

BAB III

ANALISIS DAN PERANCANGAN SISTEM

3.1 Analisis Sistem

Penelitian ini menjadikan data sinopsis buku sebagai obyek penelitian. Sinopsis buku bisa dijadikan sebagai bahan rujukan untuk melakukan penelitian baru. Sedangkan dokumen sinopsis buku merupakan atribut dari sebuah buku yang memberikan gambaran keseluruhan mengenai isi buku secara garis besar. Selama ini sistem pencarian buku di perpustakaan Universitas Muhammadiyah Gresik memang sudah modern dengan memanfaatkan perkembangan teknologi informasi dengan aplikasi digital library, kelemahan digital librarynya kurang sesuai dengan apa yang di inputkan oleh user. Meskipun tidak begitu luas dan lengkap perpustakaan Universitas Muhammadiyah Gresik masih banyak diminati oleh mahasiswa dan dosen untuk berkunjung kesana sekedar hanya membaca, meminjam ataupun mencari referensi untuk penelitian. Tata letak dan penataan buku di perpustakaan Universitas Muhammadiyah Gresik dinilai kurang baik dan efisien sehingga pencarian judul buku secara manual dirasa cukup membuang waktu, meskipun berdasarkan keilmuan sudah tersusun bagus tetapi dinilai kurang bagus dikarenakan buku-buku di perpustakaan Universitas Muhammadiyah Gresik di urutkan sesuai dengan abjad judul buku. Oleh karena itu, permasalahan yang akan diteliti adalah pembuatan sebuah sistem pencarian yang akan digunakan untuk menemukan judul buku yang relevan dan mirip dengan query inputan dari user. Penerapan sistem temu kembali informasi yang ditunjang dengan algoritma single linkage hierarchical method akan digunakan dan dicoba pengimplementasiannya pada sistem.

3.1.1 Analisis Kebutuhan Fungsional Sistem

Ada beberapa kebutuhan fungsional sistem yang akan dikembangkan, diantaranya:

1. Sistem dapat mengolah data koleksi dokumen sinopsis buku menjadi inverted index sehingga didapatkan term dari koleksi dokumen dan jumlah kemunculan term pada koleksi dokumen atau jumlah kemunculan di dokumen tertentu.
2. Sistem dapat mengambil informasi dari file inverted index dan mengolahnya untuk perhitungan pembobotan term dalam tiap dokumen.
3. Sistem dapat melakukan perhitungan matematis relevansi sebuah query dengan koleksi dokumen tertentu yang telah diolah menjadi inverted index.
4. Sistem dapat menemukan dokumen yang memiliki relevansi tertinggi dengan query yang nantinya akan direpresentasikan sebagai sebuah cluster pertama sebagai pembanding dengan dokumen lainnya dalam menemukan dokumen-dokumen yang memiliki jarak terdekat dengan dokumen yang memiliki relevansi tertinggi dengan menggunakan Single Linkage Hierarchical Method.
5. Sistem dapat menemukan hasil pencarian dokumen yang relevan dengan query dan dokumen-dokumen mirip yang didapat dari proses clustering sebelumnya. Hasil dari pencarian akan menampilkan judul buku, penulis, penerbit serta narasi sinopsis buku.

3.1.2. Deskripsi Sistem

Sistem yang akan dibangun ini merupakan aplikasi pencarian buku yang didalamnya menerapkan konsep information retrieval dan algoritma single linkage hierarchical method. Tujuan dari sistem ini adalah mengukur keeratan (similarity) dan relevansi untuk menemukan judul buku yang sesuai dengan query dari user.

Sistem ini menerapkan teknik sistem temu kembali informasi (information retrieval), yang tahapan awalnya adalah preprocessing terhadap dokumen. Dilanjutkan proses analysing yaitu menghitung kedekatan relevansi antara dokumen dengan query user. Hingga mendapatkan nilai similarity dokumen dengan query. Dan dilakukan proses pencarian dokumen-dokumen yang mirip menggunakan algoritma single linkage hierarchical method.

Sistem mengolah data berupa narasi sinopsis buku yang kemudian menghasilkan term berupa kata dasar yang disimpan dalam database untuk proses perhitungan frekuensi dan bobot tiap term dengan dokumen. Selanjutnya sistem akan melakukan perhitungan keeratan antara dokumen dengan query inputan user. Hasilnya sistem akan menghitung nilai similarity terdekat. Ditemukan sebuah dokumen yang sekiranya relevan dengan query. Dokumen ini nantinya akan dijadikan output dari sistem.

Sistem melakukan pencarian dokumen yang memiliki jarak terdekat dengan antar dokumen menggunakan algoritma single linkage hierarchical method.

3.2 Perancangan Sistem

Sistem temu kembali informasi pencarian buku berbahasa Indonesia di Universitas Muhammadiyah Gresik ini menggunakan sample yang berjumlah 90 buku dari jurusan teknik informatika yang terdiri dari judul buku, pengarang, penerbit dan sinopsis buku. Pencarian judul buku ini, menjadikan judul dan sinopsis buku sebagai acuan, beberapa tahapan perhitungan awal pada sinopsis dan judul buku seperti Preprocessing, TF-IDF dan perhitungan dengan menggunakan metode Single Linkage Hierarchical.

Perhitungan Preprocessing, ada beberapa tahapan case folding, tokenizing, filtering dan stemming. Pada tahapan filtering dilakukan pencocokan kata dengan stopwords yang telah di simpan dalam database, kata yang ada didalam stopwords bisa dilakukan penambahan atau pengurangan.

Hasil pencarian akan menghitung berapa lama proses setiap pencarian query yang di inputkan pengguna.

3.2.1. Perhitungan Manual

Contoh kasus yang akan digunakan adalah terdapat empat dokumen synopsis buku

Dokumen 1:

Interaksi Manusia & Komputer

Pengarang : Sudarwan, ST, MT & Dony Arius

Penerbit : Informatika

Materi interaksi manusia dan komputer (IMK) merupakan mata kuliah wajib yang diajarkan kepada mahasiswa di perguruan tinggi jurusan teknik informatika, sistem informasi, ilmu komputer. Interaksi manusia dan komputer bukan hanya membahas bagaimana membuat suatu interface, tetapi memiliki ruang lingkup yang jauh lebih luas.

buku ini disusun untuk memberikan pemahaman bagaimana manusia sebagai sumber daya terpenting dalam membangun sistem. Materi yang dibahas meliputi manusia, komputer, interaksi, paradigma dan prinsip penggunaan, proses desain, model kognitif, analisis tugas, desain dan notasi dialog, model sistem, dukungan implementasi, teknik evaluasi, bantuan dan dokumentasi, groupware, teori dan permasalahan pekerjaan bersama yang didukung komputer, serta sistem banyak sensor.

Dokumen 2 :

Interaksi Manusia dan Komputer (Teori dan Praktek)

Pengarang : Ir. P. Insap Sentosa, M.Sc.

Penerbit : Informatika

Bila Anda menggunakan komputer, perlu Anda sadari bahwa Anda sebenarnya sedang berdialog dengan komputer. Ketika Anda memberikan suatu perintah dan komputer mengerjakan perintah tersebut dengan menampilkan hasil di layar tampilan, sesungguhnya Anda telah berdialog dengan komputer. Dari sini kemudian muncul disiplin ilmu baru yang disebut dengan Interaksi Manusia dan Komputer. Aspek penting yang dipelajari dalam disiplin ilmu ini adalah aspek manusia, aspek komputer, dan aspek lingkungan kerja atau ergonomic. Dalam buku ini anda diajak untuk memahami ketiga aspek di atas dengan tujuan untuk mencapai kondisi yang nyaman mungkin ketika bekerja di depan terminal komputer untuk selang waktu yang cukup lama.

Pokok bahasan buku ini:

- Antarmuka antara manusia dan komputer
- Panca indera yang terlibat dalam dialog dengan komputer

- Ragam dialog pada berbagai program aplikasi
- Perancangan tampilan program aplikasi
- Peranti interaktif sebagai piranti masukan dan keluaran yang mempermudah dialog manusia dan komputer
- Aspek ergonomis untuk kenyamanan lingkungan kerja
- BGI, teknik pemrograman untuk antarmuka
- Pengoperasian mouse
- Pembuatan komponen antarmuka grafis
- Sistem pencendeleaan
- Sistem menu

Dokumen 3 :

Interaksi Manusia dan Komputer

Pengarang : Soetam Rizky

Penerbit : Graha Ilmu

Mata kuliah interaksi manusia dan komputer atau IMK merupakan sebuah mata kuliah wajib di berbagai perguruan tinggi rumpun informatika. Sayangnya, tidak banyak literatur berbahasa Indonesia yang tersedia di pasaran. Karenanya, buku ini dibuat untuk memberi kelengkapan sekaligus referensi dalam proses belajar mengajar mata kuliah IMK, khususnya bagi para akademis di Perguruan Tinggi rumpun informatika. Meski demikian, buku ini layak dijadikan referensi bagi para sistem analis ataupun programmer yang seringkali merasa kesulitan dalam melakukan perancangan antar muka sebuah aplikasi.

Dalam buku ini dibahas beragam teori yang dijadikan landasan dalam perancangan antar muka aplikasi, baik aplikasi berbasis desktop ataupun aplikasi berbasis web. Dengan kelengkapan referensi dan analisa yang dirangkum dari berbagai sumber yang

terpercaya, diharapkan buku ini dapat menjadi acuan bagi para praktisi dan akademisi informatika di Indonesia.

Pembahasan materi meliputi:

- Komponen Interaksi Manusia dan Komputer
- Prinsip Desain
- Usability
- Task Analysis
- Teknik Implementasi
- Dokumen dan Evaluasi
- Groupware dan Computer Supported Cooperative Work

Dokumen 4 :

Algoritma dan Pemrograman dalam bahasa Pascal dan C

Pengarang : Rinaldi Munir

Penerbit : Informatika

Algoritma dan Pemrograman dalam Bahasa Pascal dan C merupakan penggabungan dari dua buah buku sebelumnya, yaitu Algoritma dan Pemrograman dalam Bahasa Pascal dan C (Buku 1) dan Algoritma dan pemrograman dalam Bahasa Pascal dan C (Buku 2).

Materi yang dibahas di dalam Buku ini meliputi:

- Konsep dasar algoritma
- Tipe, Operator ,dan Ekspresi
- Struktur dasar pembangun algoritma: runtunan, pemilihan, pengulangan
- Pemrograman modular: Fungsi dan prosedur
- Larik (array)
- Matriks
- Algoritma pencarian (searching)

- Algoritma pengurutan (sorting)
- Arsip beruntun (sequentila file)
- Algoritma rekursif

Setiap materi diperkaya dengan banyak contoh pemecahan masalah. Diharapkan buku ini dapat mengajarkan pembacanya menjadi seorang pemrogram yang dapat memprogram dengan Kaidah yang benar.

Notasi algoritmik yang digunakan adalah notasi abstrak berupa pseudo-code. Notasi pseudo-code dibuat sedemikian sehingga ia mudah diterjemahkan (translasi) ke dalam notasi bahasa pemrograman. Bahasa pemrograman yang digunakan didalam buku ini adalah Bahasa Pascal dan Bahasa C. Kedua bahasa ini dipilih karena populer, banyak digunakan, dan menjadi dasar pengembangan bahasa-bahasa yang lebih bary seperti Object Pascal, Java, C+ +, dan lain-lain. Kompilator (compiler) Bahasa [ascal/C yang digunakan boleh sembarang, namun buku ini menggunakan kompilator yang free dan dapat didownload dan internet, yaitu Free Pascal dan GCC (GNU C Compiler).

Query yang dimasukkan User : Interaksi Manusia dan Komputer

Berikut dibawah ini tahapan-tahapan prosesnya :

Dilakukan tahapan yang pertama yaitu indexing:

- **Proses case-folding dan tokenizing :**

Synopsis buku maupun query inputan user dilakukan penghilangan tanda baca dan pengubahan menjadi huruf kecil semua. Selanjutnya dilakukan proses tokenizing hingga didapatkan query dan abstrak perkata.

- **Proses filtering dan stemming**

Hasil dari proses preprocessing akan dilakukan penghilangan kata stopwords, selanjutnya dilakukan stemming sehingga didapatkan term berupa kata dasar yang nantinya akan dimasukkan dalam database sistem.

Query : interaksi

D1 : materi interaksi manusia komputer teknik sistem

D2 : komputer perintah disiplin ilmu manusia aspek buku dialog program

D3 : mata kuliah interaksi manusia komputer informatika buku aplikasi

D4 : manajemen informasi jadwal teknologi internet desain aplikasi web rancang sistem

Proses pembobotan dan perhitungan similiaritas :

Term-term dokumen serta term query yang didapat dihitung frequency kemunculannya (tf) dalam tiap dokumen serta nilai inverse document frequency (IDF) menggunakan persamaan rumus (2.3).

Dibawah ini tabel 3.1 merupakan tabel perhitungan TF-IDF tiap term dalam masing-masing dokumen

Table 3.1 tabel perhitungan TF-IDF

No	Kata	Q	Tf				Df	idf
			d1	d2	d3	d4		
1	Materi		2			2	1.301029996	
2	Interaksi	1	4	1	3	3	1.124938737	
3	Manusia		3	4	2		1.124938737	
4	Komputer		5	10	2		1.124938737	
5	Teknik		2			1	1.602059991	
6	Sistem		4			1	1.602059991	
7	Perintah			2		1	1.602059991	
8	Disiplin			2		1	1.602059991	

9	Ilmu			2			1	1.602059991
10	Aspek			6			1	1.602059991
11	Buku			2	4	6	3	1.124938737
12	Dialog			5			1	1.602059991
13	Program			2		5	2	1.301029996
14	Mata				3		1	1.602059991
15	Kuliah				4		1	1.602059991
16	Informatika				3		1	1.602059991
17	Aplikasi				4		1	1.602059991
18	Algoritma					8	1	1.602059991
19	Bahasa					10	1	1.602059991
20	Pascal					6	1	1.602059991
21	Notasi					3	1	1.602059991
22	Pseudocode					3	1	1.602059991

Selanjutnya dilakukan perhitungan pembobotan masing-masing term dokumen serta term query dari tiap dokumen menggunakan rumus (2.4).

Table 3.2 Tabel Perhitungan Pembobotan

No	Kata	W			
		d1	d2	d3	d4
1	Materi	2.60206	0	0	2.60206
2	Interaksi	5.204119	1.3010299	3.9030899	0
3	Manusia	3.374816	4.499755	2.249877	0
4	komputer	5.624694	11.24939	2.249877	0
5	Teknik	3.20412	0	0	0
6	Sistem	6.40824	0	0	0
7	Perintah	0	3.20412	0	0
8	Disiplin	0	3.20412	0	0
9	Ilmu	0	3.20412	0	0
10	Aspek	0	9.61236	0	0
11	Buku	0	2.249877	4.499755	6.749632
12	Dialog	0	8.0103	0	0
13	Program	0	2.60206	0	6.50515
14	Mata	0	0	4.80618	0
15	Kuliah	0	0	6.40824	0
16	informatika	0	0	4.80618	0
17	Aplikasi	0	0	6.40824	0
18	algoritma	0	0	0	12.81648
19	Bahasa	0	0	0	16.0206

20	Pascal	0	0	0	9.61236
21	Notasi	0	0	0	4.80618
22	Pseucode	0	0	0	4.80618

Setelah di dapatkan masing masing bobot term dokumen dan query. Selanjutnya dilakukan perhitungan jarak dengan menggunakan metode Single Linkage Hierarchical Method.

$$d(x,y) = \sqrt{\sum_{i=1}^n |Xi - yi|^2}$$

1. Jarak antara dokumen 1 dengan dokumen 2 :

$$d(x,y) = \sqrt{|1.3010299 - 5.204119|^2} = 3.90309$$

2. Jarak antara dokumen 1 dengan dokumen 3 :

$$d(x,y) = \sqrt{|3.903089 - 5.204119|^2} = 1.30103$$

3. Jarak antara dokumen 2 dengan dokumen 3 :

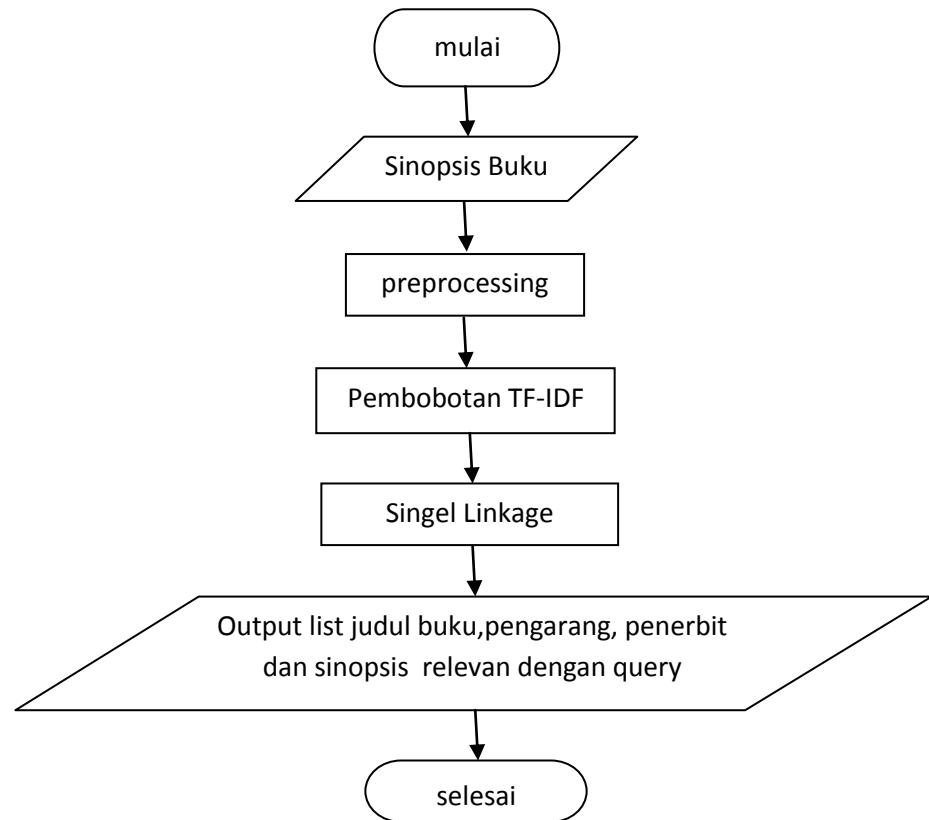
$$d(x,y) = \sqrt{|3.303089 - 1.3010299|^2} = 2.60206$$

Data	1	2	3
1	0	3.90309	1.30103
2	3.90309	0	2.60206
3	1.30103	2.60206	0

3.2.2. Flowchart Sistem (Diagram Alir Sistem)

Gambar 3.1 dibawah merupakan gambaran flowchart utama sistem pencarian buku.

1. Pertama sistem mengambil dan membaca sinopsis buku yang tersimpan di dalam tabel_buku.
2. Sistem melakukan preprocessing terhadap sinopsis, yang meliputi case-folding, tokenizing, filtering dan stemming, sehingga didapatkan kata (*terms*) dalam bentuk akar yang nantinya dilakukan pengindeksan.
3. Sistem melakukan pembobotan dari tiap terms dalam tiap dokumen berdasarkan rumus (2.4). hasil pembobotan tiap term dokumen disimpan dalam database.
4. User menginputkan query, kemudian sistem melakukan proses preprocessing hingga didapatkan *term query*, dan dilakukan pembobotan terhadap *terms query*. Hasilnya disimpan dalam database.
5. Sistem melakukan proses Single Linkage Hierarchical Method untuk mendapatkan cluster dokumen synopsis yang memiliki tingka
6. t relevansi tertinggi dengan query yang di inputkan menggunakan persamaan rumus (2.5).
7. Sistem mengeluarkan output berupa cluster dokumen synopsis (list judul buku, penulis, penerbit dan synopsis) yang memiliki tingkat relevansi tertinggi dengan inputan query.

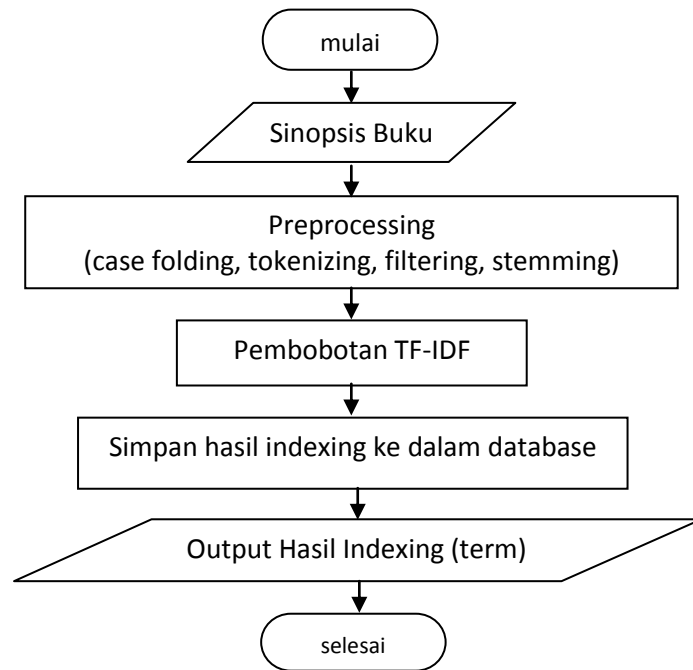


Gambar 3.1 Flowchart utama sistem

Sistem pencarian buku dibagi menjadi tiga tahapan, tahapan yang pertama yaitu indexing yang berhubungan dengan sinopsis buku. Dalam proses indexing, dilakukan beberapa tahapan :

- Proses pertama dalam sistem dimulai dengan membaca isi sinopsis buku yang disimpan dalam database sebagai *corpus*
- sistem melakukan preprocessing terhadap *corpus* tersebut, yaitu case folding, tokenizing, filtering dan stemming. Hasil proses ini berupa term sinopsis buku.
- Hasil proses tersebut kemudian dilakukan proses pembobotan tiap term. Bobot tiap term sinopsis buku disimpan dalam tabel database.

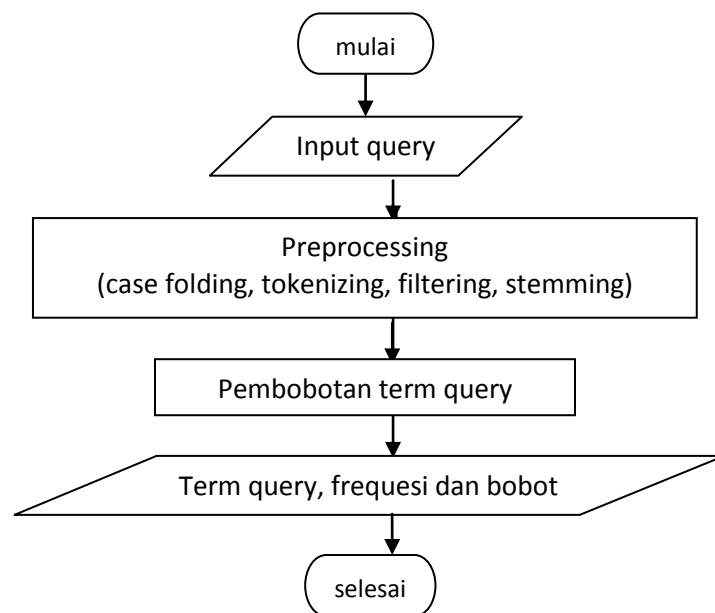
Proses indexing digambarkan dalam flowchart, lihat gambar 3.2.



Gambar 3.2 Flowchart proses indexing sinopsis buku

Tahapan selanjutnya adalah formulasi query. Gambar 3.3 merupakan gambaran flowchart tahapan formulasi query. Dalam tahapan ini dilakukan tahapan proses:

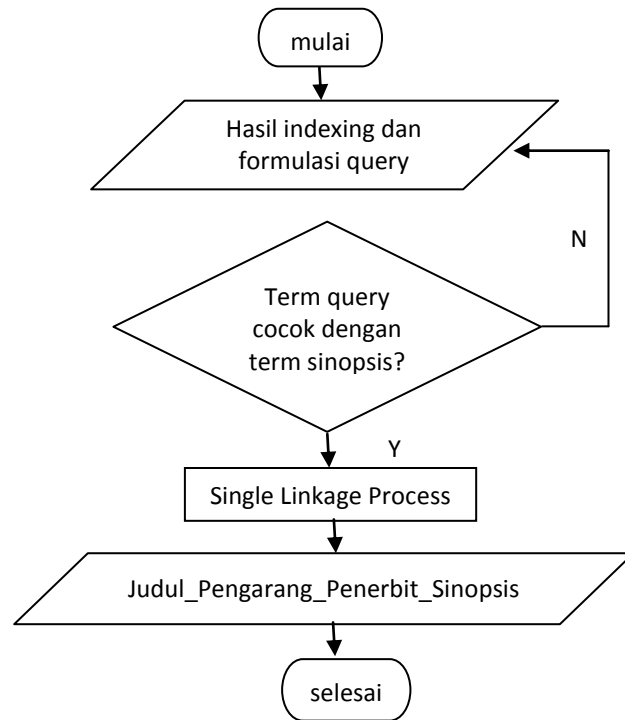
- User menginputkan query berupa judul buku.
- system melakukan preprocessing terhadap query tersebut, yaitu case folding, tokenizing, filtering dan stemming. Hasil proses ini berupa daftar term query.
- Hasil proses tersebut kemudian dilakukan proses pembobotan tiap term query. Daftar term query dan bobotnya disimpan dalam tabel database.



Gambar 3.3 Flowchart formulasi query

Tahapan yang terakhir, tahapan pencarian. Gambar 3.4 merupakan gambaran flowchart tahapan formulasi query. Dalam tahapan ini terdapat proses:

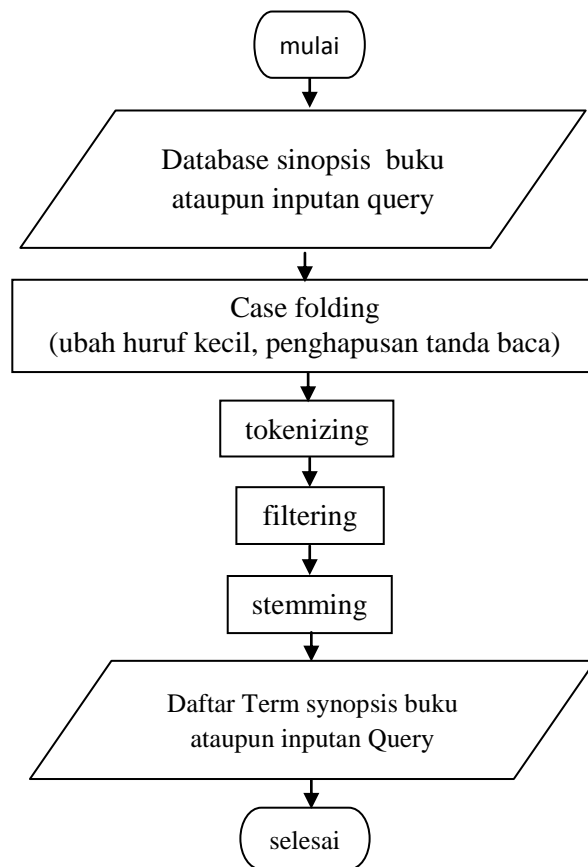
- Sistem mengambil database hasil indexing dan formulasi query, kemudian melakukan pencocokan term query dengan term sinopsis buku.
- Jika term query cocok dengan term sinopsis maka sistem melakukan perhitungan similarity antar query dengan sinopsis buku.
- Sistem melakukan pencarian sinopsis buku yang memiliki jarak terdekat antar synopsis buku dengan menggunakan single linkage hierarchical method.
- Sistem mengeluarkan output berupa list judul buku, pengarang, penerbit dan synopsis buku.



Gambar 3.4 Flowchart pencarian

Berdasarkan diagram alir sistem diatas terdapat beberapa sub proses yang dilakukan :

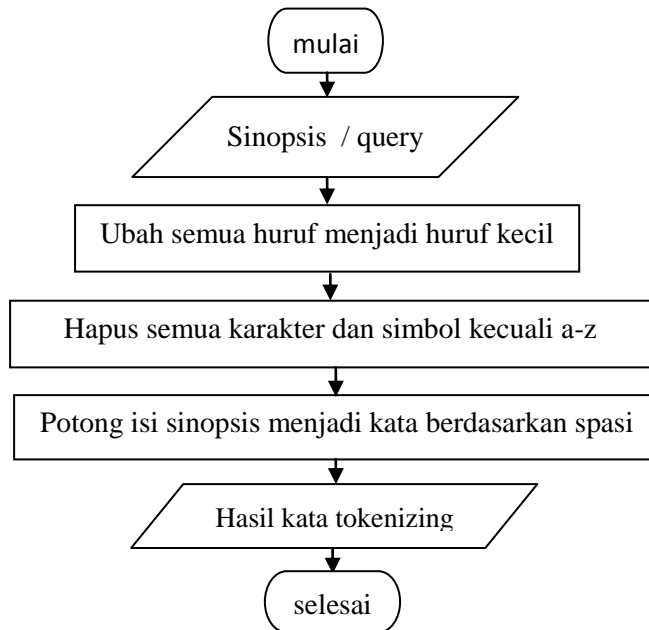
- **Proses pre-processing** merupakan proses mengolah isi sinopsis maupun query inputan user menjadi term dan keyword. Gambar 3.5 adalah gambaran diagram alir dari proses preprocessing:



Gambar 3.5 diagram alir proses preprocessing

Tahapan dalam Proses preprocessing :

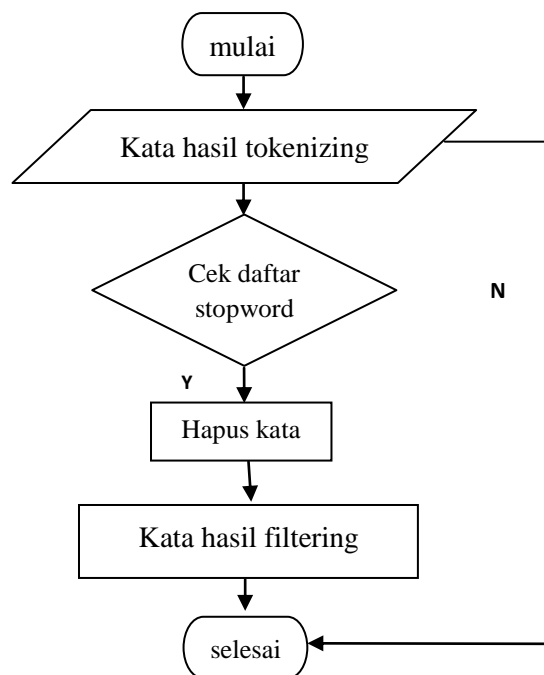
- Proses case folding, yaitu proses penghilangan karakter atau tanda baca dan juga pengubahan huruf menjadi huruf kecil dalam dokumen sinopsis maupun query.
- Proses tokenizing, merupakan proses pemecahan kalimat dokumen sinopsis menjadi kata. Begitu juga dengan query inputan user. Hasil dari proses ini akan dilakukan filtering. Proses ini digambarkan dalam flowchart gambar 3.6.



Gambar 3.6 diagram alir case folding dan tokenizing

Berdasarkan gambar 3.6 :

- Sistem memulai membaca isi synopsis / query dalam database.
 - Sistem mengubah semua hurufnya menjadi huruf kecil
 - Sistem menghapus semua karakter dan simbol kecuali huruf a-z
 - Sistem memotong isi sinopsis menjadi kata berdasarkan spasi.
 - Sistem mengambil hasil dari proses berupa kata untuk dilakukan proses filtering.
- Proses filtering, pemilahan kata dan pengeliminasian kata yang merupakan stopwords. Dalam tahap ini dilakukan pencocokan kata dengan kata stopwords yang telah disimpan dalam sebuah tabel database. Apabila ditemukan kata yang merupakan stopwords maka kata tersebut akan dieliminasi. Berikut diagram alir dari proses filtering :



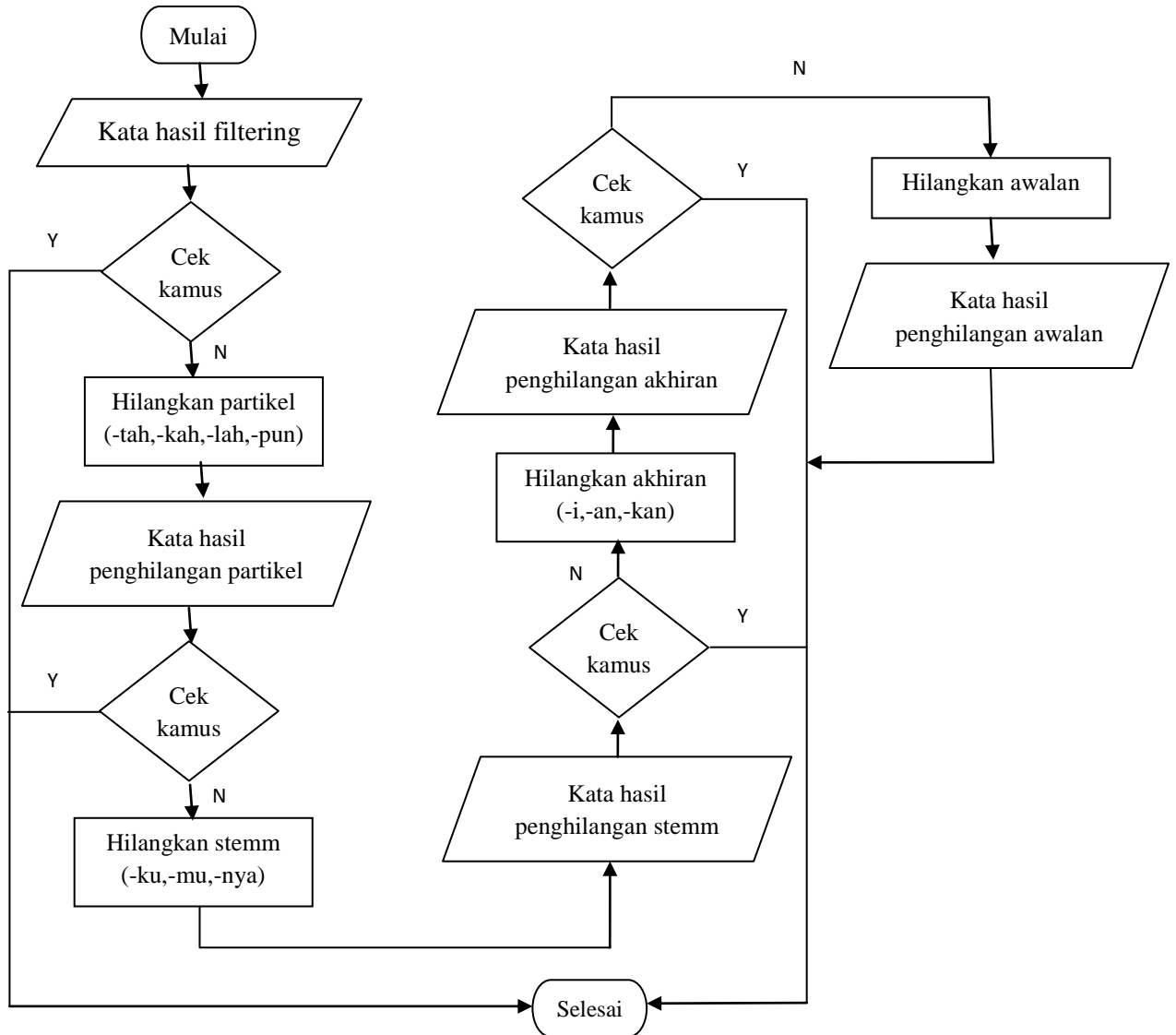
Gambar 3.7 diagram alir dari proses filtering

Keterangan:

- sistem mengambil kata dari hasil hasil tokenizing.
 - Sistem mengecek apakah kata tersebut ada yang sama dengan daftar stopwords dalam database, jika iya maka hapus kata tersebut. Namun jika tidak simpan kata tersebut untuk proses stemming.
 - Sistem melakukan perulangan dalam mengecek seluruh kata hasil tokenizing sampai tidak ada kata lagi.
- Proses selanjutnya adalah stemming, mengubah kata menjadi bentuk kata dasar atau yang disebut *term*. Term inilah nanti disimpan dalam database untuk dilakukan perhitungan pembobotan.

Dalam sistem ini digunakan algoritma stemming Nazief-Andriani. Berikut gambaran diagram alir algoritma stemming Nazief-Andriani. Dalam hal ini diperlukan sebuah database yang menampung daftar kamus kata yang kelengkapannya akan mempengaruhi ketepatan perubahan kata

ke bentuk akarnya. Berdasarkan algoritma stemming nazief-adriani pada bab 2, Gambar 3.8 diagram alir dari algoritma stemming Nazief-Andriani.



Gambar 3.8 diagram alir dari algoritma stemming Nazief-Adriani

Proses tahapan stemming nazief dan adriani yang terjadi pada gambar 3.8 :

1. Kata dicari didalam daftar kamus. Bila kata tersebut ditemukan di dalam kamus maka dapat diasumsikan kata tersebut adalah kata dasar sehingga algoritma dihentikan

2. Bila kata didalam langkah pertama tidak ditemukan didalam kamus, maka diperiksa apakah sufiks tersebut yaitu sebuah partikel (“-tah”, “-lah”, “-pun” atau “-kah”). Bila ditemukan maka partikel tersebut di hilangkan.
3. Pemeriksaan dilanjutkan pada kata ganti milik (“-ku”, “-mu” “-nya”). Bila ditemukan maka kata ganti tersebut dihilangkan.
4. Memeriksa akhiran (“-i”, “-an”, “-kan”). Bila di temukan maka akhiran tersebut dihilangkan.

Hingga langkah ke-4 dibutuhkan ketelitian untuk memeriksa apakah akhiran “-an” merupakan hanya bagian dari akhiran “-kan” dan memeriksa lagi apakah partikel (“-tah”, “-lah”, “-pun” atau “-kah”) dan kata ganti milik (“-ku”, “-mu”, “-nya”) yang telah dihilangkan pada langkah 2 dan 3 bukan merupakan bagian dari kata dasar.

5. Memeriksa awalah (“se-“, “ke-“, ”di-“, “te-“, “be-“, “pe-“, “me-“). Bila ditemukan maka awalantersebut dihilangkan. Pemeriksaan dilakukan dengan berulang mengingat adanya kemungkinan multiprefik. Langkah ke-5 ini juga membutuhkan ketelitian untuk memeriksa kemungkinan peluluhan awalan, perubahan prefix yang disesuaikan dengan huruf awal kata dan aturan kombinasi prefix-sufix yang diperbolehkan.

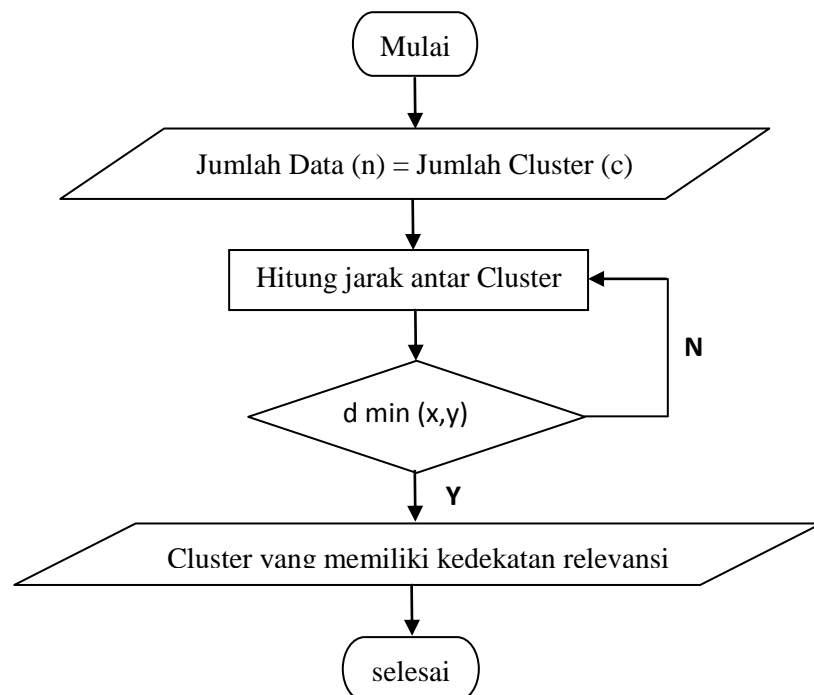
- **Proses pembobotan tf-idf (Term Frequency-Inversed Document Frequency)**

Algoritma pembobotan TF-IDF merupakan algoritma pembobotan yang merupakan perpaduan antara pembobotan local, yaitu term frequency (tf) dengan pembobotan global yaitu inverence document frequency (idf). Pembobotan TF dilakukan dengan menghitung jumlah kemunculan kata dalam synopsis buku, sedangkan pembobotan TF-IDF merupakan nilai *log* dari jumlah synopsis buku yang mengandung satu kata, dan nilai dari pembobotan TF-IDF diperoleh dengan mengalihkan nilai TF dengan IDF yang diperoleh. Dari bobot TF-IDF tersebut dilakukan shorting, bobot yang lebih tinggi memiliki tingkat kemiripan synopsis buku perlu dilakukan tahap selanjutnya,

yaitu representasi Single Linkage Hierarchical Method. Pembobotan menggunakan rumus $tf * idf$, menggunakan rumus (2.4)

- Proses pencarian dokumen mirip menggunakan algoritma Single Linkage Hierarchical Method

Cara kerja algoritma Singel Linkage Hierarchical Method :



Gambar 3.9 diagram alir proses algoritma single linkage hierarchical method

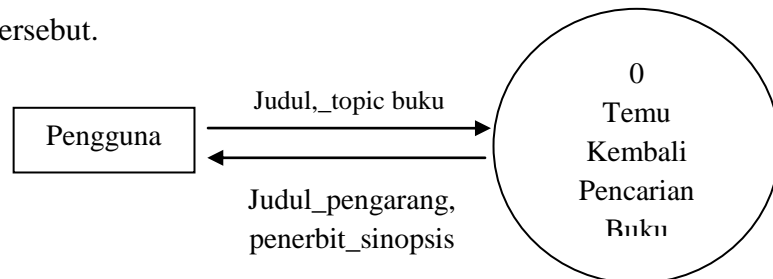
- Representasi model single linkage :
 1. Diasumsikan setiap data dianggap sebagai cluster. Jika n =jumlah data dan c =jumlah cluster, berarti ada $c=n$.
 2. Menghitung jarak antar cluster dengan menggunakan Euclidian distance

$$d(x,y) = \sqrt{\sum_{i=1}^n |X_i - y_i|^2}$$

3. Mencari 2 cluster yang mempunyai jarak antar cluster yang paling minimal dan digabungkan (merge) kedalam cluster baru (sehingga $c=c-1$).
4. Kembali ke langkah 3, dan diulangi sampai dicapai cluster yang diinginkan.

3.2.3. Context Diagram Sistem

Pada context diagram gambar 3.11 ini merupakan gambaran sistem atau aplikasi secara garis besar yang disebut sebagai top level. Dimana user memasukkan query atau keyword (berupa judul buku atau topic buku) ke dalam sistem pencarian, query inilah yang akan diproses dan kemudian akan mendapatkan hasil berupa judul buku yang relevan, didalamnya terdapat informasi berupa sinopsis, nama pengarang dan penerbit yang relevan dan mirip dengan query inputannya sebagai hasil dari proses pencarian tersebut.



Gambar 3.10 context diagram sistem pencarian judul buku

3.2.4. Diagram Berjenjang

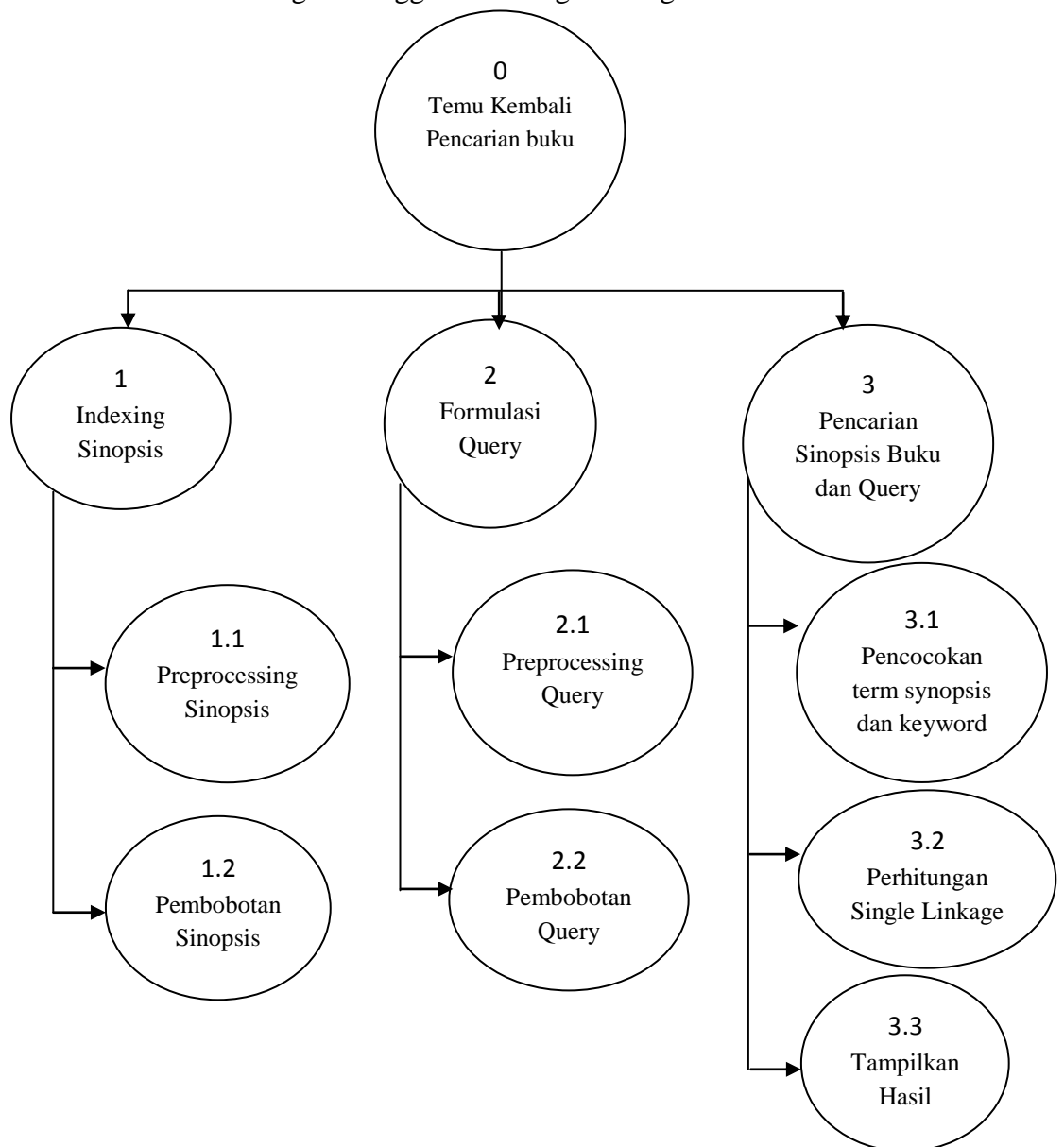
Gambar 3.12 dibawah ini merupakan gambar diagram berjenjang dari sistem pencarian judul buku yang terdiri dari 3 level, yaitu :

- Top level : Sistem Pencarian Judul buku secara global
- Level 0 : merupakan hasil break down dari proses global dari sistem yang terdiri dari beberapa sub proses, yaitu:
 - o Proses indexing dokumen synopsis buku
 - Proses indexing terdapat sub proses didalamnya, diantaranya proses preprocessing dokumen dan pembobotan
 - o Proses formulasi query / keyword yang diinput user

Proses ini juga mengandung beberapa sub proses, yaitu proses preprocessing query dan pembobotan

- o Proses pencarian

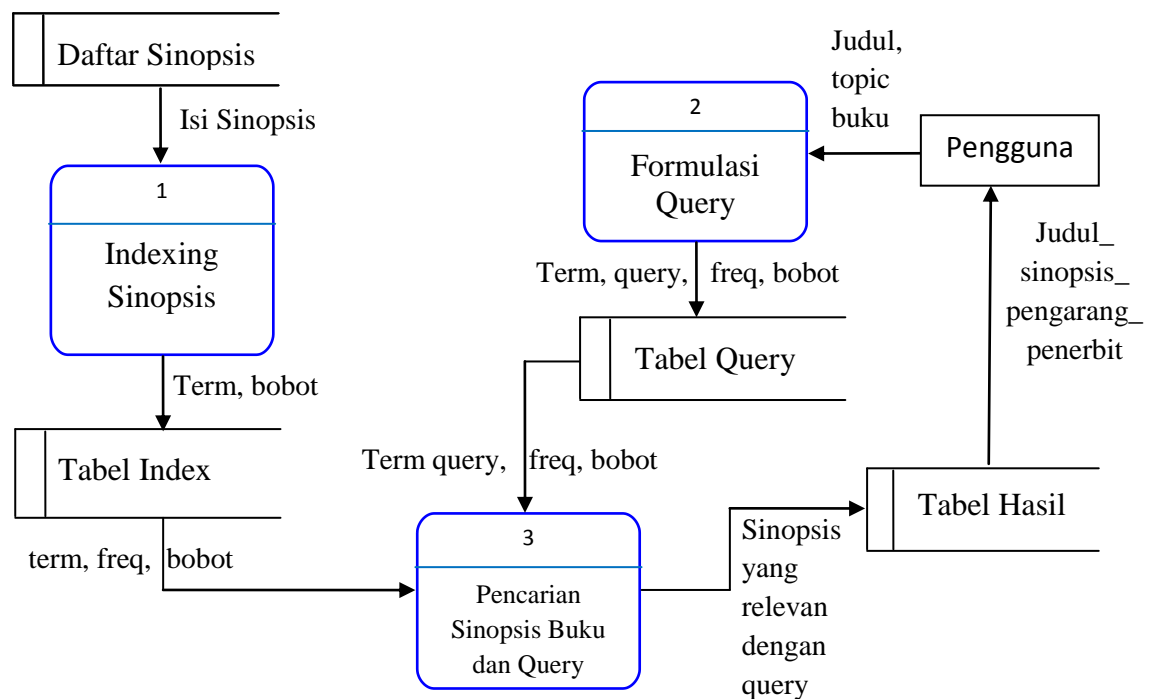
Dalam proses pencarian terdapat sub proses diantaranya yaitu pencarian term synopsis buku dan query, pencocokan term synopsis dengan keyword dan proses pencarian dokumen yang memiliki kedekatan dengan menggunakan single linkage.



Gambar 3.11 Diagram berjenjang sistem pencarian sinopsis buku

3.2.5. Data Flow Diagram

Dari diagram berjenjang akan dijabarkan lebih dalam pada DFD sistem pencarian buku. DFD (Data Flow Diagram) merupakan diagram yang menunjukkan dan memvisualisasikan secara grafis hubungan antar subproses-subproses dalam suatu sistem. Berikut gambaran data flow diagram aplikasi pencarian yang menunjukkan tahapan proses, pertama : indexing dokumen isi abstrak, kedua: pemrosesan query inputan user, ketiga: proses pencarian. Dari proses tersebut user mendapatkan informasi buku yang sesuai dengan inputan user berupa judul, pengarang, penerbit serta synopsis buku.



Gambar 3.12 Data flow diagram level 1

Berdasarkan gambar 3.12

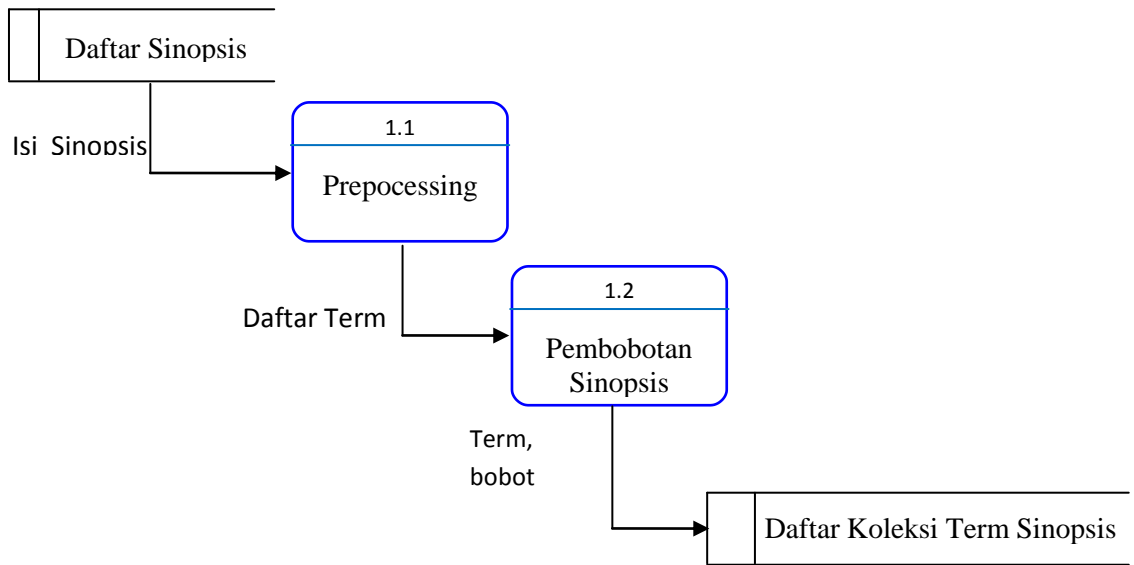
- Sistem mengambil isi synopsis dalam data base.
- Sistem melakukan proses indexing isi synopsis yang menghasilkan term isi synopsis.
- Hasil indexing tersebut disimpan dalam database berupa term dan bobotnya untuk proses pencarian

- User menginputkan query berupa judul buku yang kemudian dilakukan formulasi query, menghasilkan term query dan bobot yang disimpan dalam database.
- Sistem melakukan pencarian keceratan antara query dengan isi sinopsis yang nantinya memberikan output berupa judul buku yang relevan dan mirip dengan query inputan user

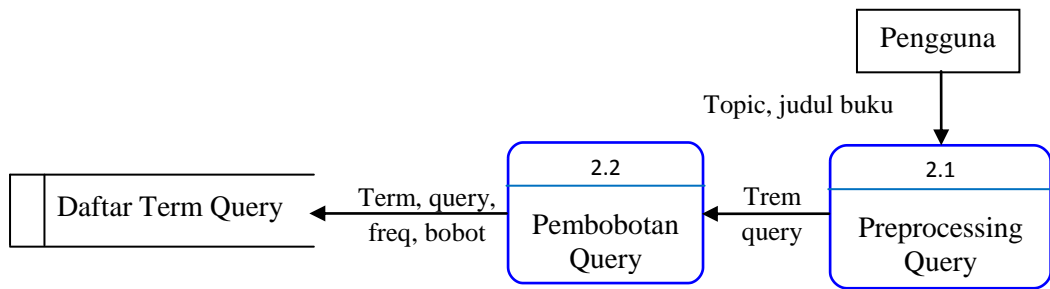
Dari tiap-tiap tahapan proses dari gambar 3.13 dapat dijabarkan lebih detail dalam sinopsis yang mana dalam proses ini dilakukan beberapa urutan proses hingga menghasilkan output berupa kumpulan term dan bobot dari tiap term terhadap narasi sinopsis.

Pada DFD level 2 proses formulasi query terdapat dua tahap proses yaitu proses preprocessing query sehingga didapatkan daftar term query, lalu dilakukan proses kedua yaitu pembobotan term query, dimana hasilnya akan disimpan dalam sebuah tabel di database untuk proses pencarian. Lihat gambar 3.14

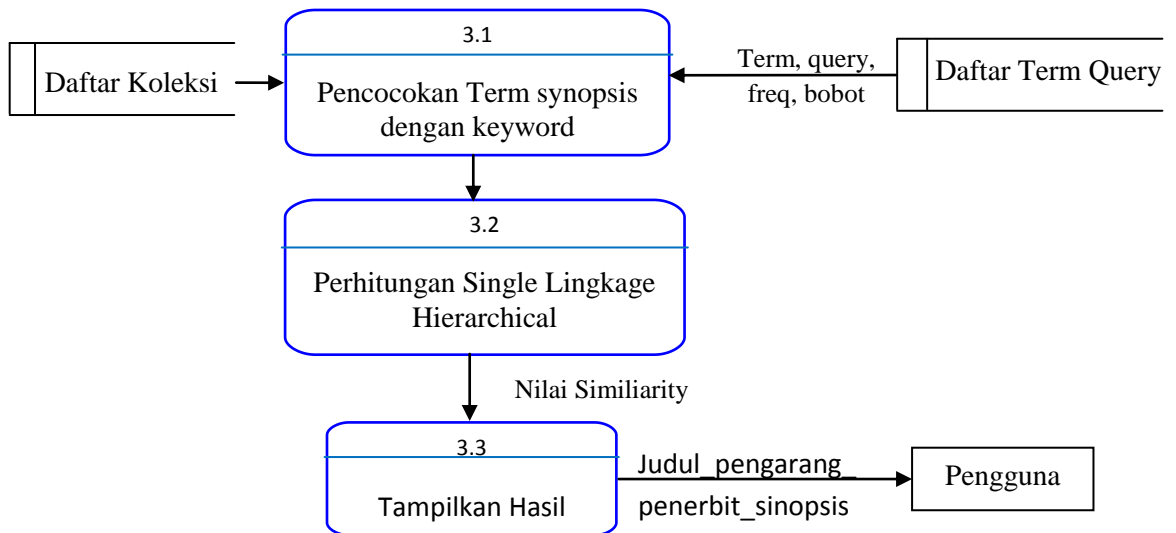
Tahapan yang ketiga sebagaimana gambar 3.15 merupakan tahapan pencarian dokumen yang sekiranya relevan dengan query dan pencarian dokumen-dokumen yang mirip. Daftar term query hasil proses query dilakukan pencocokan dengan daftar tem hasil indexing dan akan dihitung nilai simmilarity-annya terhadap koleksi term yang didapat dari tahapan indexing. Kemudian dihitung kedekatan jarak antar dokumen dengan menggunakan algoritma single linkage hierarchical method. Hasil dari tahapan pencarian ini nantinya berupa list buku yang relevan berupa informasi judul, pengarang, penerbit serta narasi synopsis buku.



3.13 DFD level 2 Proses indexing



Gambar 3.14 DFD level 2 formulasi query

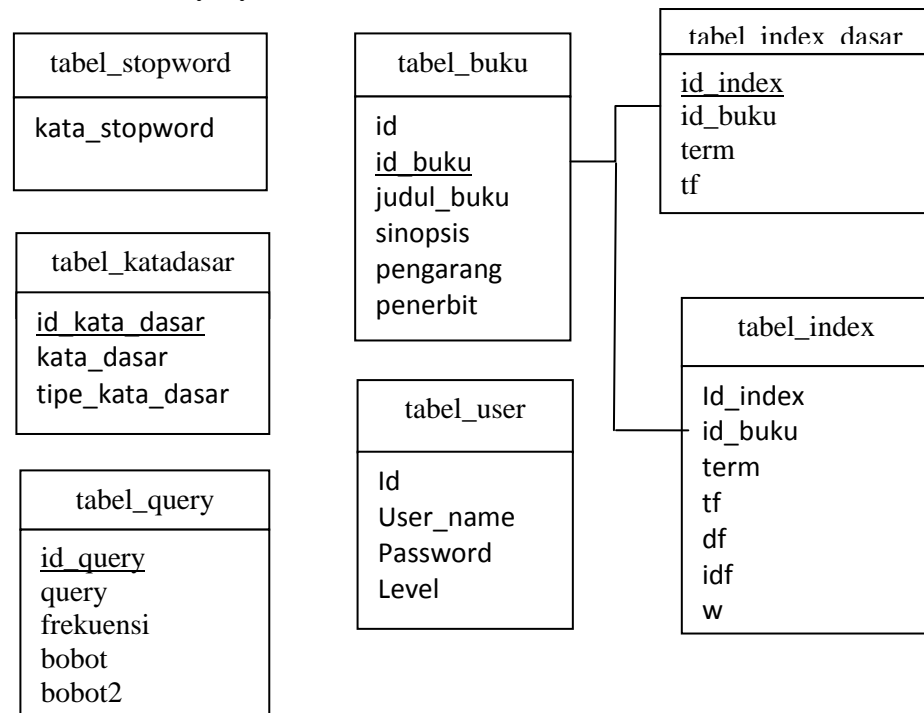


Gambar 3.15 DFD level 2 proses pencarian

Pada DFD level 2 tahapan proses preprocessing terdapat beberapa proses seperti yang dijelaskan sebelumnya. Yaitu proses case folding yang dilanjutkan tokenizing hingga mendapatkan kumpulan potongan kata dari isi synopsis buku. Hasil tokenizing tersebut dilakukan filtering penghilangan kata stopwords, dimana nantinya akan didapatkan kata-kata yang mampu merepresentasikan isi dari synopsis buku. Selanjutnya dilakukan stemming menggunakan stemming nazief-andriani sehingga mendapatkan kata dasar (*term*) yang disimpan dalam database untuk proses pembobotan. Proses yang sama juga dalam preprocessing query sehingga didapatkan term query yang disimpan dalam database.

3.2.6. Perancangan Data Base

Berikut ini ERD desain data base yang merepresentasikan relasi antara entiny-nya.



Gambar 3.16 Entity Relational Diagram Sistem Pencarian Judul Buku

Penjelasan Entity Relational Diagram Sistem :

- Tabel_stopword
Tabel ini berisikan penyimpanan kata dengan kata-kata dengan frekuensi tinggi atau kata yang irrelevant dengan dokumen.
- Tabel_katadasar
Tabel katadasar digunakan untuk menyimpan kumpulan kata dasar berdasarkan kamus bahasa Indonesia.
- Tabel_query
Untuk penyimpanan sementara query inputan user yang telah dilakukan preprocessing dan bobot dari tiap term.
- Tabel_buku
Untuk penyimpanan data buku (judul, pengarang, penerbit dan synopsis buku)
- Tabel_user
Tabel berikut digunakan untuk menyimpan data admin dan pengunjung perpustakaan.
- Tabel_index
Berfungsi untuk penyimpanan daftar term abstrak hasil proses preprocessing.
- Tabel_hasil_stemming
Digunakan untuk menyimpan hasil perhitungan stemming pada synopsis buku.

Berikut table-table yang digunakan dalam pembangunan sistem pencarian judul buku :

- a. Tabel tabel_buku
Tabel tabel_buku ini digunakan untuk penyimpanan daftar dari atribut synopsis buku, yaitu judul buku, nama pengarang dan penerbit.

Tabel 3.3 struktur tabel_buku

Field	Type	Size	Keterangan
Id	Int	10	Incremental
Id_buku	Varchar	200	
Judul_buku	Varchar	300	
Synopsis	Text		
Pengarang	Varchar	100	
Penerbit	Varchar	100	

b. Table stopword (stoplist)

Table stopword ini digunakan untuk penyimpanan kata kata dengan frekuensi tinggi atau kata yang irrelevant dengan dokumen. Terdapat satu field yang dalam table ini, yaitu kolom kata_stopword yang digunakan untuk penyimpanan kata yang akan dihilangkan pada proses filtering. Sehingga dapat dilakukan pengecekan terhadap tiap kata dari input dokumen, jika kata tersebut tidak diproses ke tahap selanjutnya. Struktur tabel ini dapat dilihat pada table 3.4

Table 3.4 struktur table_stopwrod

Field	Type	Size	Keterangan
Kata_stopword	Varchar	100	-

c. Table kata dasar

Tabel kata dasar berisi kumpulan kata dasar berdasarkan kamus bahasa Indonesia. Dalam tabel ini juga disertai keterangan kata dasar tersebut, sehingga diketahui kata yang berupa kata kerja atau kata yang berupa kata sifat. Tabel ini digunakan pada proses stemming untuk mengecek apakah kata tersebut sudah merupakan kata dasae. Jika tidak terdapat dalam daftar kata dasar maka akan dilakukan proses stem terhadap kata tersebut. struktur tabel ini dapat dilihat pada tabel 3.5

tabel 3.5 stuktur tabel_kata_dasar

Field	Type	Size	Keterangan
Id_kata_dasar	Int	11	Incremental
Kata_dasar	Varchar	20	Kata dalam bentuk akar
Tipe_kata_dasar	Varchar	20	Tipe dari kata dasar tersebut

d. Tabel index

Tabel tabel_index berfungsi untuk penyimpanan daftar term abstrak hasil proses preprocessing, nilai TF (term frequency) masing-masing term dari tiap document, nilai IDF (inverse document frequency) dari perhitungan mrnggunakan rumus (2.3) serta nilai bobot tiap term yang didapat dari perhitungan perkalian tf dan idf rumus (2.4). Tabel ini memiliki struktur sebagai berikut :

Tabel 3.6 sturktur tabel_index

Field	Type	Size	Keterangan
Id_index	Int	11	Incremental
Id_buku	Varchar	200	
Term	Varchar	50	
Tf	Int	11	
Df	Int	11	
Idf	Double		
W	Double		

e. Tabel query

Tabel tabel_query digunakan untuk penyimpanan sementara query inputan user yang telah dilakukan preprocessing dan bobot dari tiap term dalam query untuk nantinya digunakan dalam proses perhitungan cosine similiarity dengan term documents. Adapun struktur tabel_query adalah sebagai berikut:

Tabel 3.7 struktur tabel_query

Field	Type	Size	Keterangan
Id_query	Int	11	Incremental
Kata_query	Varchar	50	
Frequensi	Int	11	
Bobot	Double		

f. Tabel user

Tabel_user digunakan untuk menyimpan data admin dan pengunjung perpustakaan yang telah terdaftar sebagai mahasiswa ataupun dosen.

Berikut ini struktur tabel_user adalah sebagai berikut :

Tabel 3.8 struktur tabel_user

Field	Type	Size	Keterangan
Id	Int	11	Incremental
User_name	Varchar	50	
Password	Int	11	
Level	Varchar	15	

g. Tabel index_dasar

Tabel hasil stemming digunakan untuk menyimpan hasil perhitungan stemming pada synopsis buku. Berikut ini struktur

tabel_hasil_stemming :

Tabel 3.9 struktur tabel_stemming

Field	Type	Size	Keterangan
Id_index	Int	11	Incremental
Id_buku	Varchar	50	
Term	Varchar	50	
Tf	Varchar	20	

3.2.7. Perancangan Interface

Sistem ini akan dibangun dengan bahasa pemrograman PHP menggunakan editplus3 sebagai editor PHP-nya. Pada sistem pencarian ini terdapat beberapa halaman, sebagaimana berikut ini:

a. Halaman awal

Merupakan tampilan awal ketika user menjalankan aplikasi.

Berikut ini rancangan interface.

Gambar 3.17 Antarmuka halaman awal aplikasi

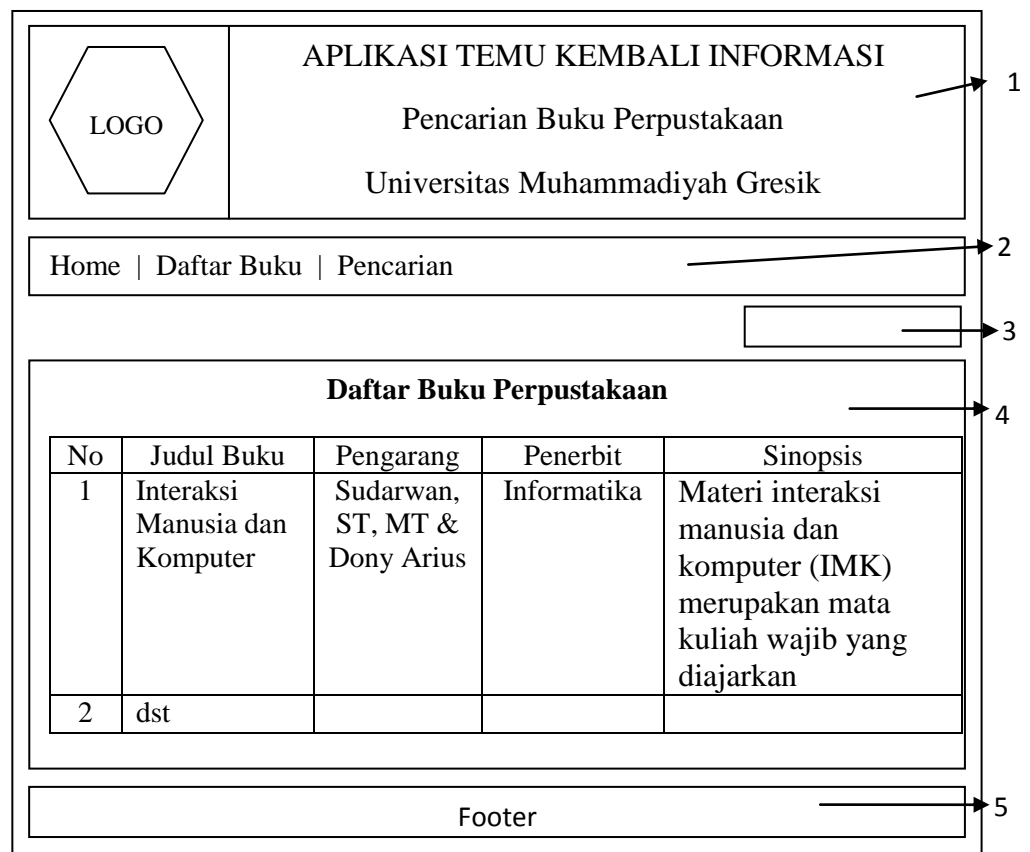
Keterangan :

1. Header berisi logo aktifitas akademik dan judul aplikasi
2. Label menu, yang terdiri dari home, daftar buku dan pencarian
3. ID , tempat user menginputkan ID yang telah disimpan didalam database sebelumnya.

4. Password, tempat user menginputkan password yang telah disimpan didalam database sebelumnya.
5. Tombol navigasi 'Ok', digunakan untuk melakukan proses login.
6. Deskripsi tentang aplikasi pencarian yang akan ditampilkan.
7. Footer.

b. Halaman Daftar Buku

Halaman ini merupakan tampilan bagi user yang menampilkan seluruh judul buku program studi teknik informatika. Perancangan tampilan interface halaman seperti dibawah ini.



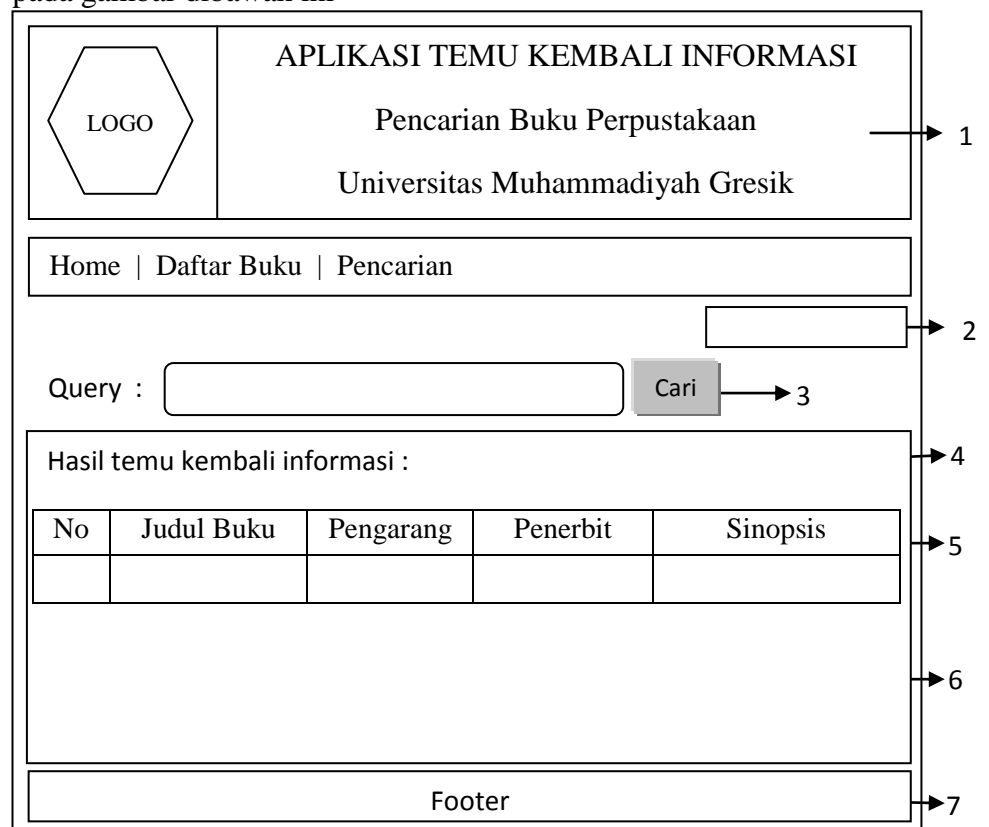
Gambar 3.18 Interface halaman aplikasi untuk melihat daftar buku yang dimiliki perpustakaan

Keterangan :

1. Header berisi logo aktifitas akademik dan judul aplikasi
2. Label menu, yang terdiri dari home, daftar buku dan pencarian
3. Tampilan ID user
4. Daftar buku yang dimiliki perpustakaan
5. Footer

c. Halaman pencarian

Halaman ini merupakan tampilan pencarian bagi user untuk mencari judul buku yang diinginkan dengan menginputkan query. Perancangan tampilan interface halaman ini bisa dilihat pada gambar dibawah ini



Gambar 3.19 Halaman interface pencarian

Keterangan :

1. Header berisi logo aktifitas akademik dan judul aplikasi
2. Label menu, yang terdiri dari home, daftar buku dan pencarian
3. Tampilan ID user
4. Label tempat query / topic buku diinputkan untuk mencari buku yang diinginkan
5. Tombol navigasi 'cari' untuk melakukan pencarian
6. Hasil pencarian
7. Footer

d. Halaman Input, Edit dan Hapus Buku

halaman berikut hanya bisa diakses oleh admin perpustakaan yang digunakan untuk memasukkan buku baru, edit dan menghapus buku di dalam perpustakaan.

The screenshot shows the 'Input, Edit and Delete Book' page. At the top, there is a header with a logo (1) and the title 'APLIKASI TEMU KEMBALI INFORMASI' (2). Below the header is a navigation menu with 'Home | Daftar Buku | Pencarian | Input buku' (3). A search bar (4) is located below the menu. The main content area contains a form for adding a new book with fields for 'Judul Buku' (5), 'Pengarang' (6), 'Penerbit' (7), and 'Sinopsis' (8). Below the form is an 'Entry' button (9). A table (10) lists existing books with columns for 'No', 'Judul Buku', 'Pengarang', 'Penerbit', 'Sinopsis', 'Edit', and 'Delete'. The 'Edit' and 'Delete' columns contain checkboxes (11). Below the table is an 'OK' button (12). The footer (13) is at the bottom of the page.

Gambar 3.20 Halaman Input, edit dan hapus buku

Keterangan :

1. Header berisi logo aktifitas akademik dan judul aplikasi
2. Label menu, yang terdiri dari home, daftar buku dan pencarian
3. Tampilan ID admin
4. Label memasukkan Judul Buku
5. Label memasukkan Pengarang buku
6. Label memasukkan Penerbit buku
7. Label memasukkan Synopsis buku
8. Tombol navigasi 'entry' untuk melakukan proses memasukkan buku kedalam database perpustakaan.
9. Daftar buku yang sudah tersimpan didalam database perpustakaan
10. Kolom pilihan untuk melakukan pengeditan buku.
11. Kolom pilihan untuk melakukan delete buku
12. Tombol navigasi 'ok' untuk melakukan proses edit atau delete buku setelah salah satu kolom edit atau delete dipilih.
13. Footer

3.2.8. Perancangan Uji Coba

Skenario Uji Coba

Skenario pengujian sistem pencarian jurnal skripsi yang dibangun ini difokuskan kepada hasil abstrak skripsi yang ditemukan. Pengujian dilakukan secara uji empiris. Dengan melakukan langkah-langkah berikut:

1. Input query dan melihat hasil pencarian
2. Mencatat hasil pencarian
3. Melakukan analisis terhadap hasil uji coba.

Analisis dilakukan dengan mencatat nilai threshold yang digunakan pada tiap percobaan, kemudian menghitung nilai precision, recall dan accuracy menggunakan persamaan rumus

(2.6, 2.7 dan 2.8) dari hasil pencarian. Hal ini dilakukan secara berulang. Dari hasil uji coba analisis dapat diketahui berapa nilai threshold terbaik untuk mendapatkan hasil pencarian yang dianggap relevan, dengan nilai precision, recall dan akurasi yang tinggi.

3.2.9. Kebutuhan Pembuatan Sistem

1. Analisis Perangkat Keras

Perangkat keras adalah device yang digunakan untuk menunjang dalam pembuatan sistem. Dalam pembuatan sistem ini perangkat keras yang digunakan yaitu laptop dengan spesifikasi :

- a. Processor AMD A8
- b. RAM 8GB
- c. HDD 750 GB
- d. Monitor 15.6"
- e. Mouse

2. Analisis Perangkat Lunak

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

- a. Windows 7
- b. Aplikasi server, dalam hal ini digunakan XAMPP
- c. Editor PHP, dalam hal ini digunakan Dreamweaver
- d. Browser, dapat menggunakan aplikasi internet seperti Internet Explorer, Mozilla Firefox, Opera dan lain-lain.
- e. SQLyog Enterprise