

PERANCANGAN WEB RANK MENGGUNAKAN COLLABORATIVE FILTERING BERDASARKAN KEMIRIPAN KONTEN

Eka Budhi Prasetya

Teknik Informatika Fakultas Teknik Universitas Muhammadiyah Jakarta
Jl. Cempaka Putih Tengah 27 Jakarta 10510
Ekabudhi96@yahoo.com

ABSTRAK

Internet berkembang pesat saat ini, ditujukan bagi manusia untuk mendapat pengetahuan dan informasi yang kompeten sesuai dengan kebutuhannya. Dalam perkembangannya, Internet memudahkan pengguna untuk memperoleh berbagai data dan informasi dalam kapasitas besar. Internet dapat mempengaruhi berbagai sektor kehidupan manusia, seperti dalam bidang bisnis, hiburan, berita dan sebagainya. Dikarenakan terlalu banyak informasi yang ada diinternet pengguna sering kali kesulitan menemukan informasi berkualitas secara cepat dan tepat. *Recommender system* mampu untuk mengurangi informasi yang terlalu berlebihan dan meningkatkan kepuasan bagi para pengguna internet. Pada *recommender system*, metode yang digunakan adalah *Collaborative Filtering* atau *Content-Based* untuk memprediksi *item* baru sesuai dengan keinginan *user*. Sistem recommender adalah bagian penting dari informasi dan e-commerce ekosistem. Sistem tersebut mewakili metode yang kuat untuk memungkinkan pengguna dalam menyaring informasi dan produk dalam ruang lingkup yang besar. Pencarian informasi yang berkualitas dan akurat menjadi hal yang sangat penting dalam berbagai aspek kehidupan. Pemberian ranking pada informasi sangatlah diperlukan agar pengguna dapat dengan mudah mencari informasi yang berkualitas sesuai dengan keinginan pengguna dan rekomendasi dari banyak pihak berdasarkan konten-konten tertentu. Pada penelitian ini akan membahas perancangan page rank menggunakan collaborative filtering berdasarkan kemiripan konten.

Kata kunci : *recommender system, collaborative filtering, content based, content-based collaborative filtering, page rank.*

ABSTRACT

Internet is growing rapidly today, is intended for humans to gain knowledge and information that is competent in accordance with their needs. In the process, the Internet allows users to obtain a variety of data and information in a large capacity. Internet can affect various sectors of human life, as in business, entertainment, news and so on. Due to too much information that is the internet users often have trouble finding quality information quickly and accurately. Recommender system is able to reduce the excessive information and improve satisfaction for Internet users. At the recommender system, the method used is Collaborative Filtering or Content-Based to predict new items in accordance with wishes of the user. Recommender system is an important part of the information and e-commerce ecosystem. The system represents a powerful method to allow users to filter the information and products in a large scope. Search quality and accurate information become very important in many aspects of life. Giving ranking the information is needed so that the user can easily search for quality information in accordance with the wishes of users and recommendations from many parties based on certain content. This research will discuss the design of a page rank using collaborative filtering based on similar content.

Keywords: *recommender system, collaborative filtering, content-based, content-based collaborative filtering, page rank*

PENDAHULUAN

1.1 Latar Belakang

Kebiasaan membaca berita di Indonesia mulai berubah seiring dengan semakin berkembangnya teknologi internet. Berawal dari media cetak kemudian tergiring menuju media elektronik berbasis internet seiring semakin maraknya penyedia layanan berita berbasis website. Menjamurnya penyedia layanan berita berbasis website di Indonesia menciptakan sebuah tantangan. Tantangan tersebut adalah kebutuhan untuk dapat memberikan rekomendasi berita sesuai dengan ketertarikan masing-masing individu.

Content-based filtering menganalisis kategori favorit pembaca untuk direkomendasikan berdasarkan pilihan berita pembaca tersebut di masa lalu. Saat ini *Google news* dapat merekomendasikan berita-berita berbagai bahasa namun tidak untuk berita-berita berbahasa Indonesia. Oleh karena itu dalam penelitian ini akan menganalisis pengimplementasian sistem rekomendasi berita berbahasa Indonesia berbasis pilihan personal menggunakan metode *collaborative filtering based content*.

Rating, Istilah ini biasa dipergunakan untuk mengukur efektivitas penggunaan pesawat televisi, radio atau media cetak sebagai sarana penyampaian pesan berita. *Rating* adalah suatu ukuran yang menunjukkan bagian dari sejumlah individu yang melihat, mendengarkan atau membaca suatu berita pada waktu tertentu, yang biasanya dinyatakan dalam persentase.

Pemberian *rating* pada informasi sangatlah diperlukan agar pengguna dapat dengan mudah mencari informasi yang berkualitas sesuai dengan keinginan pengguna dan rekomendasi dari banyak pihak berdasarkan konten-konten tertentu. Pada penulisan ini akan membahas perancangan *page rank* menggunakan *collaborative filtering* berdasarkan kemiripan konten.

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah disusun diatas maka dapat ditarik beberapa masalah yang timbul dari sistem *page rank* antara lain :

1. Menerapkan metode *collaborative filtering* pada *page rank*.

2. Menerapkan algoritma *page rank* pada web.
3. Mengimplementasikan *web page* dan menampilkan *rank* berdasarkan kemiripan content (*content based*).

1.3 Rumusan Masalah

Dari penjelasan pada bagian latar belakang diperoleh perumusan masalah sebagai berikut :

1. Bagaimana cara merancang *page rank* menggunakan *collaborative filtering based content*.
2. Bagaimana cara mengimplementasikan *web page* dan menampilkan *rank* berdasarkan kemiripan content (*content based*).

1.4 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut :

1. Menganalisa perancangan basis data yang akan digunakan dalam merancang *page rank*.
2. Mengimplementasikan metode *web spidering* yang akan digunakan pada perancangan *page rank*.
3. Merancang *web page* menampilkan *rank* berdasarkan kemiripan content (*content based*).
4. Jumlah 100 situs yang akan diindex.

1.5 Tujuan Penelitian

Tujuan dari penelitian ini adalah membuat dan merancang *web page* yang dapat menampilkan *rank* berdasarkan kemiripan content (*content based*) dengan menggunakan metode *collaborative filtering*.

1.6 Manfaat Penelitian

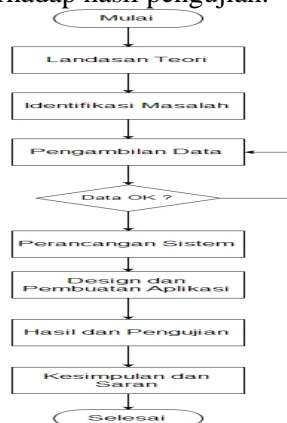
Manfaat penelitian ini adalah untuk memberikan informasi mengenai *rank* dari kumpulan web page berdasarkan kemiripan konten.

1. Membantu dengan memberikan informasi tentang kualitas dan popularitas dari sebuah website dan konten.
2. Mengoptimalkan waktu pencarian informasi yang diinginkan.
3. Memberikan hasil yang akurat terhadap informasi yang dicari.

1.7 Metodologi Penyelesaian Masalah

Metode penyelesaian masalah yang digunakan dalam menyelesaikan masalah diatas adalah sebagai berikut:

1. Studi Pustaka:
Pencarian referensi dan sumber yang berhubungan dengan konsep *collaborative filtering* yang dapat menangani *web page rank* dalam menyelesaikan penelitian ini.
2. Pengumpulan Data :
Mencari, mengumpulkan, dan mempelajari data yang berhubungan dengan proses pengerjaan penelitian ini.
3. Analisis dan Perancangan Sistem :
Melakukan analisis terhadap kebutuhan sistem, algoritma, dan karakteristik data yang digunakan sistem. Kemudian melakukan perancangan sistem berdasarkan analisis yakni pra-pengolahan dataset, arsitektur sistem, pemodelan, fungsionalitas, basisdata, pelatihan, dan pengujian.
4. Pembuatan Aplikasi :
Membuat aplikasi page rank sesuai dengan perancangan sistem, meliputi pembuatan database menggunakan phpmyadmin , script coding php menggunakan dan Adobe Dreaweaver CS5 serta webserver menggunakan Xampp.
5. Uji Coba dan Evaluasi:
Pengujian sistem untuk mengetahui tingkat akurasi prediksi yang dihasilkan dan melakukan analisis terhadap hasil pengujian.



Gambar 1.1 Flowchart Metodologi Penyelesaian Masalah

LANDASAN TEORI

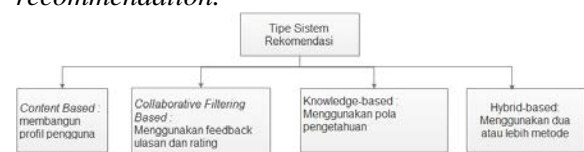
2.1 Konsep Dasar Sistem Rekomendasi

Konsep sistem rekomendasi telah digunakan secara luas oleh hampir semua area bisnis dimana seorang konsumen memerlukan informasi untuk membuat suatu keputusan (Alexandra Lilavati Pereira Okada, 2010). Sistem rekomendasi pariwisata atau perjalanan menggunakan konsep ini untuk menolong para wisatawan untuk memutuskan tujuan perjalanan mereka, akomodasi dan aktivitas di tempat tujuan.

Sistem rekomendasi merupakan model aplikasi dari hasil observasi terhadap keadaan dan keinginan pelanggan. Oleh karena itu sistem rekomendasi memerlukan model rekomendasi yang tepat agar yang direkomendasikan sesuai dengan keinginan pelanggan, serta mempermudah pelanggan mengambil keputusan yang tepat dalam menentukan produk yang akan digunakannya (McGinty, 2006).

Ada beberapa metode atau teknik yang digunakan dalam sistem rekomendasi. Setiap metode disesuaikan dengan permasalahan dalam menghasilkan sebuah informasi yang sesuai. (Adomavicius, 2005) menyatakan metode atau pendekatan yang dipilih pada sistem rekomendasi bergantung pada permasalahan yang akan diselesaikan, teknik rekomendasi yang berbeda-beda digunakan untuk aplikasi yang berbeda, dasar dari suatu tujuan dan objektif dari sebuah aplikasi. Dari penelitian terbaru metode atau teknik rekomendasi memiliki beberapa sejumlah kemungkinan klasifikasi (Adomavicius, 2005).

Berdasarkan metode rekomendasi yang sering digunakan, sistem rekomendasi dibagi dalam beberapa klasifikasi yaitu: *content-based recommendation*, *collaborative-based recommendation* dan *hybrid-based recommendation* dan beberapa peneliti menambahkan metode *knowledge based recommendation*.

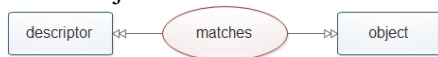


Gambar 2.1 Klasifikasi Sistem Rekomendasi (Adomavicius, 2005)

2.1.1 Content-based Recommendation

Menurut (Toine Bogers, 2007) *content-based recommendation* adalah hasil dari penelitian penyaringan informasi dalam sistem berbasis konten. Sistem rekomendasi berbasis konten dimulai dengan memahami kebutuhan *user* (pengguna), preferensi dan kendala jika ada. Informasi ini digabungkan dengan log dari interaksi *user* sebelumnya (jika ada) untuk membangun profil pengguna.

Kemudian sistem rekomendasi mencocokkan profil *user* (pengguna) dengan informasi tentang suatu produk yang telah tersimpan dalam database. Menurut (Baudisch, 1999) proses yang dilakukan dalam *content-based recommendation* yaitu melalui informasi retrieval dan sistem filtering informasi yang menampilkan keseluruhan perhitungan dari hubungan tunggal *descriptor x object*. Objek dapat berupa dokumen teks dan descriptor dapat berupa kata kunci. Gambar 2.2 berikut menunjukkan dua entitas dalam relasi antara descriptor dan object.



Gambar 3.2 Model Hubungan Entitas *Content-Based Retrieval* Informasi dan *Filtering* (Baudisch, 1999)

Namun menurut (Li, 2002) sistem rekomendasi berbasis konten memiliki kelemahan yaitu :

- Penyaringan berbasis konten sulit untuk menghasilkan rekomendasi yang tidak terduga (*serendipitous recommendation*), karena semua informasi dipilih dan direkomendasikan berdasarkan konten.
- Tidak *reliable* (handal) tidak adanya ide, pendapat serta masukan dari pengguna sebelumnya yang sulit dianalisis oleh komputer sehingga sulit untuk *user* (pengguna) pemula memakai sistem ini secara efektif.

2.1.2 Collaborative-based recommendation

Adomavicius & Tuzilin (2005) menyatakan sistem *collaborative-based recommendation* adalah metode yang digunakan untuk memprediksi kegunaan item berdasarkan penilaian pengguna sebelumnya.

Metode ini merekomendasikan item-item yang dipilih oleh pengguna lain dengan kemiripan model item dari pengguna saat ini (Sebastia, L et al., 2009).

Sistem rekomendasi berbasis COLLABORATIVE (*collaborative-based*) dibuat untuk mengatasi kelemahan dari sistem rekomendasi berbasis konten (*content-based*) yaitu:

- Pendekatan *collaborative* dapat bekerja dalam domain dimana terdapat sedikit *content* yang berasosiasi dengan item atau ditempatkan dimana *content* sulit dianalisis menggunakan komputer seperti ide, masukan atau opini sehingga menjadi *reliable*.
- Pendekatan *collaborative* mempunyai kemampuan untuk menyediakan rekomendasi yang tidak terduga atau tidak disengaja, misalnya dapat merekomendasikan item yang relevan kepada pengguna sekaligus tidak mengandung *content* dari profil pengguna tersebut.

(Baudisch ,1999) menyatakan dalam sistem *collaborative*, seluruh perhitungan berdasarkan *single relation user x object*. Gambar 2.3 berikut menunjukkan dua entitas dalam relasi antara descriptor dan object.



Gambar 4.3 Model Hubungan Entitas Dalam *Collaborative-Based System* (Baudisch ,1999)

2.1.3 Hybrid-based recommendation

Hybrid recommender system digunakan untuk menggambarkan setiap sistem rekomendasi yang menggabungkan beberapa teknik rekomendasi untuk menghasilkan sebuah output (Burk R. , 2007)). Adomavicius & Tuzilin (2005) menyatakan beberapa sistem rekomendasi menggunakan metode *hybrid* untuk menggabungkan metode *content-based* dan *collaborative* untuk membantu keterbatasan yang terdapat pada kedua metode tersebut. Menurut (Li, 2002) terdapat beberapa cara penggabungan yang dapat dilakukan dalam metode *hybrid* yaitu :

1. Penggabungan secara Linier (*Linier Combination*)

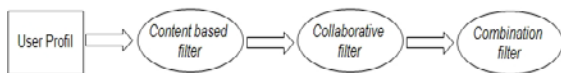
Penggabungan ini menggabungkan hasil rating dari metode *content-based* dan *collaborative*. Penggabungan ini dilakukan dengan cara pemberian ranking atau rating. Penggabungan ini digambarkan pada gambar 2.4 berikut :



Gambar 5.4 Kombinasi Linier (Li, 2002)

2. Penggabungan secara Sekuensial (*Sequential Combination*)

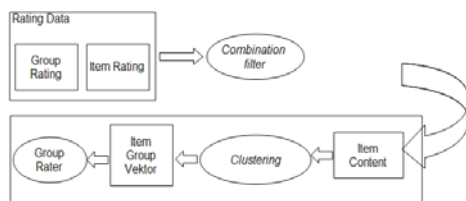
Penggabungan ini melakukan perhitungan pada salah satu metode (misalnya *content-based*) kemudian hasilnya digabungkan dengan perhitungan metode lainnya (*collaborative*). Penggabungan ini digambarkan pada gambar 2.5 berikut :



Gambar 6.5 Kombinasi Sekuensial (Li, 2002)

3. Penggabungan secara *Item-based Clustering Hybrid Method (ICHM)*

Penggabungan ini mengintegrasikan informasi item dan rating pengguna untuk menghitung kemiripan item-item. Penggabungan ini digambarkan pada gambar 2.6 berikut :



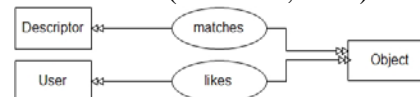
Gambar 7.6 *Item-based Clustering Hybrid Method* (Li, 2002)

(Burk R. , 2002) memperkenalkan taxonomy untuk sistem rekomendasi *hybrid*. *taxonomy* ini diklasifikasikan ke dalam tujuh kategori, yaitu:

- Weighted hybrid* : Nilai komponen dari sistem rekomendasi yang berbeda digabungkan secara numerik atau menggunakan algoritma linier.

- Switching hybrid* : Sistem memilih komponen-komponen dari setiap rekomendasi dan menerapkan komponen yang dipilih.
- Mixed hybrid* : Rekomendasi dari berbagai sistem rekomendasi disajikan bersama
- Feature Combination* : Fitur-fitur yang berasal dari berbagai sumber pengetahuan digabungkan dan diberikan algoritma rekomendasi
- Feature Augmentation* : merupakan salah satu teknik rekomendasi yang digunakan untuk menghitung sebuah fitur atau sekumpulan fitur yang kemudian menjadi bagian yang dimasukkan ke teknik berikutnya.
- Cascade* : merupakan rekomendasi yang memiliki prioritas tinggi sebagai solusi pemecahan masalah dalam melakukan perbaikan
- Meta-level* : merupakan salah satu teknik rekomendasi yang diterapkan dan menghasilkan beberapa jenis model, yang kemudian digunakan sebagai input oleh teknik berikutnya.

Berikut ini merupakan gambar model hubungan entitas dari metode *hybrid* menurut (Baudisch, 1999)



Gambar 8.7 Model Hubungan Entitas Dalam *Hybrid System* (Baudisch, 1999)

Hubungan entitas diatas menyediakan fungsionalitas dari sebuah sistem *content-based* dan juga fungsionalitas dari *collaborative system*.

2.2 Search Engine

Page rank adalah aplikasi yang didesain untuk mencari informasi dari *world wide web/ internet*. Internet terdiri dari triliun data dan informasi dalam jutaan server dan lokasi yang tersebar luas. *Page rank* menkoleksi data, gambar, informasi dalam bentuk indeks. Informasi yang disajikan dalam bentuk daftar kepada pencari berdasarkan info yang diberikan oleh pencari.

Cara kerja *page rank* berdasarkan urutan berikut ini:

- Web Crawling*

2. *Indexing*
3. *Searching*

Web crawler (disebut juga *spider*) adalah program komputer yang melihat-lihat (*browse*) halaman *web* serta mengikuti setiap *link* di halaman tersebut, dan kemudian menyimpannya dalam indeks di server *search engine*. Isi dari halaman tersebut dianalisa untuk mendapatkan index kata-kata dari judul, *header*, dan *meta tag*. *Index* memungkinkan hasil *query* yang sependan dengan cepat.

Saat pengguna menuliskan pertanyaan (*query*) ke *search engine*, maka *page rank* akan *searching* (mengevaluasi) *index* dan memberikan daftar halaman *web* yang paling sesuai dengan pertanyaan (*query*), beserta ringkasan singkat yang terdiri dari judul dokumen dan sebagian dari teks.

Saat pengguna menuliskan pertanyaan (*query*) ke *search engine*, maka *page rank* akan *searching* (mengevaluasi) *index* dan memberikan daftar halaman *web* yang paling sesuai dengan pertanyaan (*query*), beserta ringkasan singkat yang terdiri dari judul dokumen dan sebagian dari teks.

PERANCANGAN SISTEM

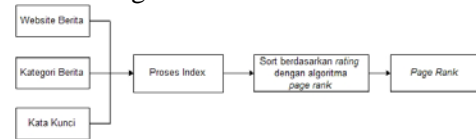
3.1 Tinjauan Umum

Secara umum, *sistem recommender* bertugas membantu para user dengan cara menyediakan rekomendasi bagi para user-nya. Rekomendasi yang diberikan berupa berita yang mungkin cocok dengan keinginan *user*. Pada sistem ini *user* dapat memberikan rating pada berbagai macam berita yang sudah dibacanya. *Rating* yang diberikan oleh masing-masing *user* tersebut digunakan sebagai dasar untuk menentukan rekomendasi berita bagi *user* yang lain. Nilai *rating* dari berita yang ditentukan secara jelas oleh *user* adalah input yang diperlukan oleh sistem.

Berdasarkan input yang diberikan oleh user, sistem akan menciptakan output berupa rekomendasi berita-berita apa saja yang diperkirakan akan disukai oleh *user* yang lain. *User* juga dapat melihat nilai prediksi *rating* yang diberikan oleh sistem terhadap masing-masing judul berita yang terdapat pada sistem tersebut. Sistem Rekomendasi berita ini dibuat karena disebabkan adanya permasalahan yang dialami oleh para user dalam menentukan berita apa saja yang ingin dibaca oleh user.

3.1.1 Konsep Perancangan *Page rank*

Konsep perancangan *page rank* dapat digambarkan sebagai berikut :



Gambar 3. 1 Konsep Perancangan *Page Rank*

3.1.2 Metode Web Spider Pada Perancangan *Page rank*

Salah satu cara yang digunakan *search engine* untuk mengideks secara cepat adalah mengumpulkan informasi tentang halaman *web* secara periodik dan otomatis dengan menggunakan program yang disebut *web spider*. Beberapa istilah lain yang digunakan seperti *ants*, *automatic indexer*, *crawler*, *bot*, *web robots*, *worms*, merujuk pada hal yang sama.

Spider ini berkelana di *web* dan berfungsi untuk mengumpulkan segala informasi tentang suatu halaman *web* dan mengindeksnya ke dalam suatu *database*. Informasi tentang halaman *web* tersebut didapat dari kata-kata yang terdapat di dalam halaman *web* tersebut. Kata-kata tersebut kemudian di indeks menjadi sebuah daftar kata-kata yang biasanya dipakai sebagai kata kunci untuk menemukan halaman *web*. Proses pengumpulan informasi tentang halaman *web* dari situs web ini disebut *web spidering*.

Pengembangan model aplikasi *web spider* ini dilakukan untuk memperlihatkan bagaimana teknologi *web spider* yang digunakan dapat secara cepat membangun *database* dari suatu situs *web* sehingga *user* dapat melakukan pencarian dengan hasil yang baik. Model *spider* ini hanya mengindeks file dan folder yang terhubung dan berada pada *web-server* tersebut dan tidak akan keluar dari *root domain* yang telah ditentukan. Sub domain tidak akan mengalami proses indeksing secara otomatis dan akan dibaca sebagai alamat situs *web* yang berbeda.

Pengembangan model dilakukan pada Windows 7 dengan *Web Server Apache* menggunakan PHP dengan *database* menggunakan MySQL *database*. Hal utama dalam melakukan *spidering* adalah cara membaca sebuah file html dan mengumpulkan *links* yang ada dalam file tersebut dengan

mengikuti *links* tersebut seperti sebuah *web browser*.

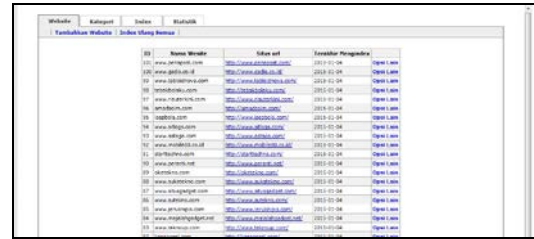
Sedangkan *indexing* mencakup proses membaca file html, mengambil isinya, mengubahnya kedalam text biasa, menyimpulkan isi yang terkandung dalam file tersebut. Kemudian mngumpulkan *keywords* yang berhubungan dan menyimpannya sebagai bahan perbandingan *query*. Langkah-langkah utama dalam melakukan *spidering* dapat digambarkan sebagai berikut :

1. Menentukan file html sebagai starting point yaitu berupa file utama yang memiliki banyak *links* terhadap isi dari *website* yang akan di index, biasanya file root dari suatu situs *web*. Dalam mengikuti *link*, perlu di tentukan seberapa dalam kita akan mengikuti *link*, untuk mudahnya ini disebut level *spidering*.
2. Membaca file tersebut dalam text biasa, kemudian mengumpulkan *links* yang ada dalam file tersebut dan menyimpannya ke dalam *database*. Ini disebut level 0. HTML merupakan file teks yang memiliki struktur yang jelas yang dapat dibaca oleh *browser*. *Links* dalam HTML files berupa kode yang memiliki tags “” dan beberapa kode script dalam vbscript seperti “windows.open” atau yang lainnya yang menunjukkan
3. *link* terhadap halaman lain. Alamat *link* inilah yang disimpan kedalam *database*.
4. Membuka *link* berikutnya seperti *web browser* dan mengulang point 2 sampai *link* yang ada dalam level ini habis. *Links* yang di kumpulkan merupakan daftar *link* pada level berikutnya.
5. Lakukan point 2 dan 3 secara berulang sampai dicapai level yang diinginkan.

PEMBUATAN APLIKASI DAN PENGUJIAN

4.1 Tampilan Output

Berikut ini adalah beranda halaman admin.



No	Nama Website	Alamat	Tanggal Pengindeks	Status
1	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
2	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
3	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
4	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
5	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
6	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
7	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
8	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
9	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
10	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
11	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
12	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
13	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
14	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
15	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
16	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
17	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
18	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
19	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1
20	www.gelita.id	http://www.gelita.id	2015-01-04	Page 1

Gambar 4. 1 Tampilan Beranda Web Admin Lihat Data Website

4.1.1 Tampilan Statistik

Berikut ini adalah halaman statistik

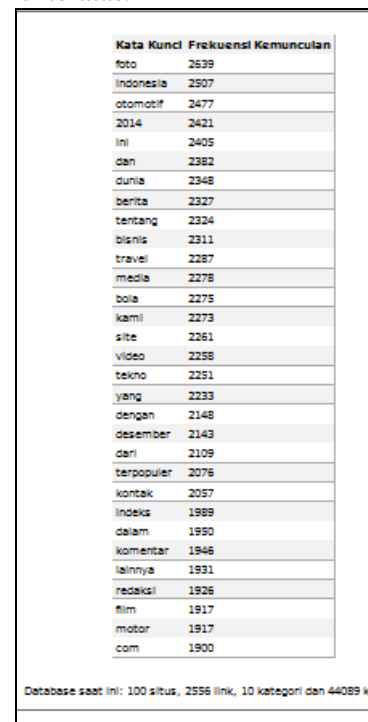


Statistik	
Pencarian Terpopuler Pencarian	
Situs	100
Link	2556
Kategori	10
Kata Kunci	44089
Relasi Kata Kunci dengan Link	902225
Total Text Chaced	38,224.80 kb
Total Ukuran Situs	134,808.66 kb
Database saat ini: 100 situs, 2556 link, 10 kategori dan 44089 kata kunci.	

Gambar 4. 2 Tampilan Halaman Statistik

4.1.2 Tampilan Kata Kunci Teratas

Berikut ini adalah tampilan halaman kata kunci teratas.



Kata Kunci	Frekuensi Kemunculan
foto	2639
Indonesia	2507
otomotif	2477
2014	2421
ini	2405
dan	2382
dunia	2348
berita	2327
tentang	2324
bisnis	2311
travel	2287
media	2278
bola	2275
kami	2273
site	2261
video	2258
teknologi	2251
yang	2233
dengan	2148
desember	2143
dari	2109
terpopuler	2076
kontak	2057
indeks	1989
dalam	1950
komentar	1946
lainnya	1931
redaksi	1926
film	1917
motor	1917
com	1900
Database saat ini: 100 situs, 2556 link, 10 kategori dan 44089 ka	

Gambar 4. 3 Tampilan Halaman Kata Kunci Teratas

4.1.3 Tampilan Halaman Terbesar

Berikut ini adalah tampilan halaman terbesar.

Kategori	Endek	Statistik
Halaman	Ukuran Yang	
http://www.kompas.com/berita/indonesia	235.91kb	125.33%
http://www.kompas.com/berita/indonesia	47.96kb	23.24%
http://www.kompas.com/berita/indonesia	44.37kb	22.00%
http://www.kompas.com/berita/indonesia	44.13kb	21.87%
http://www.kompas.com/berita/indonesia	43.92kb	21.73%
http://www.kompas.com/berita/indonesia	42.05kb	20.78%
http://www.kompas.com/berita/indonesia	39.75kb	19.65%
http://www.kompas.com/berita/indonesia	38.96kb	19.28%
http://www.kompas.com/berita/indonesia	38.92kb	19.26%
http://www.kompas.com/berita/indonesia	38.90kb	19.25%
http://www.kompas.com/berita/indonesia	38.89kb	19.24%
http://www.kompas.com/berita/indonesia	38.88kb	19.23%
http://www.kompas.com/berita/indonesia	38.87kb	19.22%
http://www.kompas.com/berita/indonesia	38.86kb	19.21%
http://www.kompas.com/berita/indonesia	38.85kb	19.20%
http://www.kompas.com/berita/indonesia	38.84kb	19.19%
http://www.kompas.com/berita/indonesia	38.83kb	19.18%
http://www.kompas.com/berita/indonesia	38.82kb	19.17%
http://www.kompas.com/berita/indonesia	38.81kb	19.16%
http://www.kompas.com/berita/indonesia	38.80kb	19.15%

Database saat ini: 100 situs, 2376 link, 10 kategori dan 44839 kata kunci.

Gambar 4. 4 Tampilan Halaman Terbesar

4.1.4 Tampilan Halaman Pencarian Terpopuler

Berikut ini adalah tampilan halaman pencarian terpopuler.

Kategori	Endek	Statistik	
Kata Pencarian	Jumlah Penelusuran	Hasil Rata-rata	Penelusuran Terakhir
jokowi	32	1.629,3	2015-01-06 10:10:10
presiden	9	1.201,1	2015-01-14 09:50:42
r	4	0,0	2015-01-03 14:52:22
j	3	0,0	2015-01-03 15:57:18
rsni	3	101,0	2015-01-04 10:44:45
DATANG BULAN	2	77,5	2015-01-04 15:51:47
sa	1	0,0	2015-01-04 15:46:35
wanita	1	221,0	2015-01-03 15:37:09
presi	1	0,0	2015-01-12 09:42:23
masak	1	23,0	2015-01-04 15:42:09
politik	1	714,0	2015-01-14 09:50:27
masak kue	1	0,0	2015-01-04 15:42:26
raffi	1	173,0	2015-01-04 15:42:42
gigi	1	117,0	2015-01-04 15:43:09

Database saat ini: 100 situs, 2356 link, 10 kategori dan 44089 kata kunci.

Gambar 4. 5 Tampilan Halaman Pencarian Terpopuler

4.1.5 Tampilan Halaman History Pencarian Kata Kunci

Berikut ini adalah tampilan halaman history pencarian kata kunci.

Kategori	Endek	Statistik	
Kata Pencarian	Hasil Penelusuran	Waktu Penelusuran	Waktu yang Terjadi
presiden	1306	2015-01-14 09:50:41	0,49
putra	714	2015-01-14 09:50:27	0,54
presiden	1206	2015-01-14 09:50:41	0,49
presiden	1306	2015-01-12 09:44:20	0,41
rasi	8	2015-01-13 09:43:33	0,86
presiden	1306	2015-01-12 09:44:20	0,41
presiden	1306	2015-01-09 11:50:23	0,59
jokowi	1481	2015-01-04 15:42:10	0,59
jokowi	1481	2015-01-04 15:42:10	0,59
jabana	8	2015-01-04 15:42:26	0,79
presiden	1306	2015-01-04 15:42:26	0,79
presiden	1306	2015-01-04 15:42:26	0,79
DATANG BULAN	8	2015-01-04 15:42:42	0,44
sa	8	2015-01-04 15:42:42	0,44
sa	8	2015-01-04 15:42:42	0,44
gigi	117	2015-01-04 15:43:09	0,19
raffi	173	2015-01-04 15:42:42	0,23
masak kue	8	2015-01-04 15:42:26	0,67
masak	23	2015-01-04 15:42:09	0,67
rsni	181	2015-01-04 10:44:45	1,39
j	8	2015-01-03 15:57:18	0,8

Gambar 4. 6 Tampilan Halaman History Pencarian Kata Kunci

5.1 KESIMPULAN

Dari hasil perancangan sistem informasi page rank, dapat kesimpulan sebagai berikut :

1. Dengan sistem ini maka akan memudahkan user untuk mencari informasi yang diinginkan secara tepat, cepat dan akurat berdasarkan perangkingan pada page rank.
2. Pada sistem ini memudahkan administrator dalam mengindex website, dan menampilkan informasi yang berguna bagi admin antara lain statistik, page rank, kata kunci teratas, kata kunci yang paling sering dicari oleh user, history penelusuran.

5.2 Saran-saran

Website ini hanya mampu memfilter kata kunci untuk dijadikan *page rank*. Dalam perkembangan *collaborative filtering* yang akan dating diharapkan mampu menyaring dokumen-dokumen berdasarkan type file dengan format doc,pdf,ppt,jpg, bmp dan lain-lain.

DAFTAR PUSTAKA

Resi Adeline, Teknik Normalisasi dalam Perancangan Basis Data Relasional, Jakarta : 2010.

Whitten, Jeffrey, Analisis dan Perancangan Sistem Informasi Manajemen, Elexmedia, Jakarta : 2004.

Yamit, Zulian, “ Dasar – Dasar Collaboration Filtering “Elexmedia, Yogyakarta : 2005.