proses awal data array tersebut akan dijadikan menjadi satu string yang gunanya adalah untuk mempermudah pemrosesan penghapusan `stoplist`. Untuk menghapus `bag of words`, maka digunakan fungsi `preg_replace()`. Dengan fungsi ini maka array `stoplist` akan dibandingkan dengan string dari array data. Jika ada yang sama, maka kata dalam string akan dihapus dengan cara mengganti kata tersebut dengan karakter kosong (“”). Setelah selesai penghapusan, maka string tersebut dijadikan array kata dengan pemotongan dengan fungsi `tokenizing()`.

4.2.2 Analisa

Untuk tahap analisa hasil pre processing di atas, maka dilanjutkan dalam penghitungan bobot tiap kata. Perhitungan bobot ini dihitung menggunakan algoritma TF/IDF dan Vector Space Model.

4.2.2.1 Algoritma TF/IDF

Dalam implementasi analisa ini ke dalam program, analisa TF/IDF dibagi menjadi beberapa fungsi. Fungsi-fungsi ini diantaranya fungsi `tf()`, `IDF()`, dan `bobot()`.

```
function tf($kunci,$kalimat){
    $kata=explode(" ",$kalimat);
    $term=0;
    foreach($kata as $word){
        if($word==$kunci){
            $term+=1;
        }
    }
    return $term;}

```

Fungsi di atas dibuat untuk menghitung term kata terhadap kata kunci pada tiap dokumen.

Setelah frekuensi kata telah dihitung, maka proses selanjutnya adalah menghitung dokumen yang mengandung kata kunci.

```
function IDF($kunci,$dok){
    $df=0;
    for($i=1;$i<count($dok);$i++){
        $sregi=strpos($kunci,$dok[$i]);
        if($sregi!=null){
            $df+=1;}}
    if($df==0){
        $df=1;}
    $idf=(count($dok)-1)/$df;$IDF=log10($idf);return $IDF;}

```

4.2.2.2 Vector Space Model

Sebagaimana telah dijelaskan pada bab selanjutnya bahwa inti dari algoritma ini adalah penghitungan bobot dokumen terhadap kata kunci. Berikut source code lengkapnya.

```
function token($dok_target){
    $token=array();
    foreach($dok_target as $kalimat){
        $kata=explode(" ",$kalimat);
        for($i=0;$i<count($kata);$i++){
            $hasil=true;
            for($j=0;$j<count($token);$j++){

```

```

if($token[$j]==$kata[$i]){

$hasil=false;

}

}

if($hasil==true){
array_push($token,$kata[$i]);
}}}
return $token;}

```

Fungsi token() di atas berfungsi sebagai pemecah dokumen dari hasil preprocessing menjadi kumpulan kata. Setelah parsing menjadi kata makadilanjutkan dalam fungsi penghitungan bobot terhadap kata kunci.

```

function kk_dot_dokumen($token,$dok,$j){
$kk_dot_dok=0;
for($i=0;$i<count($token);$i++){
$kk=bobot($token[$i],$dok[0],$dok);
$di=bobot($token[$i],$dok[$j],$dok);
$kk_dot_di=$kk*$di;
$kk_dot_dok+=$kk_dot_di;
}
return $kk_dot_dok;
}

function sqrt_dok($token,$dok,$j){

```

```

$jumlah_bobot_kuadrat=0;

for($i=0;$i<count($token);$i++){

$bobot=bobot($token[$i],$dok[$j],$dok);

$bobot_kuadrat=$bobot*$bobot;

$jumlah_bobot_kuadrat+=$bobot_kuadrat;

}

$sqrt_jumlah=sqrt($jumlah_bobot_kuadrat);

return $sqrt_jumlah;

}

function cosine($dok){

$token=token($dok);

$cosine=array();

for($i=0;$i<count($dok);$i++){

$sum=kk_dot_dokumen($token,$dok,$i);

$sqrt=sqrt_dok($token,$dok,0)*sqrt_dok($token,$dok,$i);

// tambahan

if($sqrt==0){

$cosin=0;}

else

{$cosin=$sum/$sqrt;}

}

array_push($cosine,$cosin);}

return $cosine;}

```

Dalam fungsi cosine() di atas adalah perhitungan cosines similarity dalam vector space model. Dalam penelitian ini, perhitungan ditambah bahwa jika perhitungan akar fungsi sqrt() bernilai 0 maka nilai cosine diberi nilai 0. Hal ini diberikan untuk mengatasi permasalahan pembagian 0. Dalam matematika, bahwa semua pembagian 0 adalah bernilai. Logika yang berjalan dalam sebuah pemrograman, maka pembagian 0 dianggap sebagai error programming.

Secara umum dapat disimpulkan bahwa dengan penggunaan algoritma cosine similarity sebagai pembobotan dokumen terhadap kata kunci memberikan pembobotan yang cukup baik dalam peringkat jenis obat tradisional yang cocok dengan gejala penyakit. Hal ini akan sangat membantu kepada para pencari informasi obat tradisional untuk mencari obat tradisional yang berhubungan dengan gejala yang diderita.

4.3 Uji Coba

Uji coba dilakukan menggunakan perhitungan precision dan recall. Perhitungannya (Cios, 2007) seperti berikut ini :

$$\text{Precision} = \frac{\text{dokumen relevan}}{\text{dokumen diterima}} \times 100\% \dots\dots\dots (4.1)$$

$$\text{Recall} = \frac{\text{dokumen relevanquery}}{\text{dokumen relevan}} \dots\dots\dots (4.2)$$

Perhitungan menggunakan kata kunci = “batuk”.

NO	Nama Obat	Gejala Obat	Hubungan	Nilai Cosine
1	Nanas Kerang	Bronkhitis, Batuk, TBC kelenjar, Mimsan, Disentri, Berak darah;	Ya	0.309175
2	Bunga Kenop	Asma, Batuk, Radang mata, Sakit kepala, Mimpu buruk;	Ya	0.294093

		Sakit panas, Disentri;		
3	Mengkudu	Hipertensi, Sakit kuning, Demam, Influenza, Batuk, Sakit perut; Menghilangkan sisik pada kaki	Ya	0.27757
4	Sawi Langit	Demam, panas batuk, Disentri, Hepatitis, Lelah tidak bersemangat; Susah tidur (insomnia)	Ya	0.262985
5	Meniran	Sakit kuning (lever), Malaria, Demam, Ayan, Batuk, Haid lebih; Disentri, Luka bakar, Luka koreng, Jerawat;	Ya	0.234585

Dokumen diterima : 5

Dokumen diterima relevan : 5

Dokumen relevan database : 9

Precision : $\frac{5}{5} \times 100\% = 100\%$

Recall : $\frac{5}{9} \times 100\% = 55\%$

Perhitungan menggunakan kata kunci = “panas batuk”.

NO	Nama Obat	Gejala Obat	Hubungan	Nilai Cosine
1	Sawi Langit	Demam, panas batuk, Disentri, Hepatitis, Lelah tidak bersemangat; Susah tidur (insomnia)	Ya	0.262985
2	Sawi Tanah	Radang saluran nafas, Batuk, TBC, Panas, Campak, Reumatik; Sakit tenggorokan, Hepatitis, Bisul, Memar, Luka berdarah; Gigitan ular, Kencing berkurang;	Ya	0.210001

Dokumen diterima : 2

Dokumen diterima relevan : 2

Dokumen relevan database : 2

Precision : $\frac{2}{2} \times 100\% = 100\%$

Recall : $\frac{2}{2} \times 100\% = 100\%$

Perhitungan menggunakan kata kunci = “panas batuk hipertensi”.

Dokumen diterima : 0

Dokumen diterima relevan : 0

Dokumen relevan database : 0

Precision : $\frac{0}{0} \times 100\% = 100\%$

Recall : $\frac{0}{0} \times 100\% = 100\%$

Perhitungan menggunakan kata kunci = “sakit kepala”.

NO	Nama Obat	Gejala Obat	Hubungan	Nilai Cosine
1	Bunga Kenop	Asma, Batuk, Radang mata, Sakit kepala, Mimpu buruk; Sakit panas, Disentri	Ya	0.294093
2	Kaki Kuda	Typus, Busung, Sakit kepala, Influenza, Keracunan Jengkol, Ayan	Ya	0.277864
3	Jintan/Ajeran	Asma, Batuk, Perut kembung, Sakit kepala, Sariawan, Demam; Luka, Borok	Ya	0.27757
4	Bunga Matahari	Hipertensi, Sakit kepala, Sakit gigi, nyeri menstruasi, reumatik; Nyeri lambung, radang payudara, Sulit melahirkan, Disentri, Campak; Infeksi saluran kencing, Bronkhitis, Batuk, Keputihan, Malaria	Ya	0.194944

Dokumen diterima : 4

Dokumen diterima relevan : 4

Dokumen relevan database : 4

Precision : $\frac{4}{4} \times 100\% = 100\%$

Recall : $\frac{4}{4} \times 100\% = 100\%$

Perhitungan menggunakan kata kunci = “sakit kuning”.

NO	Nama Obat	Gejala Obat	Hubungan	Nilai Cosine
1	Mengkudu	Hipertensi, Sakit kuning, Demam, Influenza, Batuk, Sakit perut; Menghilangkan sisik pada kaki	Ya	0 .26936
2	Meniran	Sakit kuning (lever), Malaria, Demam, Ayan, Batuk, Haid lebih; Disentri, Luka bakar, Luka koreng, Jerawat;	Ya	0 .215963
3	Ngokilo	Tumor, Diabetes melitus, Lever (Sakit kuning), Ambeien (Wasir); Kolesterol, Maag, Kena bisa ulat dan Semut hitam;	Ya	0 .193481
4	Keji Beling	Tumor, Diabetes melitus, Lever (Sakit kuning), Ambeien (Wasir); Kolesterol, Maag, Kena bisa ulat dan Semut hitam;	Ya	0 .193481

Dokumen diterima : 4

Dokumen diterima relevan : 4

Dokumen relevan database : 4

Precision : $\frac{4}{4} \times 100\% = 100\%$

Recall : $\frac{4}{4} \times 100\% = 100\%$

Dari hasil percobaan didapatkan hasil sebagai berikut

No	Kata Kunci	Precision	Recall
1	Batuk	100 %	55 %
2	Panas Batuk	100 %	100 %
3	Panas Batuk Hipertensi	100 %	100 %
4	Sakit Kepala	100 %	100 %
5	Sakit Kuning	100 %	100 %

Jadi dari hasil tersebut dapat disimpulkan bahwa metode Cosine Similarity cocok digunakan terhadap gejala penyakit.

4.4 Hubungan Penelitian dengan Kaidah Islam

Kesehatan merupakan kebutuhan yang sangat penting. Menurut MUI dalam Musyawarah Nasional Ulama tahun 1983 merumuskan kesehatan sebagai “Ketahanan jasmaniah, rohaniah, dan sosial yang dimiliki manusia, sebagai karunia Allah yang wajib disyukuri dengan mengamalkan (tuntunanNya) dan memelihara serta mengembangkannya”. Dari pengertian tersebut bisa disimpulkan bahwa penyakit dapat dibagi menjadi dua yaitu penyakit hati yang akan berpengaruh terhadap kehidupan rohani dan sosial seseorang dan penyakit jasmani yang berwujud fisik.

Dalam kaitannya terhadap penyakit hati, Allah berfirman dalam surat Al Isra ayat 82 :

حَسَارًا إِلَّا الظَّالِمِينَ يَزِيدُوا لِلْمُؤْمِنِينَ وَرَحْمَةً شِفَاءً هُوَ مَا الْقُرْءَانِ مِنْ وَنَزَّلَ

Artinya : “ dan Kami turunkan dari Al Quran suatu yang menjadi penawar dan rahmat bagi orang-orang yang beriman dan Al Quran itu tidaklah menambah kepada orang-orang yang zalim selain kerugian.”

Dari ayat diatas dapat disimpulkan penyembuh dari penyakit hati adalah Al Quran. Sedangkan dalam kaitannya dengan penyakit fisik, rosululloh menjelaskan melalui beberapa hadist yang diriwayatkan ke beberapa sahabatnya.

Dari riwayat Imam Muslim dari Jabir bin Abdillah dia berkata bahwa Nabi bersabda,

Artinya : “ *Setiap penyakit pasti memiliki obat. Bila sebuah obat sesuai dengan penyakitnya maka dia akan sembuh dengan seizin Allah Subhanahu wa Ta'ala.*” (HR. Muslim

Diriwayatkan pula dari musnad Imam Ahmad dari shahabat Usamah bin Suraik , bahwasanya Nabi bersabda,

عَلَيْهِ وَسَلَّمَ، وَجَاءَتِ الْأَعْرَابُ، فَقَالَ: يَا رَسُولَ اللَّهِ، أَنْتَدَاوِي؟ فَقَالَ: نَعَمْ يَا
عِبَادَ اللَّهِ، تَدَاوُوا، فَإِنَّ اللَّهَ عَزَّ وَجَلَّ لَمْ يَضَعْ دَاءً إِلَّا وَضَعَ لَهُ شِفَاءً غَيْرَ دَاءٍ وَاحِدٍ. : مَا هُوَ؟ قَالَ: لِهَرَمُ

Artinya : “ *Aku pernah berada di samping Rasulullah b. Lalu datanglah serombongan Arab dusun. Mereka bertanya, “Wahai Rasulullah, bolehkah kami berobat?” Beliau menjawab: “Iya, wahai para hamba Allah, berobatlah. Sebab Allah I tidaklah meletakkan sebuah penyakit melainkan meletakkan pula obatnya, kecuali satu penyakit.” Mereka bertanya: “Penyakit apa itu?” Beliau menjawab: “Penyakit tua.”* (HR. Ahmad, Al-Bukhari dalam Al-Adabul Mufrad, Abu Dawud, Ibnu Majah, dan At-Tirmidzi, beliau berkata bahwa hadits ini hasan shahih. Syaikhuna Muqbil bin Hadi Al-Wadi’i menshahihkan hadits ini dalam kitabnya Al-Jami’ Ash-Shahih mimma Laisa fish Shahihain, 4/486)

Hadits-hadits di atas memberikan pengertian kepada kita bahwa semua penyakit yang menimpa manusia maka Allah turunkan obatnya. Dengan adanya aplikasi ini diharapkan dapat memberikan pilihan solusi obat untuk menyembuhkan penyakit.



BAB V

KESIMPULAN

5.1 Kesimpulan

Kesimpulan yang diperoleh dari hasil pengerjaan tugas akhir tentang Implementasi Text Mining untuk mendukung Pencarian Obat Tradisional berdasarkan Gejala yaitu :

Metode yang digunakan sudah bisa diimplementasikan untuk pencarian obat menggunakan kata kunci berupa gejala penyakit. Hal ini berdasarkan hasil uji coba yang dilakukan penulis.

5.2 Saran

Tentu masih banyak kekurangan dari sistem yang dibangun menggunakan text mining ini. Penulis berharap kekurangan-kekurangan tersebut dapat dijadikan pertimbangan untuk pembangunan sistem yang lebih sempurna. Oleh sebab itu beberapa saran pengembangan lebih lanjut yang diberikan penulis adalah penggunaan database data obat yang lebih besar.

DAFTAR PUSTAKA

- Aisah, Siti. 2012. Text Mining dengan Algoritma Vector Space Model untuk Pencarian pada Ensiklopedia Hadits Berbasis Web. UIN Malang
- Aizawa, Akiko. 2002. An Information-theoretic perspective of tf-idf measures. (<http://comminfo.rutgers.edu/~muresan/IR/Docs/Articles/ipmAizawa2003.pdf>) diakses 17 Juli 2014
- Anas, Muhamad F. 2011. Implementasi Text Mining untuk Mendukung Pencarian Topik pada E-library Menggunakan Mobile Device. UIN Malang
- Cios, Krzysztof J Etc. 2007. *Data Mining A Knowledge Discovery Approach*. Springer Science+Business Media,LLC : Spring Street, New York
- Fadli, Ari. Konsep Data Mining. (http://ilmukomputer.org/wp-content/uploads/2011/03/Ari_Fadli_Konsep_Data_Mining.pdf diakses) diakses 17 Juli 2014
- Fayad, Usama Etc. 1996. From Data Mining to Knowledge Discovery in Databases. (<http://www.kdnuggets.com/gpspubs/aimag-kdd-overview-1996-Fayyad.pdf>) diakses 17 Juli 2014
- Hakim, Lukmanul. 2011. TRIK DASYAT MENGUASAI AJAXDENGAN JQUERY. Lokomedia : Yogyakarta
- Han, Jiawei Etc. 2011. Data Mining : concept and techniques. Imprint of Elsivier : Waltham
- Hariana, Arief. 2010. 812 Resep untuk mengobati 236 Penyakit. Penebar Swadaya : Depok
- <http://obtra29.wordpress.com> diakses 7 Juli 2014
- <http://www.tipssehatku.com/2013/09/efek-samping-obat-tradisional.html> diakses 7 Juli 2014
- Kadir, Abdul. 2009. Mastering Ajax dan PHP. Andi : Yogyakarta
- Katno. 2008. Tingkat Manfaat Keamanan dan Efektifitas Tanaman Obat dan Obat Tradisional. B2P2TO-OT : Karanganyar
- Robertson, Stephen. Understanding Inverse Document Frequency : On Theoretical argument of IDF.

(http://www.soi.city.ac.uk/~ser/idfpapers/Robertson_idf_JDoc.pdf)
diakses 17 Juli 2014

Tata, Sandep Etc. 2007. Estimating the Selectivity of tf-idf based Cosine Similarity Predicates. (<http://www.sigmod.org/publications/sigmod-record/0712/p75.errata-tata.pdf>) diakses 17 Juli 2014

Untung, Onny. 2010. Herbal Indonesia Berkhasiat Bukti Ilmiah & Racik. Trubus : Depok

Wiswakarma, Komang. 9 Langkah menjadi master framework codeigniter. Lokomedia : Yogyakarta

Witten, Ian H. Text Mining. (<http://www.cs.waikato.ac.nz/~ihw/papers/04-IHW-Textmining.pdf>) diakses 17 Juli 2014

