

**SISTEM REKOMENDASI ACARA TELEVISI
MENGUNAKAN ALGORITMA
ITEM-BASED TOP-N RECOMMENDATION**

SKRIPSI

Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Komputer dalam bidang Ilmu Komputer

oleh:

ADI PRATAMA

0210960002-96



**PROGRAM STUDI ILMU KOMPUTER
JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU
PENGETAHUAN ALAM
UNIVERSITAS BRAWIJAYA
MALANG
2009**

LEMBAR PENGESAHAN SKRIPSI
SISTEM REKOMENDASI
ACARA TELEVISI MENGGUNAKAN ALGORITMA
ITEM-BASED TOP-N RECOMMENDATION

oleh:

ADI PRATAMA
0210960002-96

Setelah dipertahankan di depan Majelis Penguji
pada tanggal 12 Agustus 2009
dan dinyatakan memenuhi syarat untuk memperoleh gelar
Sarjana Komputer dalam bidang Ilmu Komputer

Pembimbing I

Pembimbing II

Drs. Marji, MT
NIP. 131 993 386

Dany Primanita K, ST
NIP. 132 310 159

Mengetahui,
Ketua Jurusan Matematika
Fakultas MIPA Universitas Brawijaya

Dr. Agus Suryanto, MSc.
NIP. 132 126 049

LEMBAR PERNYATAAN

Saya yang bertanda tangan di bawah ini :

Nama : Adi Pratama
NIM : 0210960002
Jurusan : Matematika
Program Studi : Ilmu Komputer
Penulis tugas akhir berjudul : Sistem Rekomendasi Acara
Televisi menggunakan algoritma
Item-Based Top-N Recommendation

Dengan ini menyatakan bahwa :

1. Isi dari tugas akhir yang saya buat adalah benar-benar karya sendiri dan tidak menjiplak karya orang lain, selain nama-nama yang termaktub di isi dan tertulis di daftar pustaka dalam Tugas Akhir ini.
2. Apabila dikemudian hari ternyata Tugas Akhir yang saya tulis terbukti hasil jiplakan, maka saya akan bersedia menanggung segala resiko yang akan saya terima.

Demikian pernyataan ini di buat dengan segala kesadaran.

Malang, 12 Agustus 2009

Yang menyatakan,

Adi Pratama

NIM. 0210960002

SISTEM REKOMENDASI

ACARA TELEVISI MENGGUNAKAN ALGORITMA

ITEM-BASED TOP-N RECOMMENDATION

Rekomendasi acara televisi merupakan suatu proses pemberian sebuah rekomendasi terhadap acara televisi menurut minat atau kesukaan masing-masing pemirsa/user. Rekomendasi acara televisi dibutuhkan karena kemajuan pertelevisian saat ini. Kemajuan tersebut ditunjukkan dengan banyak dan beragamnya informasi atau acara yang disajikan. Beragam informasi disajikan oleh semua stasiun televisi melalui berbagai acara yang menyebabkan semakin seringnya pemirsa memindah chanel TV untuk mencari acara yang sangat sesuai dengan keinginan. Sistem rekomendasi bekerja membantu pemirsa dalam memilih atau menentukan acara yang akan dilihat dengan cara memberi informasi atau sugesti berdasarkan tingkat ketertarikan atau selera masing-masing pemirsa. Ketertarikan tersebut dapat diketahui dari data rating yang telah diberikan oleh user tersebut. Salah satu metode yang digunakan dalam memberikan rekomendasi adalah dengan menggunakan *Item-Based Top-N Recommendation*. Dengan mencari kedekatan rating antar item acara yang berkorelasi menggunakan *similarity*, maka akan diketahui seberapa dekat item acara tersebut dengan item acara yang lain. Semakin dekat item acara tersebut maka semakin besar item acara tersebut direkomendasikan. Cara penghitungan *similarity* yang digunakan adalah dengan menggunakan *cosine-based*. Setelah dilakukan penghitungan *similarity*, maka untuk user yang memberikan rating dapat dihitung prediksi item acara yang direkomendasikan. Penghitungan prediksi dilakukan dengan menggunakan cara *Weighted Sum*. Dari rekomendasi yang diberikan kepada user ada 2 macam, yaitu rekomendasi *personalized* dan *non-personalized*. Rekomendasi *personalized* bagi user terdaftar yang memberikan rating dan rekomendasi *non-personalized* bagi user terdaftar yang tidak memberikan rating, atau bagi user yang hanya berkunjung tanpa melakukan pendaftaran. Setelah dilakukan pengukuran tingkat akurasi dengan menggunakan *Mean Absolute Error (MAE)* didapati bahwa terjadi error sebesar 0,8141. Dan waktu yang dibutuhkan untuk menampilkan rekomendasi kepada user adalah 1,05 detik.

TELEVISION EVENTS RECOMMENDER SYSTEM USING ITEM-BASED TOP-N RECOMMENDATION ALGORITHM

Recommendations television is a process of giving a recommendation to the television according to the interests or preferences of each viewer / user. Recommendations television is required because progress at this time. Progress is shown with the many and various information or events are presented. Various information presented by all television stations through various events that cause the audience often move TV channel to search for events that are in accordance with the desire. Recommendation system is help the audience to work in the select or determine the events that will provide information in a way or suggestion based on the level of interest or taste of each audience can be known from the data rate that has been provided by the user. One of the methods used in providing recommendations is to use the Item-Based Top-N Recommendation. With the search for closeness between rating items correlated events using similarity, they will know how close the items event with items other events. The item near the event the more likely event of recommended items. The similarity that is used is cosine-based. After calculating the similarity be done, to provide a user rating the item can be counted input events recommended. Calculations made using the prediction is Weighted Sum. Of the recommendations given to the user there are 2 kinds, namely personalized recommendations and non-personalized. Personalized recommendations for registered users provide ratings and non-personalized recommendations for non-registered user who does not give ratings, or for users who only visit without registering. After a measurement accuracy by using the Mean Absolute Error (MAE) found that there was an error of 0,8141. And the time required to display the recommendations to the user is 1,05 seconds.

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT yang telah melimpahkan segala Rahmat, Karunia dan Hidayah-Nya sehingga Penulis dapat menyelesaikan Skripsi dengan judul: **”Sistem rekomendasi acara televisi menggunakan algoritma *Item-based Top-N Recommendation*”**

Skripsi ini diajukan sebagai syarat untuk memperoleh gelar Sarjana Komputer di Fakultas MIPA, Program Studi Ilmu Komputer Jurusan Matematika, Universitas Brawijaya Malang. Atas terselesaikannya skripsi ini, Penulis mengucapkan terima kasih kepada:

1. Drs. Marji., MT. selaku Dosen Pembimbing Tugas Akhir.
2. Dany Primanita Kartikasari, ST. selaku Dosen Pembimbing Tugas Akhir
3. Wayan Firdaus Mahmudy, S.Si., MT., selaku Ketua Program Studi Ilmu Komputer Jurusan Matematika FMIPA Universitas Brawijaya.
4. Drs. Muh. Arif Rahman, M.Kom, selaku Dosen Penasehat Akademik.
5. Dr. Agus Suryanto, MSc., selaku Ketua Jurusan Matematika FMIPA Universitas Brawijaya.
6. Segenap Bapak dan Ibu dosen yang telah mendidik dan mengajarkan ilmunya kepada Penulis selama menempuh pendidikan di Program Studi Ilmu Komputer Jurusan Matematika FMIPA Universitas Brawijaya.
7. Segenap staf dan karyawan di Jurusan Matematika FMIPA Universitas Brawijaya yang telah banyak membantu Penulis dalam pelaksanaan penyusunan tugas akhir ini.
8. Orang tua Penulis dan adik-adikku atas segala dukungan dan doa yang telah diberikan.
9. Keluarga besar PT. Jalin Mayantara Indonesia (Jayantara) atas seluruh dukungan yang telah diberikan.
10. Rekan-rekan Ilmu Komputer yang telah memberikan dukungannya kepada penulis.

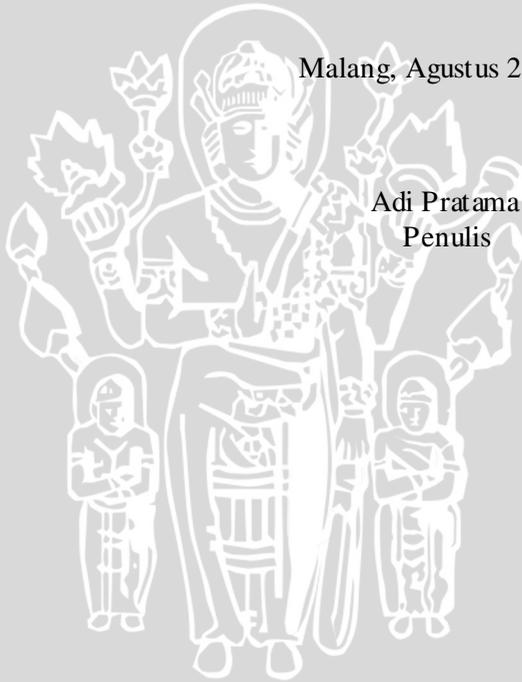
11. Semua pihak yang telah membantu terselesaikannya tugas akhir ini yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa tugas akhir ini tentunya tidak terlepas dari berbagai kekurangan dan kesalahan. Oleh karena itu, segala kritik dan saran yang bersifat membangun sangat Penulis harapkan dari berbagai pihak demi penyempurnaan penulisan tugas akhir ini.

Akhirnya penulis berharap agar tugas akhir ini dapat memberikan sumbangan serta manfaat bagi semua pihak yang berkepentingan.

Malang, Agustus 2009

Adi Pratama
Penulis



DAFTAR ISI

	Halaman
Halaman Judul	i
Halaman Pengesahan	ii
Lembar Pernyataan	iii
Abstrak	iv
Kata Pengantar	vi
Daftar Isi	viii
Daftar Gambar	xi
Daftar Tabel	xiii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah	3
1.5 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Data Mining	5
2.1.1 Association Rule	5
2.1.2 Clustering	6
2.1.3 Classification	6
2.2 Web Mining	7
2.3 Sistem Informasi Personal (Personalized Information System)	10
2.4 Recommender System	10
2.4.1 Content-Based Filtering	10
2.4.2 Collaborative Filtering	11
2.4.2.1 Model-based atau item-based	12
2.4.2.2 Memory-based atau user-based	13
2.4.2.3 Tingkat Personalisasi Rekomendasi	13
2.5 Algoritma Item-based Top-N Recommendation	14
2.5.1 Menghitung kedekatan antar item (Item-Item Similarity)	14
2.5.2 Penghitungan Prediksi	16
2.6 Pengukuran Tingkat Akurasi	17

BAB III METODE PENELITIAN	18
3.1 Tahapan pengembangan sistem rekomendasi	18
3.2 Kebutuhan sistem	18
3.2.1 Kebutuhan perangkat keras	19
3.2.2 Kebutuhan perangkat lunak	19
3.3 Perancangan Item-Based Top-N Recommendation	20
3.4 Proses Penelitian	25
3.5 Alur proses Model Item-Based Top N Recommendation	26
3.6 Alur proses pemberian rating oleh user	27
3.7 Rekomendasi Non-Personalized	28
3.8 Rekomendasi Personalized	29
3.9 Rekomendasi Top-N	30
3.10 Alur Pemberian Rekomendasi	30
3.11 Jadwal acara televisi	32
3.12 Perancangan Tabel	33
3.13 Dataset Acara Televisi	37
3.14 Rancangan Hasil Penelitian	37
3.15 Rancangan tampilan antar muka (user interface)	39
3.15.1 Home	39
3.15.2 Login	39
3.15.3 Daftar User Baru	40
3.15.4 Daftar Genre	40
3.15.5 Genre Acara per Genre	41
3.15.6 Detail Acara Bagi User yang Tidak Melakukan Login	41
3.15.7 Detail Acara Bagi User yang Melakukan Login	42
3.15.8 Halaman Profil User	43
3.15.9 Halaman Rating User	43
BAB IV HASIL & PEMBAHASAN	44
4.1 Hasil	44
4.1.1 Persiapan Data	44
4.1.2 Implementasi Basis Data	44
4.1.3 Deskripsi Program	46
4.1.3.1 Fungsi penghitungan similarity	46
4.1.3.2 Fungsi hitung Wighted SUM	47
4.1.3.3 Fungsi menampilkan rekomendasi personalized	48
4.1.3.4 Fungsi menampilkan rekomendasi non-personalized	48

4.1.4	Pembuatan Model Item	49
4.1.5	Hasil yang ditampilkan	54
4.1.5.1	Home.....	54
4.1.5.2	Login	54
4.1.5.3	Daftar User Baru	55
4.1.5.4	Daftar Genre.....	55
4.1.5.5	Daftar Acara per Genre	56
4.1.5.6	Detail Acara Bagi User yang Tidak Melakukan Login..	56
4.1.5.7	Detail Acara Bagi User yang Melakukan Login	57
4.1.5.8	Halaman Profil User	59
4.1.5.9	Halaman Rating User.....	59
4.2	Analisa Hasil	60
4.1.4	Hasil MAE (Mean Absolute Error).....	60
4.1.5	Waktu Komputasi Rekomendasi	61
BAB V PENUTUP.....		63
5.1	Kesimpulan	63
5.2	Saran	63
Daftar Pustaka		64

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Contoh decision tree sederhana	7
Gambar 2.2 Taksonomi web mining	8
Gambar 2.3 Arsitektur Content-based filtering	11
Gambar 2.4 Alur model-based/ item-based CF	12
Gambar 2.5 Alur memory-based/ user-based CF	13
Gambar 2.6 Penghitungan similarity terhadap 2 item yang berkorelasi	14
Gambar 3.1 Blok diagram proses penelitian	25
Gambar 3.2 Alur proses pembuatan model data	26
Gambar 3.3 Alur proses sistem pemberian rating	27
Gambar 3.4 Alur pemberian rekomendasi	31
Gambar 3.5 Rancangan halaman home /beranda awal	39
Gambar 3.6 Rancangan halaman login	39
Gambar 3.7 Rancangan halaman pendaftaran user baru	40
Gambar 3.8 Rancangan halaman daftar genre	40
Gambar 3.9 Rancangan halaman daftar acara	41
Gambar 3.10 Rancangan detail nama acara dan pemberian rekomendasi non-personalized	41
Gambar 3.11 Rancangan detail nama acara, pemberian rating dan pemberian rekomendasi personalized	42
Gambar 3.12 Rancangan halaman profil user	43
Gambar 3.13 Rancangan halaman daftar rating yang telah diberikan oleh user	43
Gambar 4.1 Waktu komputasi similarity terhadap peningkatan item-user	51
Gambar 4.2 Waktu komputasi prediksi terhadap peningkatan item-user	52
Gambar 4.3 Halaman home	54
Gambar 4.4 Halaman login	54
Gambar 4.5 Tampilan daftar user baru	55
Gambar 4.6 Tampilan halaman daftar genre	55
Gambar 4.7 Tampilan halaman daftar acara	56
Gambar 4.8 Halaman detail acara dan rekomendasi non-personalized	56

Gambar 4.9 Tampilan user yang login dengan pemberian rekomendasi	58
Gambar 4.10 Halaman profil user	59
Gambar 4.11 Tampilan halaman rating user.....	59
Gambar 4.12 Hasil MAE dengan beragam rasio dataset	61

UNIVERSITAS BRAWIJAYA



DAFTAR TABEL

	Halaman
Tabel 3.1 Contoh user-item rating.....	20
Tabel 3.2 Item dari user Z	20
Tabel 3.3 Item- i user Z terhadap item A	21
Tabel 3.4 Item-i user Z Terhadap item C	21
Tabel 3.5 Hasil pemasangan item dan hasil similarity	24
Tabel 3.6 Prediksi item terhadap user Z	24
Tabel 3.7 Similarity untuk menghitung prediksi terhadap user Z...29	29
Tabel 3.8 Hasil prediksi terhadap user target	29
Tabel 3.9 Jadwal item-j pada user Z.....	32
Tabel 3.10 Prediksi yang diberikan kepada user Z	32
Tabel 3.11 Total tabel database	33
Tabel 3.12 Tabel item / acara.....	33
Tabel 3.13 Tabel genre.....	34
Tabel 3.14 Tabel user.....	34
Tabel 3.15 Tabel rating.....	35
Tabel 3.16 Tabel similarity.....	35
Tabel 3.17 Tabel prediksi.....	36
Tabel 3.18 Tabel jadwal.....	36
Tabel 3.19 Rancangan hasil Mean Absolute Error	38
Tabel 3.20 Rancangan tabel hasil komputasi rekomendasi	38
Tabel 4.1 Rekap jumlah data untuk membentuk model	44
Tabel 4.2 Penghitungan similarity terhadap sensitivitas item-user..50	50
Tabel 4.3 Penghitungan prediksi terhadap sensitivitas item-user ..51	51
Tabel 4.4 Waktu komputasi pembuatan model item	53
Tabel 4.5 Hasil Mean Absolute Error	60
Tabel 4.6 Hasil komputasi rekomendasi	62

BAB I PENDAHULUAN

1.1 Latar Belakang

Kemajuan pertelevision saat ini sangatlah pesat. Kemajuan tersebut ditunjukkan dengan banyak dan beragamnya informasi atau acara yang disajikan. Beragam informasi disajikan oleh semua stasiun televisi melalui berbagai acara yang menyebabkan semakin seringnya pemirsa memindah chanel TV untuk mencari acara yang sangat sesuai dengan keinginan. Berbagai macam solusi diberikan untuk memudahkan pemilihan acara sesuai dengan keinginan setiap pemirsa, salah satunya adalah memasang daftar acara setiap stasiun TV pada koran. Hal tersebut cukup membantu para pemirsa meskipun tidak memberikan rekomendasi setiap acara. Pemasangan jadwal acara di koran tidak memberitahukan secara detil suatu acara pada salah satu stasiun TV, termasuk didalamnya adalah genre acara, rating acara dan segmen usia penonton. Dengan adanya sistem rekomendasi terhadap acara televisi maka sistem tersebut akan mampu memberikan rekomendasi kepada pemirsa sesuai keinginan pemirsa.

Dalam data mining yang merupakan proses menemukan korelasi pola dan relasi menggunakan perangkat analisis data atau metode tertentu dapat untuk membuat prediksi yang valid. Prediksi ini bisa dimanfaatkan untuk membuat prediksi terhadap kesukaan orang dalam menonton acara televisi yang satu berarti juga suka menonton acara televisi yang lain. Sistem prediksi yang memberikan rekomendasi disebut juga dengan *recommender system*.

Recommender system (RSs) merupakan sistem berbasis agen yang memberikan rekomendasi kepada user dengan cara memberi informasi atau sugesti berdasarkan tingkat ketertarikan atau selera masing-masing user. *Recommender system* bekerja membantu pemirsa dalam memilih atau menentukan acara yang akan dilihat dengan cara memberi informasi atau sugesti berdasarkan tingkat ketertarikan atau selera masing-masing pemirsa. Ketertarikan atau selera pemirsa dapat diketahui dari beberapa hal, misalnya dari cara pemirsa mengakses halaman web, dari rating yang diberikan pemirsa terhadap suatu item produk, atau dari data transaksi pemirsa pada waktu lampau. Sedangkan, informasi yang bisa

diberikan kepada user dapat berupa ulasan mendetail tentang item atau acara, stasiun TV, dan lain-lain.

Secara umum, ada 2 macam metode yang sering digunakan dalam *Recommender system*, yaitu *content-based filtering* dan *collaborative filtering*. *Content-based filtering* bekerja dengan mencari item yang mempunyai korelasi dengan item yang disukai user berdasarkan *content*/ informasi tekstual tiap item. Sedangkan *collaborative filtering (CF)* merekomendasikan item kepada seorang user berdasarkan rating yang diberikannya terhadap item. *Collaborative filtering* lebih banyak digunakan karena dalam beberapa domain, seperti musik dan film, sulit dilakukan ekstraksi *content* tiap item yang merupakan langkah utama dari *content-based filtering*.

Salah satu metode dalam *collaborative filtering* adalah *Item based Top N-Recommendation*, yaitu model rekomendasi yang dibangun berdasarkan kedekatan item yang sering dipilih oleh pemirsa. Algoritma *Item-based Top N-Recommendation* ini bekerja dengan cara membuat daftar “top-N list” yang memuat N item yang paling banyak diminati atau dipilih oleh user, dimana N mewakili item yang akan dimuat dalam daftar yang paling banyak diminati atau dipilih. Dengan menggunakan *Top-N Recommendation* maka dapat ditentukan rating acara televisi yang paling banyak ditonton untuk kemudian direkomendasikan kepada.

Item based Top N-Recommendation melibatkan rating dari user terhadap setiap acara televisi (item) sehingga rekomendasi yang dihasilkan optimal berdasarkan penilaian user terhadap item tersebut. Dengan algoritma ini, akan dibuat 2 macam bentuk rekomendasi kepada user, yaitu rekomendasi *personalized* dan *non-personalized*. Berdasarkan latar belakang yang dipaparkan maka tugas akhir ini diberi judul “**Sistem rekomendasi acara televisi menggunakan algoritma *Item-based Top-N Recommendation*”**.

1.2 Rumusan Masalah

Rumusan masalah pada penelitian tugas akhir ini adalah :

1. Bagaimana merancang sebuah sistem rekomendasi untuk dapat memberikan rekomendasi acara televisi sesuai dengan personalisasi masing-masing user dalam memberikan rating terhadap item / acara televisi.

2. Bagaimana menghasilkan rekomendasi dengan waktu yang cepat dan memiliki *error* / penyimpangan yang kecil.

1.3 Tujuan Penelitian

Tujuan yang ingin dicapai dari penulisan Tugas Akhir ini antara lain:

1. Untuk melakukan perancangan sistem rekomendasi dalam menghasilkan rekomendasi acara televisi.
2. Melakukan implementasi dan uji coba sistem rekomendasi pada sebuah platform website dan menganalisa hasil percobaan.
3. Menghasilkan rekomendasi *personalized* acara televisi terhadap user yang pernah memberikan rating, dan rekomendasi *non-personalized* bagi user yang belum pernah memberikan rating.

1.4 Batasan Masalah

Batasan Masalah pada tugas akhir ini adalah sebagai berikut :

1. Hanya menggunakan metode *Item-based Top-N Recommendation*.
2. Perhitungan prediksi dilakukan per-genre item.
3. Sistem rekomendasi ini dibangun dengan data acara stasiun televisi nasional.
4. Skala rating yang digunakan adalah 1 s/d 5 dengan kenaikan 1.
5. Rekomendasi menampilkan acara televisi sesuai jadwal acara.

1.5 Sistematika Penulisan

Sistematika penulisan tugas akhir terbagi atas 5 bab, yaitu sebagai berikut:

BAB I PENDAHULUAN

Di dalam bab ini diuraikan mengenai latar belakang masalah yang akan dibahas, perumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, serta sistematika penulisan.

BAB II

TINJAUAN PUSTAKA

Pada bab ini dijelaskan tentang teori-teori yang mendasari sistem rekomendasi.

BAB III

METODE PENELITIAN

Bab ini menjelaskan metode sistem rekomendasi yang dipakai dalam membuat aplikasi ini.

BAB IV

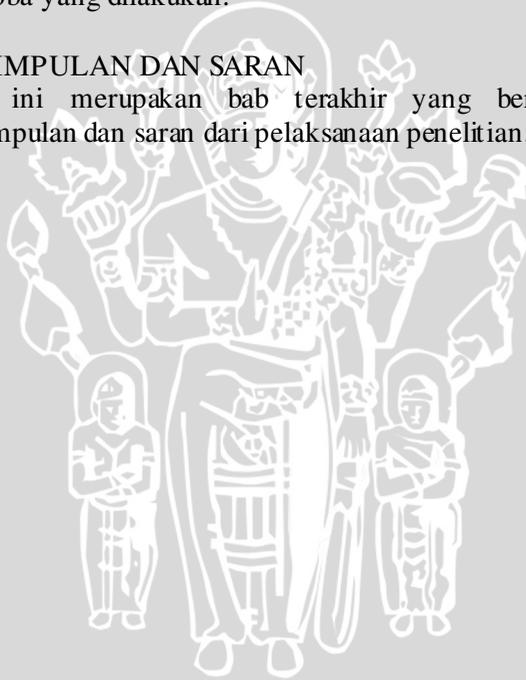
HASIL DAN PEMBAHASAN

Bab ini membahas implementasi rekomendasi dan uji coba yang dilakukan.

BAB V

KESIMPULAN DAN SARAN

Bab ini merupakan bab terakhir yang berisi kesimpulan dan saran dari pelaksanaan penelitian.



BAB II TINJAUAN PUSTAKA

2.1 Data Mining

Data mining adalah proses menemukan korelasi pola dan tren dengan mencarinya dalam tempat penyimpanan data menggunakan teknologi pengenalan pola atau juga teknik statistik dan matematis (Larose, 2005). Definisi lain dari data mining adalah suatu proses menggunakan berbagai perangkat analisis data untuk menemukan pola atau relasi dalam suatu data yang mungkin berguna dalam membuat prediksi yang valid (www.twocrows.com). Dari kedua pengertian di atas, maka data mining secara umum adalah proses menemukan korelasi pola dan relasi menggunakan berbagai perangkat analisis data untuk membuat prediksi yang valid.

Data mining juga populer dengan nama *knowledge discovery in database* (KDD). Istilah data mining dan *knowledge discovery in databases* (KDD) sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, akan tetapi berkaitan satu sama lain, dimana salah satu tahapan dalam keseluruhan proses KDD adalah data mining.

Beberapa teknik dalam data mining yang sering digunakan dalam memecahkan masalah antara lain (Larose, 2005) :

- *association rule*
- *clustering*
- *classification*

2.1.1 Association Rule

Association rule mining adalah suatu prosedur untuk menemukan aturan asosiatif antara suatu kombinasi item dalam suatu dataset yang ditentukan (Weiyang, 2000). Contoh dari aturan asosiatif dapat dilihat dari analisa pemesanan di sebuah restoran adalah bisa diketahui berapa besar kemungkinan seorang pengunjung memesan roti bersamaan dengan susu. Dengan pengetahuan tersebut pemilik restoran dapat mengatur penempatan menu makanan atau merancang strategi pemasaran untuk kombinasi

pemesanan tertentu. Metode ini disebut dengan *market basket analysis*.

Market basket analysis adalah teknik data mining dalam mengidentifikasi asosiasi antar item dalam sebuah *shopping basket* atau catatan transaksi penjualan (Pramudiono, 2003). Proses analisis ini sering digunakan untuk mengetahui item-item mana saja yang sering dibeli secara bersamaan sehingga hasil analisis yang didapat dapat digunakan untuk mengembangkan strategi pemasaran.

Dalam menentukan suatu *association rule*, terdapat suatu *interestingness measure* (ukuran kepercayaan) yang didapatkan dari hasil pengolahan data dengan perhitungan tertentu. Umumnya ada dua ukuran, yaitu *support* dan *confidence*. Ukuran ini nantinya berguna dalam menentukan *interesting association rules*, yaitu untuk dibandingkan dengan batasan (*threshold*) yang ditentukan oleh user. Batasan tersebut umumnya terdiri dari *minimum support* dan *minimum confidence*.

Algoritma *association rule* yang populer dan banyak digunakan adalah *Apriori* dengan paradigma *generate and test*, yaitu pembuatan kandidat kombinasi item yang mungkin berdasarkan aturan tertentu lalu diuji apakah kombinasi item tersebut memenuhi syarat *minimum support*. Kombinasi item yang memenuhi syarat tersebut disebut *frequent itemset*, yang nantinya dipakai untuk membuat aturan-aturan yang memenuhi syarat *minimum confidence*.

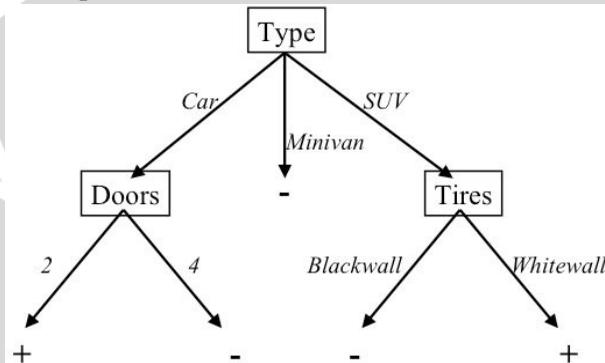
2.1.2 Clustering

Berbeda dengan *association rule* dimana kelas data telah ditentukan sebelumnya, *clustering* melakukan pengelompokan data tanpa berdasarkan kelas data tertentu (Pramudiono, 2003). Bahkan *clustering* dapat dipakai untuk memberikan label pada kelas data yang belum diketahui tersebut. Oleh karena itu, *clustering* sering digolongkan sebagai metode *unsupervised learning*.

2.1.3 Classification

Classification atau klasifikasi adalah proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui (Pramudiono, 2003). Model itu sendiri bisa berupa aturan "*if-then*", berupa *decision tree*, formula matematis, atau *neural network*.

Decision tree adalah salah satu metode klasifikasi yang paling populer karena mudah untuk diinterpretasi oleh manusia. Contoh dari *decision tree* dapat dilihat di Gambar 2.1. Disini setiap percabangan menyatakan kondisi yang harus dipenuhi dan tiap ujung pohon menyatakan kelas data. Contoh pada Gambar 2.1 adalah contoh pemilihan kendaraan mobil.

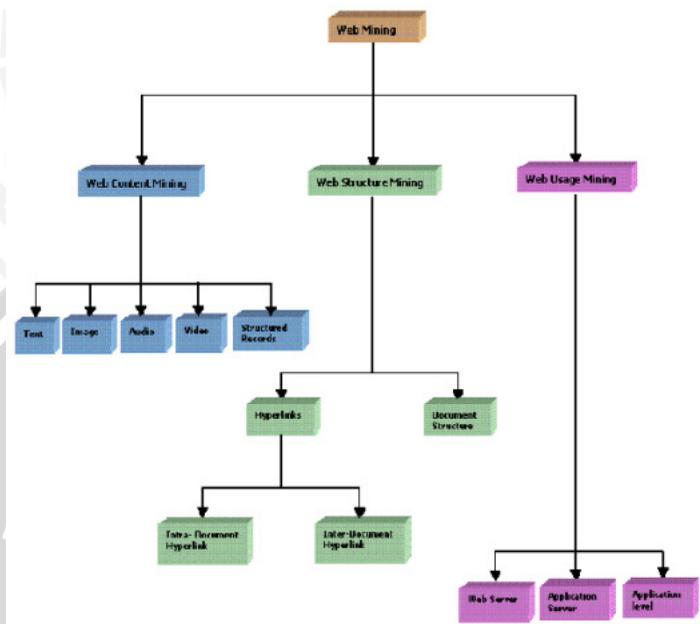


Gambar 2.1 Contoh decision tree sederhana
(sumber : <http://kur2003.if.itb.ac.id/file/pohon.pdf>)

2.2 Web Mining

Web adalah koleksi dari *inter-related files* dalam satu atau lebih web *servers*. Sedangkan web mining adalah aplikasi data mining teknik untuk diaplikasikan pada data web. Data web itu sendiri dapat berupa text, image, hyperlink, tags, atau logs (Srivastava, 2000).

Tujuan web mining adalah bagaimana membuat sistem yang bersifat personal bagi setiap user yang disebut dengan *personalized*. *Personalized* adalah sebuah sistem informasi yang bertujuan meningkatkan interaksi dalam sebuah website sesuai dengan keinginan user, berdasarkan informasi yang ada dalam item maupun informasi user.



Gambar 2.2T aksonomi web mining
(sumber gambar : Srivastava, J. 2000)

Berbagai data diolah dan dianalisa untuk menghasilkan sistem yang bersifat personal bagi user. Berdasarkan jenis data web yang dianalisa atau yang digali, web mining dapat dibagi menjadi 3 kategori (Eirinaki and Vazirgiannis, 2003), yaitu:

1. *Web Content Mining*

Ada 2 macam *web content mining* yaitu *web page content* yang mengekstrak informasi penting dari content (isi) dokumen web yang bisa berupa teks biasa atau halaman HTML/XML, dan *search result* yang melakukan analisa terhadap hasil pencarian.

2. *Web Structure Mining*

Web structure mining mengidentifikasi pengenalan graph atau merepresentasikan lubang graph web dengan metrics seperti *PageRank*.

3. *Web Usage Mining*

Web usage mining merekam kunjungan ke setiap halaman web dalam sebuah website, termasuk waktu kunjungan, alamat IP, dan

lain-lain. Data ini direkam seluruhnya oleh Web server logs, tetapi dapat juga diperoleh dari *cookies* atau tools *tracking session* lainnya. Sedangkan cara penyajian *personalized* terbagi dalam 4 bagian utama (Nasroui, 2002), yaitu:

1. *Memorization*

Dalam penyajian sederhana ini, informasi dari user seperti nama, dan *browsing history* disimpan (menggunakan *cookies*), yang kemudian digunakan untuk mengenali dan dikembalikan kembali kepada user.

2. *Customization*

Rekomendasi dilakukan berdasarkan input user yang diambil dari registrasi user dan selanjutnya digunakan untuk menentukan content dan struktur dari halaman web. Proses ini terkesan statis dan manual atau semi otomatis. Biasanya diimplementasikan dalam web server. Contoh nyata dapat dilihat pada web portal seperti My Yahoo!.

3. *Guidance or Recommender System*

Sistem berusaha untuk merekomendasikan *hyperlink* secara otomatis yang sesuai dengan keinginan user, dalam arti untuk memfasilitasi akses ke informasi yang dibutuhkan dalam website yang besar, hal ini biasanya diimplementasikan dalam web server, dan diterjemahkan dalam data yang mencerminkan keinginan user *implicit* (*browsing history* direkam dalam web server logs) atau *explicit* (*user profile* yang terdapat pada form registrasi atau polling). Penyajian ini merupakan cara yang efektif dalam perancangan web *personalized*.

4. *Task Performance Support*

Dalam sistem *client-side personalized*, asisten personal menganalisa tingkah laku user yang akan digunakan untuk memfasilitasi user dengan sistem. Penyajian ini mempunyai batasan, dimana tidak dapat menggunakan informasi user lain dengan kesamaan keinginan.

2.3 Sistem Informasi Personal (Personalized Information System)

Personalized Information System merupakan system informasi yang menyediakan dan melayani informasi berdasarkan ketertarikan pengguna, preferensi dan informasi diri pengguna yang dapat dioptimalkan selanjutnya, sehingga memberi hasil yang optimal untuk pengguna dan sesuai keinginan (J. V. Barneveld, 2003).

2.4 Recommender System

Tipe spesifik dari Sistem Informasi Personal adalah *recommender system*. *Recommender System* berdasar pada perpaduan dari *artificial intelligence*, *human computer interaction*, *sociology*, *information retrieval* dan website teknologi, memberikan sebuah rekomendasi yang berasal dari teman, kolega, publikasi dan sumber lainnya untuk membantu menentukan pilihan.

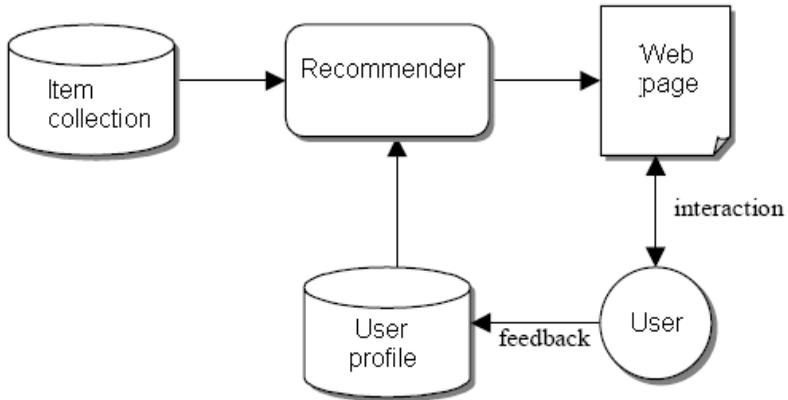
Recomender System (Sistem rekomendasi) adalah teknologi pemilihan informasi personal yang digunakan untuk memprediksi apakah seorang *user* tertentu akan tertarik dengan sesuatu (Han.dkk, 2005). Tujuan utama dari sistem rekomendasi adalah memperkirakan item atau bagian mana yang paling disukai atau yang paling berguna. Dengan melakukan tugas tersebut maka sistem ini bisa mengurangi informasi yang berlebihan yang membuat orang/pengguna bingung untuk menentukan pilihan.

Berbagai teknik data mining seperti *association rule* dan *market basket analysis* (Weiyang, 2000), hingga *clustering* telah digunakan dalam membangun sebuah RS. Namun, secara garis besar teknik yang digunakan dalam RS ada 2 macam, yakni *content-based filtering* dan *collaborative filtering* (Schafer, 2001).

2.4.1 Content-Based Filtering

Content-based filtering (CB) digunakan untuk menyediakan rekomendasi dengan membandingkan gambaran kesamaan item dengan item yang disukai oleh user berdasarkan *content/* isi tiap item yang ada (Nasroui, 2002). Item yang mempunyai kesamaan biasanya didasarkan pada *item attributes* (seperti artis, genre untuk item movie, penulis dan judul untuk item buku). Keuntungan menggunakan metode ini adalah tidak memperdulikan rating *implicit* atau *explicit* dari user atau data pembelian untuk

pembuatan rekomendasi. Skema untuk content-based filtering dapat dilihat pada Gambar 2.3.



Gambar 2.3 Arsitektur *Content-based filtering*

(sumber gambar : Nasroui, 2002)

2.4.2 Collaborative Filtering

Collaborative filtering (CF) adalah suatu metode dalam membuat sebuah prediksi otomatis untuk memperkirakan ketertarikan atau selera seorang user terhadap suatu item dengan cara mengumpulkan informasi dari user-user yang lain yang direpresentasikan dalam bentuk nilai rating (Sarwar, 2001). Secara umum, ada 2 proses yang dilakukan dalam CF, yaitu:

Mencari user lain yang mempunyai kemiripan pola rating dengan user target (user yang akan diberikan prediksi).

Menggunakan nilai rating dari user lain yang didapat dari langkah 1 di atas untuk menghitung prediksi bagi user target.

Sedangkan output yang dihasilkan yaitu prediksi dan rekomendasi (Sarwar, 2001).

Prediksi (P_{Ip}) adalah suatu nilai yang menyatakan prediksi besarnya rating item p yang mungkin didapat dari user I , dimana item p telah dirating oleh user lain.

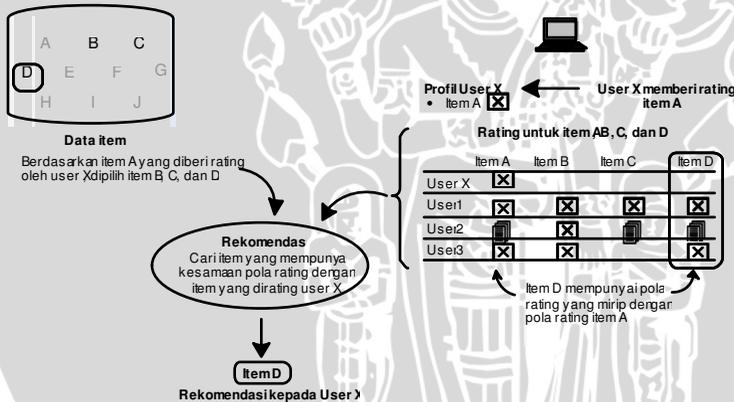
Rekomendasi adalah sebuah daftar yang berisi N item yang mempunyai kemungkinan terbesar untuk disukai oleh user aktif, dimana item-item tersebut pernah dirating oleh user lain.

Dengan menggunakan metode ini dimaksudkan informasi yang akan disaring (*filter*) dapat menampung grup user yang besar dan terdapat *individual application* untuk setiap grup user. Contoh nyata dari aplikasi menggunakan metode ini adalah *search service Google* yang memuat 50 bahasa dan terdapat *individual Google site* untuk negara terdekat di seluruh dunia.

Secara umum metode *collaborative filtering* dikelompokkan menjadi 2 jenis, yaitu:

2.4.2.1 Model-based atau item-based

Algoritma ini mencari hubungan antara item berdasarkan tabel rating untuk membuat rekomendasi. Sehingga metode ini disebut juga *item-based* (Masruri, 2007). Sebelum melakukan rekomendasi, algoritma ini melakukan korelasi terlebih dahulu untuk mengetahui hubungan antar item berdasarkan nilai rating yang ada dalam sebuah item. Proses pembuatan dapat dilakukan dengan berbagai teknik seperti *association rule*, *clustering*, atau *clasification*. Alur *model-based* atau *item-based* dapat dilihat pada Gambar 2.4.

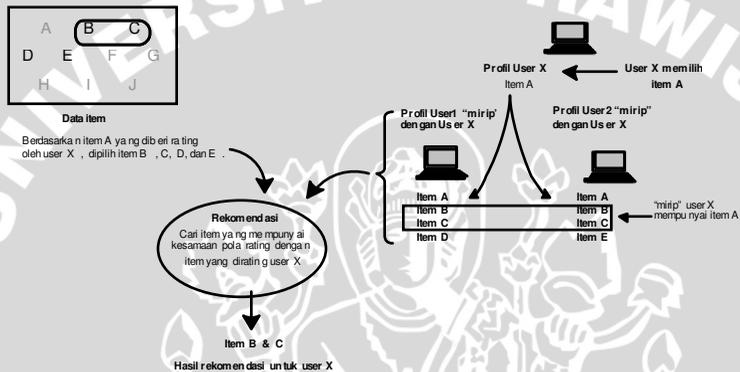


Gambar 2.4 Alur *model-based/ item-based* CF

(sumber : masruri, 2007)

2.4.2.2 Memory-based atau user-based

Berbeda dengan *item-based* yang mencari hubungan antara item, metode ini menggunakan seluruh data rating untuk melakukan prediksi. Metode ini menggunakan teknik statistik dalam mencari sekumpulan user, yang mirip dengan user target. Berdasarkan kemiripan tersebut, dibuat prediksi dan rekomendasi untuk user target. Skema *memory-based* atau *user-based* dapat dilihat pada Gambar 2.5.



Gambar 2.5 Alur *memory-based/ user-based* CF

(sumber : masruri, 2007)

2.4.2.3 Tingkat Personalisasi Rekomendasi

Tingkat personalisasi rekomendasi yang diberikan kepada user ada 2 macam, yaitu:

1. *Personalized* : rekomendasi yang diberikan kepada user tidak sama antara satu dengan yang lain, bergantung pada masing-masing profil user.
2. *Non-personalized*: bentuk rekomendasi ini tidak melihat profil masing-masing user, dengan kata lain rekomendasi bersifat umum sehingga dapat diberikan bagi pengunjung/ visitor.

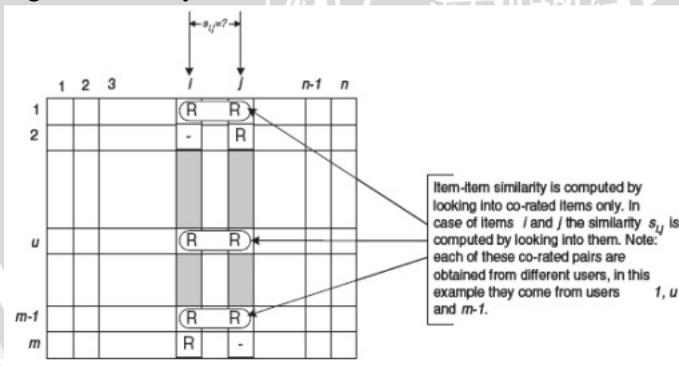
2.5 Algoritma Item-based Top-N Recommendation

Algoritma item-based top-N rekomendasi juga bisa disebut dengan model-based top-N rekomendasi. Algoritma ini menggunakan kedekatan antar item untuk dilakukan perhitungan hubungan antar item-item yang berbeda. Alasan utama dibalik adanya algoritma ini adalah bahwa faktanya customer(user) lebih senang untuk membeli item yang sama/dekat dengan item yang telah dia beli sebelumnya. Dengan melakukan analisa historis informasi pembelian (direpresentasikan dalam user-item matrik) maka akan dapat diketahui kedekatan antar item dan menggunakannya untuk top-N rekomendasi.

2.5.1 Menghitung kedekatan antar item (*Item-Item Similarity*)

Keberhasilan (efektif) dari penggunaan item-based top-N rekomendasi terletak pada metode yang digunakan untuk menghitung *similarity* antara berbagai macam item yang ada.

Salah satu metode yang bisa digunakan untuk penghitungan *similarity* adalah dengan menggunakan algoritma *cosine-based*. Algoritma ini merupakan termasuk dalam kategori penghitungan yang *symetric* yaitu dimana $sim(i,j) = sim(j,i)$. Cara kerja dari algoritma ini adalah dengan memperlakukan setiap item sebagai sebuah vektor pada customer/user tersebut, dan kemudian menggunakan fungsi *cosine* antara vektor-vektor tersebut untuk dihitung kedekatannya.



Gambar 2.6 Penghitungan *similarity* terhadap 2 item yang berkorelasi

(sumber : Deshpande, Mukhun & Karypis, George. 2004)

Secara umum jika R adalah $n \times m$ user-item matrik, maka *similarity* antara dua item i dan j dapat dihitung dengan rumus sbb :

$$\text{sim}(I_i, I_j) = \cos(\vec{R}_{*,i}, \vec{R}_{*,j}) = \frac{\vec{R}_{*,i} \cdot \vec{R}_{*,j}}{\|\vec{R}_{*,i}\|_2 \|\vec{R}_{*,j}\|_2} \quad (2.1)$$

Keterangan :

- Sim = similarity
- I_i = item i
- I_j = item j
- \vec{R} = matrix item

dimana ‘.’ Adalah vektor dengan operasi *dot-product*, sementara pembagiannya adalah operasi *cross-product*.

Dalam pembentukan model algoritma-nya untuk pemberian rekomendasi dapat dijelaskan sebagai berikut :

```

for j → 1 to m
  for i → 1 to m
    do {
      if i ≠ j
        then Mi,j → sim (R*,i, R*,j)
        else Mi,j → 0
    }
  for i → 1 to m
    do {
      if Mi,j ≠ among the k largest values in M*,j
        then Mi,j → 0
    }
return (M)

```

(2.2)

Keterangan :

m = user

$M_{i,j}$ = pasangan item i dan item j

dimana algoritma diatas menjelaskan apabila terdapat $n \times m$ user-item matrik R dan k adalah nilai dari perbandingan item *similarity* yang disimpan di masing-masing item, output yang dihasilkan direpresentasikan dengan $n \times m$ (matrik M). (Deshpande, Mukhun & Karypis, George. 2004)

2.5.2 Penghitungan Prediksi

Selain menghitung *similarity* langkah yang penting berikutnya adalah menghitung prediksi yang akan diberikan kepada user. Ketika *similarity* antar item telah diketahui, selanjutnya adalah melihat kepada target user (user yang akan diberi rekomendasi) dan menggunakan metode untuk menghitung prediksi sebuah item kepada user tersebut. Metode yang digunakan antara lain adalah *Weighted Sum*, yaitu metode yang digunakan untuk menghitung prediksi item i untuk user u dengan menghitung *sum* (jumlah) dari rating yang diberikan user u pada item *similar* pada i . Setiap rating di bobot dengan melihat korelasi *similarity* $S_{i,j}$ antara item i dan j . Dapat digambarkan dengan rumus sebagai berikut :

$$P_{u,i} = \frac{\sum_{all\ similar\ items, N} (S_{i,N} * R_{u,N})}{\sum_{all\ similar\ items, N} (S_{i,N})} \quad (2.3)$$

Keterangan :

$P_{u,i}$ = weighted sum

$S_{i,N}$ = nilai rating item dari user

$R_{u,N}$ = similariy item

\sum = jumlah (*sum*)

Pada dasarnya penghitungan prediksi ini mencoba menggambarkan bagaimana rating item dari aktif user ini serupa (*similar*) dengan item yang lain.

2.6 Pengukuran Tingkat Akurasi

Tingkat keakurasian rekomendasi yang dihasilkan dapat diukur dengan menggunakan salah satu metode standar dalam statistika yang disebut dengan *Mean Average Error* atau MAE (Sarwar, 2001). Secara mendasar, MAE menghitung kesalahan/*error* absolut antara rating yang sebenarnya (p) dan rating hasil prediksi (q), dimana semakin kecil nilai MAE yang didapat, maka prediksi yang dihasilkan semakin akurat. Jika pengukuran dilakukan terhadap N data, maka MAE dapat dirumuskan dengan:

$$MAE = \frac{\sum_{i=1}^N |p_i - q_i|}{N} \quad (2.4)$$

Keterangan :

- \sum = jumlah (*sum*)
- p_i = prediksi item
- q_i = nilai rating sebenarnya
- N = banyaknya rating

BAB III METODE PENELITIAN

3.1 Tahapan pengembangan sistem rekomendasi

Tahapan pengembangan sistem rekomendasi dengan menggunakan *item based top-n* rekomendasi ini adalah sebagai berikut :

1. Mempelajari konsep sistem rekomendasi dengan menggunakan *item based top-n recommendation*.
2. Mempelajari konsep penghitungan kedekatan (*similarity*) antar item dengan menggunakan algoritma *cosine-based*.
3. Mempelajari konsep penghitungan prediksi dengan menggunakan *weighted sum* untuk menghitung prediksi yang digunakan sebagai rekomendasi *personalized*.
4. Menganalisa dan merancang sistem rekomendasi pada lokal komputer berdasar pada konsep dan metode yang telah dipelajari.
5. Melakukan uji coba terhadap sistem rekomendasi yang telah dibuat dengan dataset yang ada.
6. Mengukur dan menganalisa apakah sistem rekomendasi yang dihasilkan berhasil dengan mengukur tingkat *error* yang dihasilkan.
7. Mengukur dan menganalisa apakah sistem rekomendasi yang dihasilkan berhasil dengan mengukur waktu komputasi yang dihasilkan.
8. Menganalisa dan menyimpulkan apakah sistem rekomendasi yang dihasilkan dapat diterapkan dan berhasil atau tidak untuk dijadikan sebuah sistem rekomendasi.

3.2 Kebutuhan sistem

Kebutuhan sistem untuk penelitian ini meliputi kebutuhan perangkat keras dan perangkat lunak

3.2.1 **Kebutuhan perangkat keras**

Kebutuhan perangkat keras yang digunakan oleh peneliti dalam melakukan penelitian ini adalah 2 (satu) buah komputer dengan spesifikasi sebagai berikut :

1. Komputer 1

Jenis : Notebook /Laptop
Nama : Fujitsu Lifebook P Series P5020
Prosesor : Intel Pentium Mobile 1GHz
Memory : 1 GB (Mobile Memory)

2. Komputer 2

Jenis : Server
Nama : Intel CPU E3110
Prosesor : Intel(R) Xeon(R) Dual Core CPU E3110
@ 3.00GHz
Memory : 1 GB (Server Memory)

3.2.2 **Kebutuhan perangkat lunak**

Kebutuhan perangkat lunak yang digunakan dalam melakukan penelitian ini adalah sebagai berikut :

- a. Webserver : Apache httpd (2.2.11)
- b. Database : MySQL (5.1.30)
- c. Database manajemen : phpMyadmin (3.1.1)
- d. Bahasa pemrograman : PHP (5.2.8)
- e. Browser : Google Chrome (2.0) / Mozilla Firefox (2.0) / Microsoft Internet Explorer (7.0)

Sedangkan untuk sistem operasi digunakan sebagai berikut :

1. Komputer 1

Sistem operasi : Microsoft Windows XP Professional with SP2

2. Komputer 2

Sistem operasi : Linux CentOS release 5.3 (Final)

Kebutuhan perangkat lunak untuk poin a s/d d tersebut diatas telah disediakan dalam satu paket software yaitu xampp (versi 1.7.0). Sedangkan untuk browser digunakan sebagai client untuk menampilkan aplikasi.

3.3 Perancangan Item-Based Top-N Recommendation

Pembuatan model item-graph dilakukan untuk mengetahui persamaan/kedekatan antara user-item matrik yang satu dengan user-item matrik yang lainnya. Metode yang digunakan adalah dengan menggunakan metode algoritma *Cosine-based /symmetric* (2.1). Misalkan diketahui data dari rating yang dilakukan oleh 3 user (w,x,y,z) terhadap item (a,b,c,d) dengan skala rating 1-5 seperti ditunjukkan pada tabel 3.1.

Tabel 3.1 Contoh *user-item* rating

Item \ User	A	B	C	D
W	1	3	5	3
X	5	null	3	4
Y	3	2	2	4
Z	?	4	?	3

Maka langkah-langkah yang dilakukan untuk mencari *similarity* dan prediksi dari user-item matrik diatas adalah sebagai berikut :

1. Dicari item-item yang telah dirating oleh user target (Z) yang ditunjukkan pada tabel 3.2

Tabel 3.2 Item dari user Z

Item	Rating
B	4
D	3

2. Untuk setiap item yang telah dirating oleh user target (Z) dicari item lain yang mempunyai korelasi, yaitu item pada user yang lain yang juga merating item B & D.

3. Setelah dilihat maka yang berkorelasi adalah pada item di user Y, yaitu :

- Terhadap item A ditunjukkan pada tabel 3.3

Tabel 3.3 Item- i user Z terhadap item A

Item_i	Item_j
B	A
D	A

- Terhadap item C ditunjukkan pada tabel 3.4

Tabel 3.4 Item-i user Z Terhadap item C

Item_i	Item_j
B	C
D	C

Item pada user X tidak berkorelasi, karena pada item B user X tidak merating item tersebut.

4. Dari tabel diatas maka di ketahui untuk item matriknya, yaitu :

- matrik item_i (item B) adalah $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$
- matrik item_i (item D) adalah $\begin{bmatrix} 3 \\ 4 \\ 4 \end{bmatrix}$
- matrik item_j (item A) adalah $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$
- matrik item_j (item C) adalah $\begin{bmatrix} 5 \\ 3 \\ 2 \end{bmatrix}$

5. Dari item matrik yang telah diketahui maka dicari kedekatan (*similarity*) terhadap item A & C, yaitu :

a. Terhadap item A ($B \rightarrow A$)

$$\text{sim}(I_B, I_A) = \cos(R_{*,B}, R_{*,A})$$

$$= \cos \begin{bmatrix} 3 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$= \frac{(3 \times 1) + (2 \times 3)}{\sqrt{(3^2 + 1^2) \times (2^2 + 3^2)}}$$

$$= \frac{9}{\sqrt{130}}$$

$$= \frac{9}{11,40175}$$

$$= 0,7894$$

b. Terhadap item A ($D \rightarrow A$)

$$\text{sim}(I_D, I_A) = \cos(R_{*,D}, R_{*,A})$$

$$= \cos \begin{bmatrix} 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

$$= \frac{(3 \times 1) + (4 \times 3)}{\sqrt{(3^2 + 1^2) \times (4^2 + 3^2)}}$$

$$= \frac{15}{\sqrt{250}}$$

$$= \frac{15}{15,811388}$$

$$= 0,9487$$

c. Terhadap item C ($B \rightarrow C$)

$$\begin{aligned}\text{sim}(I_B, I_C) &= \cos(R_{*,B}, R_{*,C}) \\ &= \cos \begin{bmatrix} 3 \\ 2 \end{bmatrix} \begin{bmatrix} 5 \\ 2 \end{bmatrix} \\ &= \frac{(3 \times 5) + (2 \times 2)}{\sqrt{(2^2 + 3^2) \times (2^2 + 5^2)}} \\ &= \frac{19}{\sqrt{377}} \\ &= \frac{19}{19,41649} \\ &= 0,9785\end{aligned}$$

d. Terhadap item C ($D \rightarrow C$)

$$\begin{aligned}\text{sim}(I_D, I_C) &= \cos(R_{*,D}, R_{*,C}) \\ &= \cos \begin{bmatrix} 3 \\ 4 \\ 4 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ 2 \end{bmatrix} \\ &= \frac{(3 \times 5) + (4 \times 3) + (4 \times 2)}{\sqrt{(3^2 + 4^2 + 4^2) \times (5^2 + 3^2 + 2^2)}} \\ &= \frac{35}{\sqrt{1558}} \\ &= \frac{35}{39,4715} \\ &= 0,8867\end{aligned}$$

6. Dari perhitungan pada langkah 5 diatas, maka hasil *similarity* item_i ke item_j seperti ditunjukkan tabel 3.5

Tabel 3.5 Hasil pemasangan item dan hasil *similarity*

item_i	item_j	Similarity
B	A	0,7894
D	A	0,9487
B	C	0,9785
D	C	0,8867

7. Setelah diketahui maka dihitung prediksinya dengan menggunakan *weighted sum* :

$$P_{z,a} = \frac{(0,7894 \times 4) + (0,9487 \times 3)}{0,7894 + 0,9487}$$

$$= \frac{6,0037}{1,7381} = 3,4542$$

$$P_{z,c} = \frac{(0,9785 \times 4) + (0,8867 \times 3)}{0,9785 + 0,8867}$$

$$= \frac{6,5741}{1,8652} = 3,5246$$

penghitungan prediksi dengan menggunakan *weighted sum* ini menghasilkan seperti yang ditampilkan pada tabel 3.6.

Tabel 3.6 Prediksi item terhadap user Z

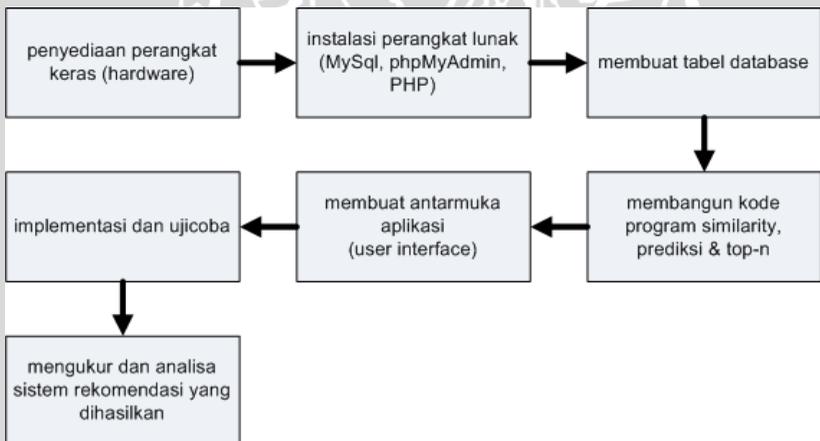
nourut	user	item_j	prediksi
1	Z	C	3,5246
2	Z	A	3,4542

dari data pada tabel 3.6 maka bisa diketahui bahwa item C diprediksikan kepada user Z daripada item A, karena item C mempunyai nilai lebih dekat/besar 3,5246 dibandingkan dengan item A yang mempunyai nilai 3,4542.

3.4 Proses Penelitian

Tahapan proses penelitian yang dilakukan adalah sebagai berikut :

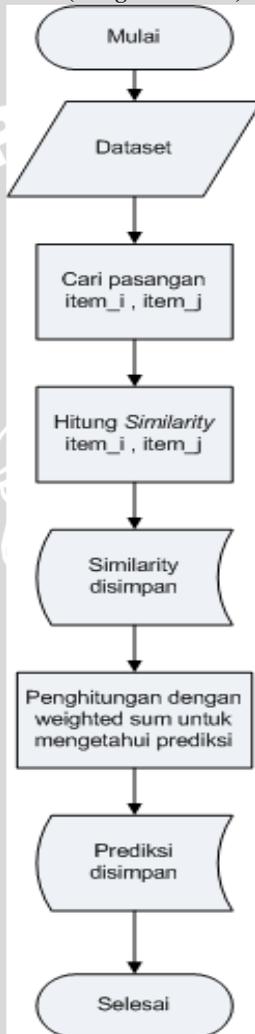
1. Menyediakan perangkat keras (komputer) yang dibutuhkan dan disertai dengan sistem operasi-nya.
2. Melakukan instalasi perangkat lunak xampp for windows (versi 1.7.0) yang didalamnya termasuk juga melakukan instalasi :
 - a. MySQL (5.1.30)
 - b. phpMyadmin (3.1.1)
 - c. PHP (5.2.8)
3. Membuat tabel database pada phpMyadmin
4. Membuat penghitungan *similarity* antar item
5. Membuat penghitungan prediksi
6. Membuat antar muka aplikasi (user interface)
7. Implementasi dataset dan ujicoba pada browser
8. Mengukur dan menganalisa sistem rekomendasi yang telah dihasilkan.



Gambar 3.1 Blok diagram proses penelitian

3.5 Alur proses Model Item-Based Top-N Recommendation

Berikut adalah alur dari proses pembuatan model data rekomendasi item-based top-N dengan menggunakan similarity (*cosine-based*) dan prediksi (*weighted sum*).



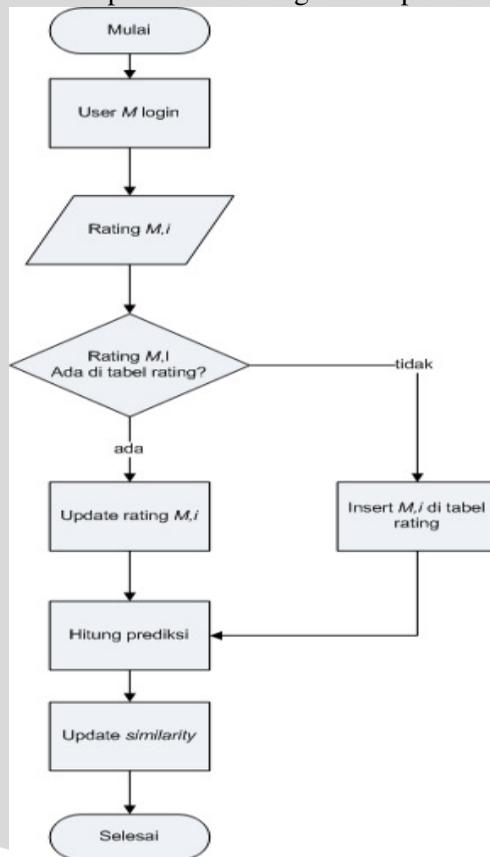
Gambar 3.2 Alur proses pembuatan model data

Penjelasan dari alur diatas adalah sebagai berikut :

1. Data acara televisi , rating, dan user
2. data acara televisi tersebut dicari pasangan item i dan j
3. Dilakukan penghitungan *similarity*, kemudian dilakukan penyimpanan pada *stored data (database)*
4. Dari hasil *similarity*, dilakukan penghitungan prediksi untuk setiap item terhadap setiap user. Kemudian disimpan dalam *stored data (database)*

3.6 Alur proses pemberian rating oleh user

Berikut adalah pemberian rating terhadap item oleh user :



Gambar 3.3 Alur proses sistem pemberian rating

Penjelasan dari alur diatas adalah sebagai berikut :

1. User melakukan login dengan user_id dan password-nya.
2. Kemudian user akan disediakan tampilan beberapa acara televisi yang disediakan untuk diberi rating.
3. Dilakukan pengecekan apakah item tersebut pernah di rating oleh user tersebut sebelumnya.
4. Apabila ada/pernah dirating sebelumnya maka dilakukan update rating di tabel_rating.
5. Apabila belum pernah dirating, maka merupakan item baru yang dirating dan dilakukan insert pada tabel_rating.
6. Melakukan penghitungan *similarity* terhadap item yang dirating oleh user dengan item-item lainnya yang dirating oleh user lainnya
7. Melakukan penghitungan prediksi dari user tersebut terhadap item lain.
8. Selesai.

3.7 Rekomendasi Non-Personalized

Rekomendasi *non-personalized* adalah rekomendasi yang dibuat tidak berdasarkan atas preferensi user dalam merating item. Kegunaan rekomendasi *non-personalized* adalah untuk memberikan rekomendasi kepada user yang tidak terdaftar (pengunjung) untuk tetap bisa mendapatkan rekomendasi yang akurat. Rekomendasi *non-personalized* seperti ini dapat diilustrasikan dengan: “user yang menyukai item I juga menyukai item J dan K”.

Untuk dapat memberikan rekomendasi *non-personalized* digunakan dengan memilih dari data *similarity* dimana item *i* yang sedang dilihat oleh user akan dicari *similarity* terdekat dengan item *j* kemudian merekomendasikan item *j* tersebut.

Dapat dicontohkan dari tabel *similarity* diatas dimana apabila dicari item *i* yang dikunjungi dengan memasangkan terhadap item *j*, maka akan diketahui seperti pada tabel 3.7

Tabel 3.7 *Similarity* untuk menghitung prediksi terhadap user Z

item_i	item_j	Similarity
B	C	0,9785
B	A	0,7894
D	A	0,9487
D	C	0,8867

rekomendasi yang diberikanurut berdasarkan kedekatan *similarity* nya, yaitu semakin mendekati 1 maka lebih utama untuk direkomendasikan. Dengan kata lain untuk *similarity* diurutkan dari besar ke kecil (*descending*).

Sehingga untuk pengunjung yang melihat item B maka akan direkomendasikan item C kemudian item A, sedangkan untuk pengunjung yang melihat item D maka akan direkomendasikan item A kemudian item C.

3.8 Rekomendasi Personalized

Rekomendasi *personalized* diberikan kepada user yang terdaftar dan telah memberikan rating item dalam sistem. Setelah diketahui korelasi antar user dari data model *similarity* yang telah terbentuk, dapat dilakukan perhitungan prediksi untuk selanjutnya dari hasil prediksi tersebut dibuat sebagai daftar rekomendasi bagi user target.

Rekomendasi *personalized* ini diberikan dengan cara menghitung *weighted sum* (persamaan 2.3) dari *similarity* yang telah ada. Hasil dari prediksi tersebut seperti ditunjukkan pada tabel 3.8.

Tabel 3.8 Hasil prediksi terhadap user target

nourut	user	item_j	prediksi
1	Z	C	3,5246
2	Z	A	3,4542

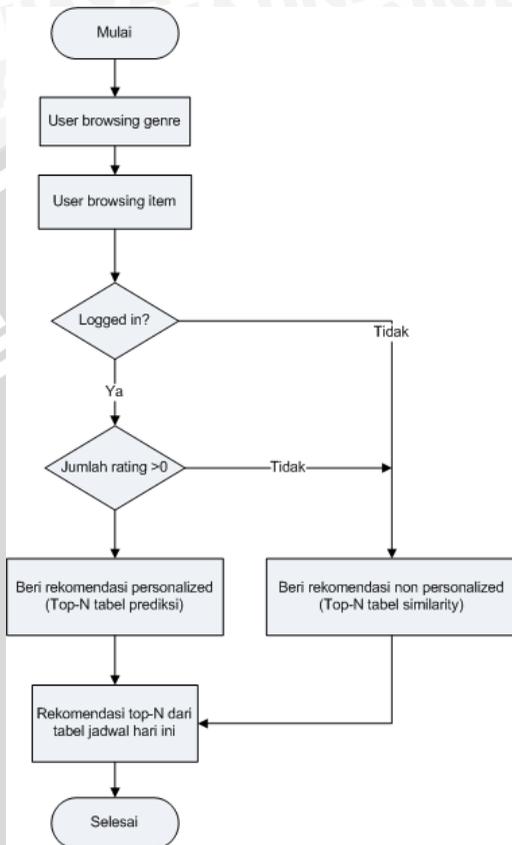
3.9 Rekomendasi Top-N

Rekomendasi yang diberikan kepada user adalah Top-N, dimana rekomendasi ini mengambil N item yang nilainya paling mendekati / paling tinggi. Dalam hal ini penulis mengambil N sebagai 5 teratas dengan nilai *similarity* yang paling mendekati yaitu *similarity* yang paling dekat dengan nilai 1. Atau untuk rekomendasi *personalized* diambil dari tabel prediksi.

3.10 Alur Pemberian Rekomendasi

Alur data proses pemberian rekomendasi sistem terhadap user ditunjukkan pada gambar 3.4





Gambar 3.4 Alur pemberian rekomendasi

Pada saat user memilih item dalam sebuah genre, diperiksa apakah user telah terdaftar atau belum. Jika user telah terdaftar, maka perlu diperiksa lagi apakah user telah memberi paling tidak sebuah rating pada sebuah item lain pada genre yang sama. Jika kondisi benar, user akan diberi rekomendasi *personalized*. Namun jika belum, user akan diberi rekomendasi *non-personalized*. Bagi user yang tidak terdaftar (tidak login), akan selalu diberikan rekomendasi *non-personalized*. Dari rekomendasi yang diberikan, akan dilihat pada jadwal acara televisi. Apabila acara tersebut tidak memiliki jadwal maka acara tidak direkomendasikan.

3.11 Jadwal acara televisi

Pada setiap item acara televisi tentunya memiliki jadwal acara yang terdiri dari hari, tanggal, dan jam tayang. Hal ini juga merupakan elemen yang penting dalam pemberian rekomendasi, dimana acara yang telah lewat masa tayangnya tidak perlu untuk diberikan rekomendasi. Maka disini akan dilihat pada jadwal acara dimana item yang direkomendasikan adalah item yang memiliki jadwal acara pada hari sekarang. Untuk user Z diatas, item C & A dapat ditunjukkan pada tabel 3.9.

Tabel 3.9 Jadwal item-j pada user Z

Item	Jadwal tayang
C	1 januari 2009 08:00
C	2 januari 2009 08:00
A	2 januari 2009 08:00
A	3 januari 2009 08:00

Sedangkan untuk prediksi nya ditunjukkan pada tabel 3.10.

Tabel 3.10 Prediksi yang diberikan kepada user Z

nourut	user	item_j	prediksi	Tgl tayang terdekat
1	Z	A	3,4542	3 januari 2009 08:00
2	Z	C	3,5246	2 januari 2009 08:00

Jadi apabila diketahui semisal hari ini adalah tanggal 3 januari 2009, maka untuk user Z item yang direkomendasikan adalah item A. Karena item C telah lewat tgl tayangnya.

3.12 Perancangan Tabel

Perancangan tabel ini dibutuhkan untuk menyimpan data dalam sebuah database sebagai data induk yang digunakan untuk penelitian.

Total Tabel

Berikut adalah total tabel beserta record yang digunakan, data type yang digunakan adalah MyISAM dimana data type ini dikoneksikan secara manual.

Tabel 3.11 Total tabel database

No	Tabel	Records	Type
1	tabel_genre	11	MyISAM
2	tabel_item	110	MyISAM
3	tabel_jadwal	3.357	MyISAM
4	tabel_prediksi	11.000	MyISAM
5	tabel_rating	11.000	MyISAM
6	tabel_similarity	11.990	MyISAM
7	tabel_user	100	MyISAM

Tabel item / acara

Berikut adalah tabel item/ acara yang digunakan untuk menyimpan acara televisi

Tabel 3.12 Tabel item / acara

Field	Type	KEY
item_id	int(11)	PRIMARY
nama_acara	char(30)	
stasiun_tv	char(10)	
genre_id	int(11)	

Keterangan :

- item_id : kode item/ acara
- nama_acara : nama acara
- stasiun_tv : nama stasiun televisi
- genre_id : kode genre

Tabel genre

Tabel genre untuk menyimpan genre-genre yang digunakan untuk mengkategorikan acara.

Tabel 3.13 Tabel genre

Field	Type	KEY
genre_id	int(11)	PRIMARY
nama	varchar(50)	

Keterangan :

genre_id : kode genre

nama_genre : nama genre

Tabel user

Tabel user digunakan untuk menyimpan data-data user

Tabel 3.14 Tabel user

Field	Type	KEY
user_id	int(11)	PRIMARY
username	varchar(255)	PRIMARY
nama	varchar(100)	
kelamin	char(1)	
tgl_lahir	date	
passwd	varchar(255)	

Keterangan

user_id : kode user

username : username untuk login

nama : nama user

kelamin : jenis kelamin user

tgl_lahir : tanggal lahir user

passwd : password untuk login

Tabel rating

Tabel rating digunakan untuk menyimpan rating yang diberikan oleh user terhadap item

Tabel 3.15 Tabel rating

Field	Type	KEY
user_id	int(11)	PRIMARY
item_id	int(11)	PRIMARY
genre_id	int(11)	
rating	int(1)	

Keterangan :

user_id : kode user
item_id : kode item
genre_id : kode genre
rating : angka rating

Tabel similarity

Tabel similarity digunakan untuk menyimpan data hasil penghitungan kedekatan item_i dan item_j

Tabel 3.16 Tabel *similarity*

Field	Type	KEY
item_i	int(10)	PRIMARY
item_j	int(10)	PRIMARY
genre_id	int(11)	
similarity	double	

Keterangan

item_i : item ke -i yang dirating oleh user
item_j : item ke -j yang dihitung kedekatannya terhadap item -i
genre_id : kode genre acara
similarity : angka *similarity*

Tabel prediksi

Tabel prediksi digunakan untuk menyimpan hasil penghitungan *weighted sum* (prediksi).

Tabel 3.17 Tabel prediksi

Field	Type	KEY
user_id	int(11)	PRIMARY
item_j	int(11)	PRIMARY
prediksi	double	

Keterangan :

user_id : kode user

item_j : item /acara yang di prediksi

prediksi : angka prediksi

Tabel jadwal

Tabel jadwal digunakan untuk menyimpan jadwal acara televisi.

Tabel 3.18 Tabel jadwal

Field	Type	KEY
item_id	int(11)	PRIMARY
wkt_tayang	datetime	PRIMARY
episode	text	
sinopsis	text	

Keterangan :

item_id : kode item

wkt_tayang : waktu tayang

episode : nama episode

sinopsis : keterangan singkat tentang acara

3.13 Dataset acara televisi

Dataset ini adalah data-data yang diperlukan untuk *database* sebagai sebuah data awal sistem rekomendasi. Data diambil dari acara televisi di Indonesia yang terdapat dalam website masing-masing stasiun televisi. Kemudian dikelompokkan genre / genre dari acara-acara televisi tersebut dengan mengacu pada genre yang sudah ada (www.rcti.tv)

Untuk sebuah sistem rekomendasi maka diperlukan adanya pemberian rating. Dengan memberikan rating maka user tersebut dianggap telah melihat acara televisi tersebut. Nilai untuk rating yang diberikan oleh user adalah 1,2,3,4, & 5, sementara untuk user yang tidak memberikan rating terhadap item tersebut maka nilai ratingnya adalah *NULL*.

Untuk menguji sistem rekomendasi ini digunakan dataset dari *movielens* (www.grouplens.org) dengan mengambil jumlah rating sebanyak 10.000 rating dan jumlah user sebanyak 567 user, jumlah item sebanyak 110 item.

3.14 Rancangan Hasil Penelitian

Jika seorang user belum merating satu pun item dalam sebuah genre, maka tidak memperoleh rekomendasi *personalized*. Hal ini disebabkan karena sistem tidak dapat mengetahui preferensi rating dari user yang bersangkutan sehingga tidak ada input pada rating untuk diikutsertakan dalam perhitungan prediksi.

Pengujian dilakukan untuk mengetahui nilai keakuratan prediksi dan waktu komputasi dari metode *item-based* yang digunakan. Untuk mengukur keakuratan prediksi, digunakan rumus *mean absolute error* /MAE (2.3).

Tabel 3.19 Rancangan hasil *Mean Absolute Error*

No	Rasio Data Set	Jumlah Prediksi	MAE
1
2
...
n
Rata-rata		

Dari data pada tabel 3.19 maka dapat diketahui rata-rata *error* yang ditimbulkan sehingga mengalami penyimpangan berapa persen, apabila prosentase penyimpangan sangat kecil maka sistem rekomendasi yang dihasilkan telah berhasil. Demikian pula apabila *error* yang dihasilkan semakin kecil maka semakin baik rekomendasi yang dihasilkan.

Sedangkan untuk mengetahui waktu komputasi, dilakukan x kali percobaan dengan menampilkan prediksi rating item bagi m user yang berbeda. M user diambil secara random dan dicatat berapa waktu yang diperlukan untuk menampilkan keseluruhan prediksi berdasarkan genre dan item yang sedang dilihat/ dipilih.

Tabel 3.20 Racangan tabel hasil komputasi rekomendasi

Uji Coba	UserID	ItemID yang Dilihat	Waktu 1 (detik)	Waktu 2 (detik)
1
2
...
N
Rata-rata		

Dari data tabel 3.20 maka dapat diketahui bahwa sistem rekomendasi yang dibuat dapat menghasilkan rekomendasi bagi user dalam waktu x detik. Waktu 1 menunjukkan komputasi rekomendasi untuk item dikategorikan genre, dan waktu 2 untuk item yang tidak dikategorikan genre.

3.15 Rancangan tampilan antar muka (*user interface*)

Berikut adalah rancangan tampilan yang digunakan dalam sistem rekomendasi ini :

3.15.1 Home

Home Login Daftar Acara	Halaman Home Selamat Datang ! Di SistemRekomendasi Acara Televisi (ItemBased Top-N Recommendation)
-------------------------------	--

Gambar 3.5 Rancangan halaman home /beranda awal

3.15.2 Login

Home Login Daftar Acara	Halaman Login Username : Password : LOGIN (daftar baru) → <div style="border: 1px solid black; padding: 2px; display: inline-block;">mendaftar sebagai pengguna baru</div>
-------------------------------	--

Gambar 3.6 Rancangan halaman login

3.15.3 Daftar User Baru

Home Login Daftar Acara	Halaman Pendaftaran User Baru Nama : _____ Kelamin : _____ Tgl Lahir : _____ Username : _____ Password : _____ DAFTAR
-------------------------------	---

Gambar 3.7 Rancangan halaman pendaftaran user baru

3.15.4 Daftar Genre

Home Login Daftar Acara	Halaman Daftar Acara Genre Acara : Sinetron Berita Musik Program anak
-------------------------------	--

Gambar 3.8 Rancangan halaman daftar genre

3.15.5 Genre Acara per Genre

Home Login Daftar Acara	Daftar Acara >> Sinetron				
	No	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam
	1	Sinetron 1	TV 1	Tgl 1	08:00
	2	Sinetron 2	TV 2	Tgl 2	09:00
	3	Sinetron 3	TV 3	Tgl 3	10:00
	4	Sinetron 4	TV 4	Tgl 4	11:00
	5	Sinetron 5	TV 5	Tgl 5	12:00

Gambar 3.9 Rancangan halaman daftar acara

3.15.6 Detail Acara Bagi User yang Tidak Melakukan Login

Home Login Daftar Acara	Daftar Acara >> Sinetron >> Nama Acara					
	Nama Acara : Melati Untuk Marvell					
	Stasiun TV : TV1					
	Hari / Tanggal : Selasa, 12 Mei 2009					
	Jam : 18:00					
	Sinopsis : Keterangan singkat acara					
	Rekomendasi :					
No	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam	Prediksi	
1	Sinetron 1	TV 1	Tgl 1	08:00	0.9	
2	Sinetron 2	TV 2	Tgl 2	09:00	0.8	
3	Sinetron 3	TV 3	Tgl 3	10:00	0.7	
4	Sinetron 4	TV 4	Tgl 4	11:00	0.6	
5	Sinetron 5	TV 5	Tgl 5	12:00	0.5	

top-n rekomendasi
non-personalized

Gambar 3.10 Rancangan detail nama acara dan pemberian rekomendasi non-personalized

3.15.7 Detail Acara Bagi User yang Melakukan Login

Home Daftar Acara >> Sinetron >> Nama Acara

Profil Anda

Rating Anda Nama Acara : Melati Untuk Marvell

Daftar Stasiun TV : TV1

Acara Hari / Tanggal : Selasa, 12 Mei 2009

Logout Jam : 18:00

 Sinopsis : Keterangan singkat acara

Rating Anda : 1 2 3 4 5 |RATE| → user memberikan rating

Rekomendasi :

No	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam	Rating
1	Sinetron 1	TV 1	Tgl 1	08:00	0.9
2	Sinetron 2	TV 2	Tgl 2	09:00	0.8
3	Sinetron 3	TV 3	Tgl 3	10:00	0.7
4	Sinetron 4	TV 4	Tgl 4	11:00	0.6
5	Sinetron 5	TV 5	Tgl 5	12:00	0.5

top-n rekomendasi *personalized*

Gambar 3.11 Rancangan detail nama acara, pemberian rating dan pemberian rekomendasi personalized

3.15.8 Halaman Profil User

Home	Halaman Profil Nama : User1 Kelamin : Laki-Laki Tgl Lahir : 13 Pebruari 1985
Profil Anda	
Rating Anda	
Daftar Acara	
Logout	

Gambar 3.12 Rancangan halaman profil user

3.15.9 Halaman Rating User

Home Profil Anda Rating Anda Daftar Acara Logout	Halaman Rating Anda			
	No	Nama Acara	Stasiun TV	Rating Anda
	1	Acara 1	TV 1	5
	2	Acara 2	TV 2	3
	3	Acara 3	TV 3	4
	4	Acara 4	TV 4	1
5	Acara 5	TV 5	2	

Gambar 3.13 Rancangan halaman daftar rating yang telah diberikan oleh user

BAB IV HASIL DAN PEMBAHASAN

4.1 Hasil

Hasil dari penelitian ini akan menampilkan hasil dari persiapan data, penghitungan model data, rekomendasi yang dihasilkan, waktu yang dihasilkan dan rata-rata error atau penyimpangan dari prediksi yang dihasilkan.

4.1.1 Persiapan Data

Untuk mengetahui bagaimana kinerja *recommender system*, dibutuhkan sebuah *data awal* rating yang telah terisi dengan *record* rating item dari user ($n \times m$). Data yang digunakan adalah sesuai dengan yang telah disampaikan pada bab 3.13.

Tabel 4.1 Rekap jumlah data untuk membentuk model

Jumlah user	567
Jumlah item acara	110
Banyaknya rating	10000
Rata-rata rating per item	3,7287
Rata-rata banyaknya rating per user	17,63
Rata-rata banyaknya rating per item	90,90
Banyaknya jadwal acara	3357

4.1.2 Implementasi Basis Data

Tabel utama yang dirancang untuk pembuatan rekomendasi berbasis metode *item-based top-N recommendation* adalah yaitu tabel *tabel_rating*, *tabel_similarity* dan *tabel_prediksi*. Tabel *rating* menyimpan data rating user terhadap item, sedangkan tabel *similarity* untuk menyimpan data korelasi antar item (*item_i* dan *item_j*) atau data model item yang nantinya akan digunakan dalam perhitungan prediksi. Dan *tabel_prediksi* menyimpan data prediksi *item_j* kepada user target.

Tabel *rating* digunakan sebagai dasar pembuatan model item pada metode *item-based*. Cuplikan kode dalam bentuk query di

bawah ini bertujuan membuat tabel rating beserta *field* yang dibutuhkan.

1	CREATE TABLE `tabel_rating` (
2	`user_id` int(11) NOT NULL,
3	`item_id` int(11) NOT NULL,
4	`genre_id` int(11) NOT NULL,
5	`rating` int(1) unsigned DEFAULT NULL)

Berdasarkan skema penghitungan menggunakan *cosine-based* yang sudah dibahas, maka dibuat tabel *similarity*. Dengan demikian, setelah nilai korelasi antara 2 item yang berpasangan dihitung dan diketahui, maka nilai *similarity* nya disimpan pada tabel ini. Query untuk membuat tabel *similarity* adalah sebagai berikut.

1	CREATE TABLE `tabel_similarity` (
2	`item_i` int(10) unsigned DEFAULT NULL,
3	`item_j` int(10) unsigned DEFAULT NULL,
4	`genre_id` int(11) NOT NULL,
5	`similarity` double DEFAULT NULL)

Setelah tabel *similarity* terbentuk berikutnya adalah membuat tabel *prediksi* untuk menyimpan data penghitungan prediksi dengan menggunakan metode *weighted sum*. Query untuk membuat tabel prediksi adalah sebagai berikut.

1	CREATE TABLE `tabel_prediksi` (
2	`user_id` int(11) DEFAULT NULL,
3	`item_j` int(11) DEFAULT NULL,
4	`prediksi` double DEFAULT NULL)

4.1.3 Deskripsi Program

Dalam penulisan ini akan dijelaskan deskripsi program yang meliputi fungsi-fungsi utama dalam pembangunan sistem rekomendasi dengan menggunakan *item-based top-N*.

4.1.3.1 Fungsi penghitungan *similarity*

Berikut adalah fungsi penghitungan *similarity*, yaitu fungsi dalam menghitung kedekatan antara item_i dan item_j.

```
1 $sql = "  
2 SELECT (  
3     SELECT SUM( a.rating * b.rating )  
4     FROM tabel_rating a,tabel_rating b  
5     WHERE a.user_id = b.user_id  
6           AND a.item_id = ". $this->Db->quote($itemI) . "  
7           AND b.item_id = ". $this->Db->quote($itemJ) . " ) / (  
8     SELECT sqrt( (  
9     SELECT sum( pow( rating ,2 ) )  
10    FROM tabel_rating  
11    WHERE item_id = ". $this->Db->quote($itemI) . " and user_id in (  
12    SELECT user_id  
13    FROM tabel_rating  
14    WHERE item_id IN ( ". $this->Db->quote($itemI) . " , ". $this->Db->  
15    >quote($itemJ) . " )  
16    GROUP BY user_id  
17    HAVING count( * ) = 2 ) ) * (  
18    SELECT sum( pow( rating ,2 ) ) )  
19    FROM tabel_rating  
20    WHERE item_id = ". $this->Db->quote($itemJ) . " and user_id in (  
21    SELECT user_id  
22    FROM tabel_rating  
23    WHERE item_id IN ( ". $this->Db->quote($itemI) . " , ". $this->  
24    >Db->quote($itemJ) . " )  
25    GROUP BY user_id  
26    HAVING count( * ) = 2 ) ) );
```

4.1.3.2 Fungsi hitung Wighted SUM

Berikut adalah fungsi penghitungan weighted sum dalam memprediksi item/ acara kepada user.

```
1 $sql = "  
2     SELECT (SELECT SUM(a.rating * b.similarity)  
3     FROM tabel_rating a, tabel_similarity b  
4     WHERE b.item_j = ". $this->Db->quote($itemId) . "  
5     AND a.item_id = b.item_i  
6     AND a.item_id IN (  
7     SELECT item_id  
8     FROM tabel_item  
9     WHERE genre_id = ". $this->Db->quote($genreId) . "  
10    AND item_id <> ". $this->Db->quote($itemId) . "  
11    AND user_id = ". $this->Db->quote($userId) . ") /  
12 (SELECT SUM( similarity )  
13 FROM tabel_rating a, tabel_similarity b  
14 WHERE b.item_j = ". $this->Db->quote($itemId) . "  
15 AND a.item_id = b.item_i  
16 AND a.item_id IN (  
17 SELECT item_id  
18 FROM tabel_item  
19 WHERE genre_id = ". $this->Db->quote($genreId) . "  
20 AND item_id <> ". $this->Db->quote($itemId) . "  
21 AND user_id = ". $this->Db->quote($userId) . ")");
```

4.1.3.3 Fungsi menampilkan rekomendasi personalized

Berikut adalah fungsi untuk mengambil data pada tabel prediksi yang telah dilakukan penghitungannya untuk ditampilkan kepada user yang melakukan login, sehingga mendapat rekomendasi *personalized*.

```
1 $sql = "  
2 SELECT b.item_id, c.nama_acara, c.stasiun_tv, b.episode, b.sinopsis,  
3 a.prediksi,  
4 b.wkt_tayang, DAYOFWEEK(b.wkt_tayang) hari_ke,  
5 TIME_FORMAT(b.wkt_tayang, '%H:%i') jam  
6 FROM tabel_prediksi a, tabel_jadwal b, tabel_itemc  
7 WHERE a.item_j=b.item_id  
8 AND a.item_j=c.item_id  
9 AND DATE(b.wkt_tayang) = CURRENT_DATE  
10 AND a.user_id = ". $this->Db->quote($userId) . "  
11 AND c.genre_id = ". $this->Db->quote($genreId) . "  
12 ORDER BY a.prediksi DESC" . ($limit ? " LIMIT " . (int)$limit : ");
```

4.1.3.4 Fungsi menampilkan rekomendasi non-personalized

Berikut adalah fungsi untuk mengambil data pada tabel similarity yang telah dilakukan penghitungannya untuk ditampilkan kepada user yang tidak melakukan login, sehingga mendapatkan rekomendasi *non-personalized*.

```
1 $sql = "  
2 SELECT b.item_id, c.nama_acara, c.stasiun_tv, b.episode, b.sinopsis,  
3 a.similarity prediksi,  
4 b.wkt_tayang, DAYOFWEEK(b.wkt_tayang) hari_ke,  
5 TIME_FORMAT(b.wkt_tayang, '%H:%i') jam  
6 FROM tabel_similarity a, tabel_jadwal b, tabel_itemc  
7 WHERE a.item_j=b.item_id  
8 AND a.item_j=c.item_id  
9 AND DATE(b.wkt_tayang) = CURRENT_DATE  
10 AND c.genre_id = ". $this->Db->quote($genreId) . "  
11 AND a.genre_id = ". $this->Db->quote($genreId) . "  
12 AND a.item_i = ". $this->Db->quote($itemId) . "  
13 ORDER BY a.similarity DESC" . ($limit ? " LIMIT " . (int)$limit : ");
```

4.1.4 Pembuatan Model Item

Model item dibuat berdasarkan data rating item dalam tabel rating. Oleh karena perhitungan prediksi rating dilakukan untuk masing-masing genre, maka item-item yang diperbandingkan/dipasangkan juga dalam satu genre yang sama. Secara mendetail, langkah-langkah yang dilakukan adalah sebagai berikut:

1. Untuk setiap user, dicari item-item dalam satu genre yang sama.
2. Selanjutnya, untuk masing-masing item, dipasangkan satu sama lainnya lalu dihitung *similarity*-nya.
3. Setelah diketahui pasangan item yang terbentuk dan *similarity*-nya, dilakukan penyimpanan ke dalam tabel *similarity*.
4. Setelah diketahui *similarity*-nya, maka dilakukan penghitungan *weighted sum*. Hasil dari penghitungan *weighted sum* dilakukan penyimpanan ke dalam tabel *prediksi*.
5. Selanjutnya membuat rekomendasi dari item top-N yang diambil dari tabel *similarity (non-personalized)* dan tabel *prediksi (personalized)*.

Pembuatan model item dilakukan secara *offline* untuk mencari semua pasangan item sehingga dapat digunakan untuk perhitungan prediksi dan pembuatan rekomendasi secara *online*. Pembuatan awal model item dilakukan dengan menjalankan sebuah file *batch*.

Banyaknya item dapat diketahui dari tabel *rating*, yakni dengan melihat banyaknya item yang dipasangkan. Penulis menggunakan komputer (*hardware*) seperti yang telah disebutkan pada sub bab 3.2.1. Penulis menggunakan 2 kali pengujian yaitu :

1. Item tidak dibedakan oleh genre
Ini dilakukan untuk mengetahui *sensitivitas* dari perhitungan komputasi apabila jumlah item dan user bertambah. Ujicoba dilakukan dengan menggunakan matrik item – user ($n \times m$) dengan iterasi 10, dimana data semua user merating seluruh item yang disediakan.

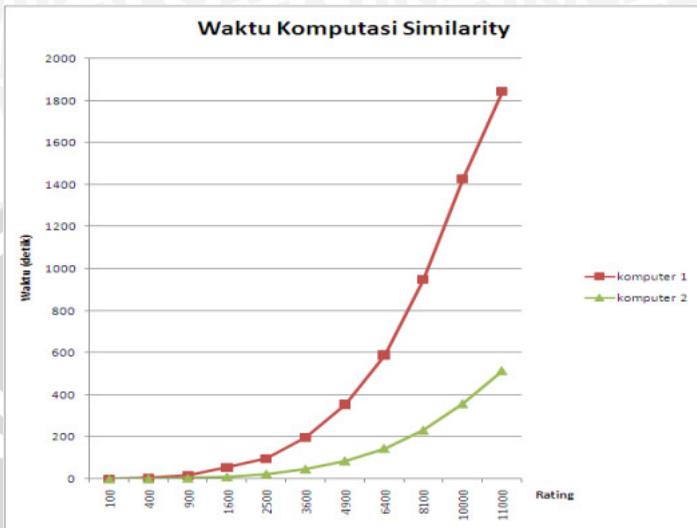
2. Item dibedakan oleh genre

Ini dilakukan untuk mengetahui waktu ideal yang dibutuhkan untuk seluruh item yang dihitung berdasarkan genre, dengan menggunakan dataset yang telah disebutkan pada bab 3.13.

Untuk ujicoba 1 dilakukan set genre adalah merupakan 1 genre yang sama dimana hasil penghitungan waktu komputasi nya ditampilkan pada tabel 4.2.

Tabel 4.2 Penghitungan similarity terhadap sensitivitas item-user

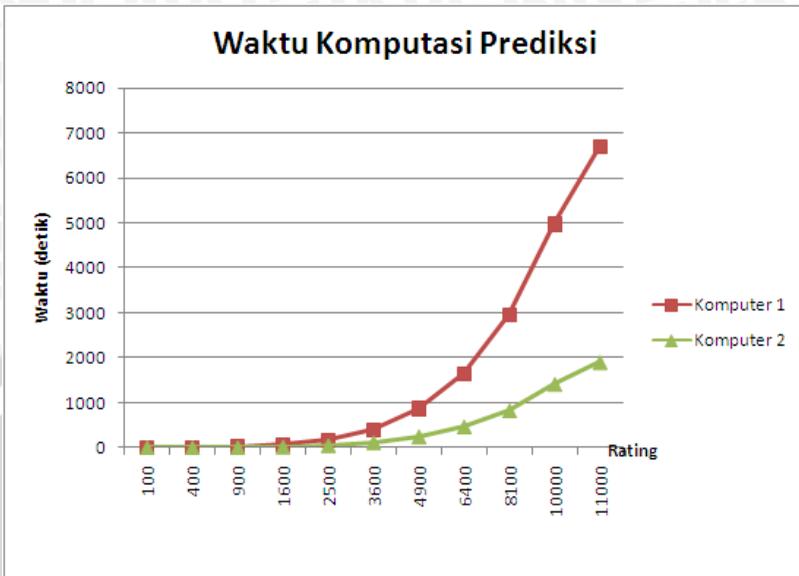
user	item	rating	pasangan item	similarity			
				Waktu komputer 1	rata2	Waktu komputer 2	rata2
10	10	100	90	0,5908	0,006564	0,0631	0,0007
20	20	400	380	3,1036	0,008167	0,6483	0,00171
30	30	900	870	15,5264	0,017846	3,9966	0,00459
40	40	1600	1560	56,4148	0,036163	15,2993	0,00981
50	50	2500	2450	162,3695	0,066273	45,8747	0,01872
60	60	3600	3540	392,7553	0,110948	111,0099	0,03136
70	70	4900	4830	862,7665	0,178627	240,5094	0,04979
80	80	6400	6320	1647,4751	0,260676	464,281	0,07346
90	90	8100	8010	2960,5222	0,369603	824,8624	0,10298
100	100	10000	9900	4976,4772	0,502674	1412,9546	0,14272
100	110	11000	10890	6692,70106	0,614573	1903,4208	0,17479



Gambar 4.1 Waktu komputasi similarity terhadap peningkatan item-user

Tabel 4.3 Penghitungan prediksi terhadap sensitivitas item-user

user	item	rating	pasangan item	prediksi			
				waktu komputer1 (detik)	rata2	waktu komputer2 (detik)	rata2
10	10	100	90	0,7889	0,0079	0,1639	0,0016
20	20	400	380	4,4218	0,0111	0,8736	0,0022
30	30	900	870	15,8033	0,0176	3,3266	0,0037
40	40	1600	1560	54,7621	0,0342	9,7994	0,0061
50	50	2500	2450	97,8175	0,0391	23,0075	0,0092
60	60	3600	3540	198,4205	0,0551	46,6374	0,013
70	70	4900	4830	355,2198	0,0725	85,7994	0,0175
80	80	6400	6320	590,6924	0,0923	145,4286	0,0227
90	90	8100	8010	948,4698	0,1171	231,6797	0,0286
100	100	10000	9900	1427,9682	0,1428	357,061	0,0357
100	110	11000	10890	1843,204	0,1676	514,1855	0,0467



Gambar 4.2 Waktu komputasi prediksi terhadap peningkatan item-user

Dari tabel 4.2 dan tabel 4.3 dapat diketahui bahwa semakin banyak peningkatan item dan user maka waktu yang dibutuhkan untuk menghitung komputasi model data akan lebih besar pula. Sehingga semakin banyak jumlah item dan user maka dengan menggunakan komputer 2 yang lebih baik spesifikasinya, penghitungan dapat dilakukan lebih cepat. Diketahui bahwa waktu rata-rata setiap pemasangan item (*similarity*), paling tinggi yang dibutuhkan untuk menghitung 110 item x 100 user untuk komputer 1 adalah 0,6145 detik dan komputer 2 adalah 0,1747 detik. Sedangkan untuk penghitungan prediksi paling tinggi untuk user-item adalah untuk komputer 1 sebesar 0,1676 detik dan komputer 2 sebesar 0,0467 detik.

Sedangkan untuk ujicoba ke -2 dilakukan penghitungan item yang dipisahkan ke dalam setiap genre (11 genre), dan waktu proses komputasi yang diperlukan untuk memproses item pada setiap genre ditunjukkan pada tabel 4.4.

Tabel 4.4 Waktu komputasi pembuatan model item

No	Genre	Pasangan Item	Waktu Komputer 1 (detik)	Waktu Komputer 2 (detik)	Selisih waktu (detik)
1	Sinetron	90	13,4051	3,6254	9,7797
2	Musik	90	12,6841	3,6681	9,016
3	Reality show	90	12,7187	3,651	9,0677
4	Program anak	90	12,7629	3,6655	9,0974
5	Berita	90	12,7721	3,6739	9,0982
6	Infotainment	90	12,8300	3,6924	9,1376
7	Religi	90	12,8680	3,7075	9,1605
8	Olahraga	90	12,8827	3,7196	9,1631
9	Variety show	90	12,9361	3,7298	9,2063
10	Sitkom / komedi	90	13,1437	3,7516	9,3921
11	Talk show	90	13,1766	3,763	9,4136
Rata-rata Total		90	12,9254	3,6952	9,2302
Rata-rata per item		1	0,1436	0,041	0,10255
Total		990	142,18	40,6478	101,5322

Sehingga apabila dihitung keseluruhan total waktu yang dibutuhkan oleh komputer 1 adalah 142,18 detik, dengan rata-rata setiap genre adalah 12,9254 detik. Sedangkan rata-rata setiap pemasangan item_i dengan item_j untuk penghitungan *similarity* oleh perangkat komputer 1 membutuhkan waktu 0,1436 detik.

Sedangkan dengan menggunakan perangkat komputer 2 menghasilkan waktu yang lebih cepat, yaitu total waktu 40,6478 detik. Dengan rata-rata setiap genre 3,6952 detik, dan rata-rata setiap pemasangan item_i dan item_j untuk penghitungan *similarity* membutuhkan waktu 0,041 detik.

Dihasilkan waktu yang lebih cepat terhadap komputer 2, yaitu lebih cepat 0,10255 detik setiap pemasangan item_i dan item_j pada penghitungan similarity dan total waktu lebih cepat 101,5322 detik.

Diketahui bahwa dengan mengkategorikan item pada genre maka akan lebih mempercepat model penghitungan karena item yang dipasangkan hanya dibatasi per-genre saja, sehingga tidak memunculkan pasangan yang banyak.

4.1.5 Hasil yang ditampilkan

Berikut akan diuraikan hasil dari tampilan sistem rekomendasi acara televisi menggunakan *item-based top-N*.

4.1.5.1 Home

Halaman ini menampilkan halaman beranda / home

Ilmu Komputer Universitas Brawijaya login Home Daftar Acara	Halaman Home
	Selamat Datang ! Ini merupakan Sistem Rekomendasi Acara Televisi, yang dapat memberikan rekomendasi acara televisi sesuai dengan rating yang diberikan oleh user.

Gambar 4.3 Halaman home

4.1.5.2 Login

Halaman ini digunakan oleh user untuk melakukan login

Ilmu Komputer Universitas Brawijaya login Home Daftar Acara	Halaman Login
	Username <input type="text"/> Password <input type="password"/> <input type="button" value="LOGIN"/> (Daftar baru)

Gambar 4.4 Halaman login

4.1.5.3 Daftar User Baru

Halaman ini digunakan untuk pendaftaran bagi user baru

Ilmu Komputer Universitas Brawijaya	Halaman Pendaftaran User Baru	
login	Nama	<input type="text" value="user1"/>
Home	Kelamin	<input type="text" value="Laki-Laki"/>
Daftar Acara	Tgl Lahir	<input type="text" value="19-02-1998"/> (dd-mm-yyyy)
- Adi Pratama -	Username	<input type="text" value="user1"/>
	Password	<input type="password" value="*****"/>
		<input type="button" value="DAFTAR"/>

Gambar 4.5 Tampilan daftar user baru

4.1.5.4 Daftar Genre

Daftar genre menampilkan daftar genre yang urut berdasarkan abjad.

Ilmu Komputer Universitas Brawijaya	Halaman Daftar Acara
login	Genre Acara
Home	<ul style="list-style-type: none">◆ Berita◆ Infotainment◆ Musik◆ Olahraga◆ Program Anak◆ Reality Show◆ Religi◆ Sinetron◆ Sitkom / Komedi◆ Talk Show◆ Variety Show
Daftar Acara	
- Adi Pratama -	

Gambar 4.6 Tampilan halaman daftar genre

4.1.5.5 Daftar Acara per Genre

Daftar acara menampilkan daftar item/acara yang terurut berdasarkan genre.

Ilmu Komputer Universitas Brawijaya Home Daftar Acara - Adi Pratama -	Halaman Daftar Acara																																																										
	Genre Acara » Sinetron																																																										
	<table border="1"> <thead> <tr> <th>No.</th> <th>Nama Acara</th> <th>Stasiun TV</th> <th>Hari (Tanggal)</th> <th>Jam</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Kepompong : Episode 54</td> <td>SCTV</td> <td>Jumat, 7 Agustus 2009</td> <td>16:00 WIB</td> </tr> <tr> <td>2</td> <td>Tarzan Klik : Episode 19</td> <td>RCTI</td> <td>Jumat, 7 Agustus 2009</td> <td>17:30 WIB</td> </tr> <tr> <td>3</td> <td>Manohara : Episode 36</td> <td>RCTI</td> <td>Jumat, 7 Agustus 2009</td> <td>18:00 WIB</td> </tr> <tr> <td>4</td> <td>Tangisan Isabella : Episode 35</td> <td>INDOSIAR</td> <td>Jumat, 7 Agustus 2009</td> <td>18:00 WIB</td> </tr> <tr> <td>5</td> <td>Ben 7 : Episode 13</td> <td>TPI</td> <td>Jumat, 7 Agustus 2009</td> <td>18:00 WIB</td> </tr> <tr> <td>6</td> <td>Cinta Dan Anugerah : Episode 6</td> <td>RCTI</td> <td>Jumat, 7 Agustus 2009</td> <td>19:00 WIB</td> </tr> <tr> <td>7</td> <td>Melati Untuk Marvel : Episode 64</td> <td>SCTV</td> <td>Jumat, 7 Agustus 2009</td> <td>19:30 WIB</td> </tr> <tr> <td>8</td> <td>Dewi : Episode 34</td> <td>RCTI</td> <td>Jumat, 7 Agustus 2009</td> <td>20:30 WIB</td> </tr> <tr> <td>9</td> <td>Cinta Fitri : Episode 74</td> <td>SCTV</td> <td>Jumat, 7 Agustus 2009</td> <td>20:30 WIB</td> </tr> <tr> <td>10</td> <td>Terlanjur Cinta : Episode 21</td> <td>SCTV</td> <td>Jumat, 7 Agustus 2009</td> <td>21:30 WIB</td> </tr> </tbody> </table>	No.	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam	1	Kepompong : Episode 54	SCTV	Jumat, 7 Agustus 2009	16:00 WIB	2	Tarzan Klik : Episode 19	RCTI	Jumat, 7 Agustus 2009	17:30 WIB	3	Manohara : Episode 36	RCTI	Jumat, 7 Agustus 2009	18:00 WIB	4	Tangisan Isabella : Episode 35	INDOSIAR	Jumat, 7 Agustus 2009	18:00 WIB	5	Ben 7 : Episode 13	TPI	Jumat, 7 Agustus 2009	18:00 WIB	6	Cinta Dan Anugerah : Episode 6	RCTI	Jumat, 7 Agustus 2009	19:00 WIB	7	Melati Untuk Marvel : Episode 64	SCTV	Jumat, 7 Agustus 2009	19:30 WIB	8	Dewi : Episode 34	RCTI	Jumat, 7 Agustus 2009	20:30 WIB	9	Cinta Fitri : Episode 74	SCTV	Jumat, 7 Agustus 2009	20:30 WIB	10	Terlanjur Cinta : Episode 21	SCTV	Jumat, 7 Agustus 2009	21:30 WIB			
No.	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam																																																							
1	Kepompong : Episode 54	SCTV	Jumat, 7 Agustus 2009	16:00 WIB																																																							
2	Tarzan Klik : Episode 19	RCTI	Jumat, 7 Agustus 2009	17:30 WIB																																																							
3	Manohara : Episode 36	RCTI	Jumat, 7 Agustus 2009	18:00 WIB																																																							
4	Tangisan Isabella : Episode 35	INDOSIAR	Jumat, 7 Agustus 2009	18:00 WIB																																																							
5	Ben 7 : Episode 13	TPI	Jumat, 7 Agustus 2009	18:00 WIB																																																							
6	Cinta Dan Anugerah : Episode 6	RCTI	Jumat, 7 Agustus 2009	19:00 WIB																																																							
7	Melati Untuk Marvel : Episode 64	SCTV	Jumat, 7 Agustus 2009	19:30 WIB																																																							
8	Dewi : Episode 34	RCTI	Jumat, 7 Agustus 2009	20:30 WIB																																																							
9	Cinta Fitri : Episode 74	SCTV	Jumat, 7 Agustus 2009	20:30 WIB																																																							
10	Terlanjur Cinta : Episode 21	SCTV	Jumat, 7 Agustus 2009	21:30 WIB																																																							

Gambar 4.7 Tampilan halaman daftar acara

4.1.5.6 Detail Acara Bagi User yang Tidak Melakukan Login

Halaman ini menampilkan detail acara beserta rekomendasi *non-personalized* bagi user yang tidak melakukan login. Rekomendasi yang bersifat *non-personalized* ini ditampilkan berupa 5 item teratas yang paling berkorelasi dengan item yang sedang dilihat/ ditampilkan, yaitu diambil dari *tabel_similarity*. Untuk mengetahui semua item yang berkorelasi, disediakan tampilan lengkap seperti pada Gambar 4.8.

Ilmu Komputer Universitas Brawijaya Home Daftar Acara - Adi Pratama -	Halaman Daftar Acara																																								
	Genre Acara » Sinetron » Kepompong																																								
	Nama Acara : Kepompong Stasiun TV : SCTV Hari / Tanggal : Senin, 3 Agustus 2009 Jam : 16:00 WIB Sinopsis : keterangan cerita singkat tentang Kepompong : Episode 54																																								
	Rekomendasi :																																								
	<table border="1"> <thead> <tr> <th>No.</th> <th>Nama Acara</th> <th>Stasiun TV</th> <th>Hari (Tanggal)</th> <th>Jam</th> <th>Prediksi</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Cinta Dan Anugerah : Episode 2</td> <td>RCTI</td> <td>Senin, 3 Agustus 2009</td> <td>19:00 WIB</td> <td>0.855</td> </tr> <tr> <td>2</td> <td>Melati Untuk Marvel : Episode 60</td> <td>SCTV</td> <td>Senin, 3 Agustus 2009</td> <td>19:30 WIB</td> <td>0.843</td> </tr> <tr> <td>3</td> <td>Cinta Fitri : Episode 70</td> <td>SCTV</td> <td>Senin, 3 Agustus 2009</td> <td>20:30 WIB</td> <td>0.842</td> </tr> <tr> <td>4</td> <td>Tarzan Klik : Episode 15</td> <td>RCTI</td> <td>Senin, 3 Agustus 2009</td> <td>17:30 WIB</td> <td>0.832</td> </tr> <tr> <td>5</td> <td>Manohara : Episode 32</td> <td>RCTI</td> <td>Senin, 3 Agustus 2009</td> <td>18:00 WIB</td> <td>0.829</td> </tr> </tbody> </table>	No.	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam	Prediksi	1	Cinta Dan Anugerah : Episode 2	RCTI	Senin, 3 Agustus 2009	19:00 WIB	0.855	2	Melati Untuk Marvel : Episode 60	SCTV	Senin, 3 Agustus 2009	19:30 WIB	0.843	3	Cinta Fitri : Episode 70	SCTV	Senin, 3 Agustus 2009	20:30 WIB	0.842	4	Tarzan Klik : Episode 15	RCTI	Senin, 3 Agustus 2009	17:30 WIB	0.832	5	Manohara : Episode 32	RCTI	Senin, 3 Agustus 2009	18:00 WIB	0.829				
No.	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam	Prediksi																																				
1	Cinta Dan Anugerah : Episode 2	RCTI	Senin, 3 Agustus 2009	19:00 WIB	0.855																																				
2	Melati Untuk Marvel : Episode 60	SCTV	Senin, 3 Agustus 2009	19:30 WIB	0.843																																				
3	Cinta Fitri : Episode 70	SCTV	Senin, 3 Agustus 2009	20:30 WIB	0.842																																				
4	Tarzan Klik : Episode 15	RCTI	Senin, 3 Agustus 2009	17:30 WIB	0.832																																				
5	Manohara : Episode 32	RCTI	Senin, 3 Agustus 2009	18:00 WIB	0.829																																				
	Halaman ini di generasi dalam waktu : 0.1281 detik																																								

Gambar 4.8 Halaman detail acara dan rekomendasi non-personalized

4.1.5.7 Detail Acara Bagi User yang Melakukan Login

Halaman ini menampilkan detail acara bagi user yang melakukan login. User yang melakukan login pada halaman ini bisa memberikan rating pada item/acara yang sedang dilihat. Halaman detail acara bagi user yang melakukan login ini ditampilkan dalam bentuk 2 output/tampilan, yaitu :

1. Bagi user yang login dan tidak memiliki rating akan ditampilkan rekomendasi dari hasil perhitungan *similarity* yang diambil dari tabel_similarity, ini adalah merupakan rekomendasi *non-personalized*.
2. Bagi user yang login dan memiliki rating akan ditampilkan rekomendasi teratas dari hasil prediksi yang diambil pada tabel_prediksi, ini adalah merupakan rekomendasi *personalized*.

Jika seorang user belum merating satu pun item dalam sebuah genre, tidak memperoleh rekomendasi *personalized*. Hal ini disebabkan karena sistem tidak dapat mengetahui preferensi rating dari user yang bersangkutan sehingga tidak ada input *item_i* untuk diikutsertakan dalam perhitungan prediksi pada genre tersebut. Sebagai solusinya, user masih dapat menerima rekomendasi berupa daftar item-item yang berkorelasi dengan item yang sedang dilihat, seperti halnya pada rekomendasi *non-personalized*.

Ilmu Komputer Universitas Brawijaya

Anda login sebagai **nicholas** (logout)

Home

Profil

Daftar Acara

Rating Anda

- Adi Pratama -

Halaman Daftar Acara

Genre Acara » Sinetron » Kepompong

Nama Acara : Kepompong
 Stasiun TV : SCTV
 Hari / Tanggal : Senin, 3 Agustus 2009
 Jam : 16:00 WIB
 Sinopsis : keterangan cerita singkat tentang Kepompong : Episode 54

Rating Anda : 1 2 3 4 5

Rekomendasi :

No.	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam	Prediksi
1	Cinta Dan Anugerah : Episode 2	RCTI	Senin, 3 Agustus 2009	19:00 WIB	0.855
2	Melati Untuk Marvel : Episode 60	SCTV	Senin, 3 Agustus 2009	19:30 WIB	0.843
3	Cinta Fitri : Episode 70	SCTV	Senin, 3 Agustus 2009	20:30 WIB	0.842
4	Tarzan Cilik : Episode 15	RCTI	Senin, 3 Agustus 2009	17:30 WIB	0.832
5	Manohara : Episode 32	RCTI	Senin, 3 Agustus 2009	18:00 WIB	0.829

Halaman ini di generasi dalam waktu : 0.1274 detik

(a) tampilan user yang login dengan rekomendasi non-personalized

Ilmu Komputer Universitas Brawijaya

Anda login sebagai **user1** (logout)

Home

Profil

Daftar Acara

Rating Anda

- Adi Pratama -

Halaman Daftar Acara

Genre Acara » Sinetron » Kepompong

Nama Acara : Kepompong
 Stasiun TV : SCTV
 Hari / Tanggal : Jumat, 7 Agustus 2009
 Jam : 16:00 WIB
 Sinopsis : keterangan cerita singkat tentang Kepompong : Episode 54

Rating Anda : 1 2 3 4 5

Rekomendasi :

No.	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam	Prediksi
1	Melati Untuk Marvel : Episode 64	SCTV	Jumat, 7 Agustus 2009	19:30 WIB	0.851
2	Kepompong : Episode 54	SCTV	Jumat, 7 Agustus 2009	16:00 WIB	0.831
3	Manohara : Episode 36	RCTI	Jumat, 7 Agustus 2009	18:00 WIB	0.826
4	Cinta Fitri : Episode 74	SCTV	Jumat, 7 Agustus 2009	20:30 WIB	0.826
5	Tarzan Cilik : Episode 19	RCTI	Jumat, 7 Agustus 2009	17:30 WIB	0.822

Halaman ini di generasi dalam waktu : 0.08 detik

(b) tampilan user ke-1 yang login dengan rekomendasi personalized

Ilmu Komputer Universitas Brawijaya

Anda login sebagai **user10** (logout)

Home

Profil

Daftar Acara

Rating Anda

- Adi Pratama -

Halaman Daftar Acara

Genre Acara » Sinetron » Kepompong

Nama Acara : Kepompong
 Stasiun TV : SCTV
 Hari / Tanggal : Jumat, 7 Agustus 2009
 Jam : 16:00 WIB
 Sinopsis : keterangan cerita singkat tentang Kepompong : Episode 54

Rating Anda : 1 2 3 4 5

Rekomendasi :

No.	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam	Prediksi
1	Melati Untuk Marvel : Episode 64	SCTV	Jumat, 7 Agustus 2009	19:30 WIB	0.841
2	Manohara : Episode 36	RCTI	Jumat, 7 Agustus 2009	18:00 WIB	0.832
3	Kepompong : Episode 54	SCTV	Jumat, 7 Agustus 2009	16:00 WIB	0.830
4	Cinta Fitri : Episode 74	SCTV	Jumat, 7 Agustus 2009	20:30 WIB	0.824
5	Cinta Dan Anugerah : Episode 6	RCTI	Jumat, 7 Agustus 2009	19:00 WIB	0.822

Halaman ini di generasi dalam waktu : 0.1427 detik

(c) tampilan user ke-2 yang login dengan rekomendasi personalized

Gambar 4.9 Tampilan user yang login dengan pemberian rekomendasi

Dari gambar 4.9(b) dan 4.9(c) dapat dilihat bahwa user yang berbeda akan ditampilkan rekomendasi yang berbeda pula, hal ini sesuai dengan rating yang telah diberikan oleh user tersebut. Sehingga tampilan yang diberikan adalah personalisasi masing-masing user tersebut.

4.1.5.8 Halaman Profil User

Halaman profil user ini menampilkan profil dari user yang sedang melakukan login

Ilmu Komputer Universitas Brawijaya	Halaman Profil
Anda login sebagai user1 (logout)	Nama : User1
Home	Kelamin : Laki-laki
Profil	Tanggal lahir : 1 Oktober 1981
Daftar Acara	
Rating Anda	

Gambar 4.10 Halaman profil user

4.1.5.9 Halaman Rating User

Halaman rating user menampilkan banyaknya item yang dirating, rata-rata rating yang diberikan pada setiap itemnya, serta daftar semua item yang telah diberi rating oleh user yang bersangkutan.

	Halaman Daftar Acara																																				
Anda login sebagai rudy (logout)	Genre Acara » Sinetron » Kepompong																																				
Home	Nama Acara : Kepompong																																				
Profil	Stasiun TV : SCTV																																				
Daftar Acara	Hari / Tanggal : Senin, 3 Agustus 2009																																				
Rating Anda	Jam : 16:00 WIB																																				
	Sinopsis : keterangan cerita singkat tentang Kepompong : Episode 50																																				
	Rating Anda : <input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input checked="" type="radio"/> 4 <input type="radio"/> 5 <input type="button" value="Ratel"/>																																				
	Rekomendasi :																																				
	<table border="1"> <thead> <tr> <th>No.</th> <th>Nama Acara</th> <th>Stasiun TV</th> <th>Hari (Tanggal)</th> <th>Jam</th> <th>Prediksi</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>Cinta Dan Anugerah : Episode 2</td> <td>RCTI</td> <td>Senin, 3 Agustus 2009</td> <td>19:00 WIB</td> <td>0.855</td> </tr> <tr> <td>2</td> <td>Melati Untuk Marvel : Episode 60</td> <td>SCTV</td> <td>Senin, 3 Agustus 2009</td> <td>19:30 WIB</td> <td>0.842</td> </tr> <tr> <td>3</td> <td>Cinta Fitri : Episode 70</td> <td>SCTV</td> <td>Senin, 3 Agustus 2009</td> <td>20:30 WIB</td> <td>0.842</td> </tr> <tr> <td>4</td> <td>Tarzan Klik : Episode 15</td> <td>RCTI</td> <td>Senin, 3 Agustus 2009</td> <td>17:30 WIB</td> <td>0.832</td> </tr> <tr> <td>5</td> <td>Manohara : Episode 32</td> <td>RCTI</td> <td>Senin, 3 Agustus 2009</td> <td>18:00 WIB</td> <td>0.829</td> </tr> </tbody> </table>	No.	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam	Prediksi	1	Cinta Dan Anugerah : Episode 2	RCTI	Senin, 3 Agustus 2009	19:00 WIB	0.855	2	Melati Untuk Marvel : Episode 60	SCTV	Senin, 3 Agustus 2009	19:30 WIB	0.842	3	Cinta Fitri : Episode 70	SCTV	Senin, 3 Agustus 2009	20:30 WIB	0.842	4	Tarzan Klik : Episode 15	RCTI	Senin, 3 Agustus 2009	17:30 WIB	0.832	5	Manohara : Episode 32	RCTI	Senin, 3 Agustus 2009	18:00 WIB	0.829
No.	Nama Acara	Stasiun TV	Hari (Tanggal)	Jam	Prediksi																																
1	Cinta Dan Anugerah : Episode 2	RCTI	Senin, 3 Agustus 2009	19:00 WIB	0.855																																
2	Melati Untuk Marvel : Episode 60	SCTV	Senin, 3 Agustus 2009	19:30 WIB	0.842																																
3	Cinta Fitri : Episode 70	SCTV	Senin, 3 Agustus 2009	20:30 WIB	0.842																																
4	Tarzan Klik : Episode 15	RCTI	Senin, 3 Agustus 2009	17:30 WIB	0.832																																
5	Manohara : Episode 32	RCTI	Senin, 3 Agustus 2009	18:00 WIB	0.829																																

Gambar 4.11 Tampilan halaman rating user

4.2 Analisa Hasil

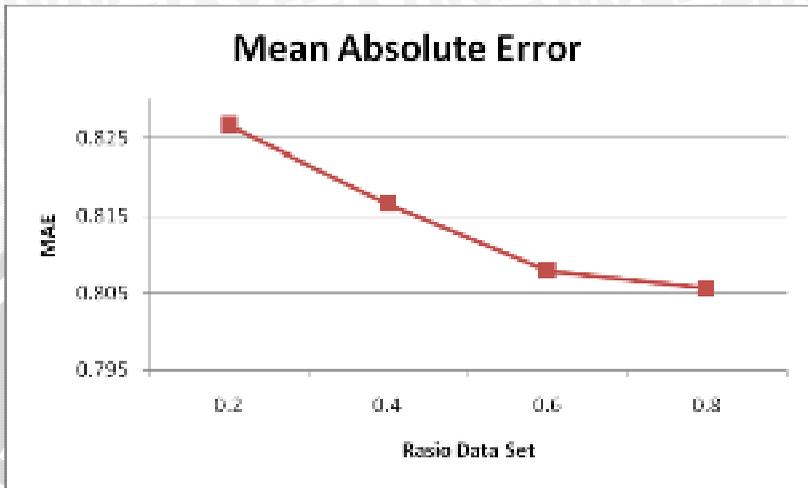
Pada analisis hasil ini akan ditunjukkan hasil dari *Mean Absolute Error (MAE)* yaitu penyimpangan hasil prediksi rekomendasi terhadap rating sebenarnya. Dan juga akan ditunjukkan hasil waktu komputasi untuk dapat menampilkan sebuah rekomendasi.

4.2.1 Hasil MAE (Mean Absolute Error)

Pengujian dilakukan untuk mengetahui nilai keakuratan prediksi dan waktu komputasi dari metode *item-based top-n recommendation* yang digunakan. Untuk mengukur keakuratan prediksi, digunakan ukuran *mean absolute error (MAE)* seperti yang disebutkan pada persamaan 2.4. Hasil *MAE* ditunjukkan pada tabel 4.5. Pengujian untuk mengetahui hasil *MAE* dilakukan sebanyak 4 kali dengan percobaan yang berbeda dengan training set dan test set yang berbeda. Dimana pengertian nilai 0,2 adalah 20% training set dan 80% test set, dengan iterasi (kenaikan) 0,2. Sehingga hasil pengujian ditunjukkan pada tabel 4.5.

Tabel 4.5 Hasil *Mean Absolute Error*

No	Rasio Data Set	Jumlah Prediksi	MAE
1	0,2	7989	0,8266
2	0,4	9986	0,8166
3	0,6	5990	0,8079
4	0,8	1997	0,8056
Rata-rata			0,8141



Gambar 4.12 Hasil MAE dengan beragam rasio dataset

Hasil pada gambar 4.12 menunjukkan hasil pengujian yang telah dilakukan, besarnya *MAE* yang didapatkan adalah rata-rata sebesar 0,8141. Sehingga dapat disimpulkan bahwa rating setiap item yang diprediksi mengalami rata-rata penyimpangan sebesar 0,8141 dari rating user yang sebenarnya dalam skala rating 1-5. Dari rasio dataset juga diketahui dengan semakin besar *training set* maka *MAE* akan semakin kecil, hal ini karena semakin banyak data pembelajarannya (*knowledge based*).

4.2.2 Waktu Komputasi Rekomendasi

Untuk mengetahui waktu komputasi, dilakukan 10 kali percobaan dengan menampilkan prediksi rating item bagi 10 user yang berbeda pada halaman web, 10 user diambil secara random dan dicatat berapa waktu yang diperlukan untuk menampilkan prediksi berdasarkan item yang sedang dirating. Waktu akan dihitung berdasarkan item yang dikategorikan berdasarkan genre dan tidak dikategorikan genre.

Tabel 4.6 Hasil komputasi rekomendasi

Uji Coba	UserID	ItemID yang Dilihat	Waktu 1 (detik)	Waktu 2 (detik)
1	1	5	1,0363	16,5033
2	23	45	1,0696	16,3649
3	45	27	1,0365	15,9771
4	30	73	1,0578	16,1091
5	10	67	1,0426	16,0612
6	15	21	1,1000	16,3540
7	21	20	1,0286	15,9661
8	37	83	1,0953	15,9920
9	50	51	1,0326	16,0584
10	39	49	1,0293	16,0724
Rata-rata			1,0528	16,1458

Keterangan :

Waktu 1 : waktu untuk item yang dikategorikan genre

Waktu 2 : waktu untuk item yang tidak dikategorikan genre

Hasil pada tabel 4.6 menunjukkan bahwa waktu yang dibutuhkan rata-rata untuk menghitung dan menampilkan rekomendasi untuk item yang dikategorikan genre adalah 1,0528 detik. Sedangkan rata-rata waktu untuk menampilkan rekomendasi untuk item yang tidak dikategorikan genre adalah 16,1458. Sehingga ditunjukkan waktu rekomendasi untuk item yang dikategorikan genre lebih cepat daripada item yang tidak dikategorikan berdasarkan genre.

BAB V PENUTUP

5.1 Kesimpulan

Kesimpulan yang dapat diambil dari implementasi dan uji coba yang dilakukan adalah sebagai berikut :

1. Model korelasi antar item dapat di-*update* secara langsung dan tidak hanya bergantung pada proses pembuatan model item dari awal, artinya sebuah rating baru dapat mempengaruhi hasil prediksi rating sebuah item.
2. Rekomendasi yang dibuat dengan metode *item-based* yang digunakan menghasilkan 2 macam output, yakni rekomendasi *personalized* dan *non-personalized*.
3. Hasil perhitungan prediksi dicapai dengan nilai *mean absolute error* yang didapatkan sebesar 0,8141 pada skala rating 1 sampai dengan 5 dengan kenaikan 1 angka.
4. Hasil prediksi yang di tampilkan kepada user rata-rata waktunya akan lebih cepat apabila item tersebut dikategorikan berdasarkan genre, yaitu 1,0528 detik. Sehingga mampu memberikan rekomendasi yang cepat kepada user.

5.2 Saran

Berikut ini adalah beberapa saran untuk pengembangan penelitian tentang *item-based top-N rekomendasi acara televisi* lebih lanjut:

1. Untuk membuat sistem rekomendasi yang lebih akurat, metode *item-based* yang digunakan dapat dipadukan dengan metode-metode lain, semisal *user-based*.
2. Untuk menghasilkan hasil rekomendasi televisi yang lebih selektif, dapat dilakukan pengklasifikasian user berdasarkan data usia/ tanggal lahir, sehingga sistem rekomendasi akan merekomendasikan berdasarkan genre tontonan acara yang layak ditonton untuk kategori usia, Misalkan Semua Umur (SU), Remaja (RJ), Dewasa 18+/17+(D), Bimbingan Orang Tua (BO).

DAFTAR PUSTAKA

- David Bueno, Ricardo Conejo, Javier G. Recuenco. *An architecture for a TV Recommender System*. Interactive TV: A Shared Experience TICSP Adjunct Proceedings of EuroIT V. 2007.
- Deshpande, Mukund & Karypis, George. *Item-Based Top-N Recommendation Algorithms*. ACM Transactions on Information System, Volume 22, Issue 1, pp. 143 - 177. 2004.
- Eirinaki, M., Vazirgiannis, M. 2003. *Web mining for personalization*. ACM on Internet Technology (TOIT)
- Han, Eui-Hong. dan G. Karypis. Feature Based Recommendation System. <http://glaros.dtc.umn.edu/gkhome/fetch/papers/fbrsCIKM05.pdf>, 2005.
- J. V. Barneveld and M. V. Setten: 'User Interfaces for Personalized Information Systems'. 2003.
- John Zimmerman, Kaushal Kurapati, Anna L. Buczak, Dave Schaffer, Jacquelyn Martino, Srinivas Gutta. *TV Personalization System: Design of a TV Show Recommender Engine and Interface*. In Liliana Ardissono, Alfred Kobsa, Mark Maybury (eds) Personalized Digital Television: Targeting Programs to Individual Viewers. 2004
- Karypis, George. *Evaluation of Item-Based Top-N Recommendation Algorithm*. 10th Conference of Information and Knowledge Management (CIKM), pp. 247 – 254. 2001.
- Larose, Daniel T. *Discovering Knowledge in Data: An Introduction to Data Mining*. John Wiley & Sons, Inc. New Jersey: 2005.
- Masruri, Farid. 2006. *Perancangan Recommender System untuk personalisasi web e-commerce menggunakan metode Item-Based Collaborative Filtering*. Universitas Brawijaya.

Nasraoui, O. 2002. *World Wide Web Personalization*. Louisville University

Pramudiono, Iko. *Pengantar Data Mining: Menambang Pematana Pengetahuan di Gunung Data*. Ilmukomputer.com. 2003.

Sarwar Badrul, Karypis George, Joseph Konstan, and John Riedl. *Item-Based Collaborative Filtering Recommendation Algorithms*. WWW10, pp. 285 – 295. 2001.

Schafer, J. B., J. A. Konstan, dan J. Riedl. 2001. *E-Commerce Recommendation Application*. Data Mining and Knowledge Discovery.

Srivastava, J. 2000. *Introduction to Web Mining*. University of Minnesota. USA

Weiyang, Lin. *Association Rule Mining for Collaborative Recommender Systems*. Thesis. Worcester Polytechnic Institute. 2000.

Internet literatur :

<http://www.cs.ualberta.ca/~zaiane/courses/cmput690/slides/Chapter9/sld009.htm>, 15-05-2009

<http://kur2003.if.itb.ac.id/file/pohon.pdf>, 02-07-2009

<http://www.grouplens.org>, 14-08-2009

<http://www.movielens.org>, 14-08-2009

<http://www.rcti.tv>, 01-07-2009

<http://www.twocrows.com> , 01-07-2009