**MAKALAH SISTEM RECOMMENDER**



**Disusun Oleh :**

**Agus Mudiyanto 14.01.53.0040**

**Mukhamad Iqbal Muzadi 14.01.53.0036**

**Vikko Dwi Prakoso 14.01.53.0023**

**UNIVERSITAS STIKUBANK SEMARANG (UNISBANK)**

**FALKUTAS TEKNOLOGI INFORMASI**

**TEKNIK INFORMATIKA**

**2017**

**BAB I**

**PENDAHULUAN**

**1.1 Latar Belakang**

Beragam informasi tersedia di internet dan dapat diakses secara luas oleh penggunanya. Banyaknya data baik berupa teks, gambar, maupun video yang ada menyebabkan terjadinya luapan informasi yang membuat pengguna/user internet seringkali mengalami kesulitan dalam mencari informasi yang benar-benar sedang dicari atau dibutuhkan. Dalam perdagangan toko online atau e-commerce, seperti penjualan CD atau DVD film misalnya, kesulitan yang dihadapi user sebagai calon konsumen adalah ketika mencari sebuah produk film yang tepat dan sesuai dengan selera atau kesukaannya. Kebanyakan web e-commerce yang ada saat ini mempunyai struktur yang besar dan kompleks sehingga membuat informasi yang dibutuhkan oleh calon konsumen justru tidak tersampaikan. Oleh karena itu, personalisasi informasi di internet menjadi hal yang penting saat ini. Bagi sebuah web e-commerce, dibutuhkan sebuah sistem yang mampu beradaptasi dengan profil masing-masing konsumen dan dapat memberi saran atau sugesti kepada konsumen tentang sebuah produk yang cocok bagi konsumen. Dalam kajian web mining, sistem personalisasi web semacam ini disebut dengan recommender system (RSs).

RSs bekerja untuk membantu user dalam memilih item yang akan dibeli dengan cara memberi informasi atau sugesti berdasarkan tingkat ketertarikan atau selera masing-masing user [2]. Ketertarikan atau selera user dapat diketahui dari beberapa hal, misalnya dari cara user mengakses halaman web, dari rating yang diberikan user terhadap suatu item produk, atau dari data transaksi user pada waktu lampau. Sedangkan informasi yang bisa diberikan kepada user dapat berupa ulasan mendetail tentang item, angka penjualan, atau dengan menunjukkan nilai rating dari user-user yang lain (komunitas lain). Contoh web e- commerce populer yang telah cukup sukses mengimplementasikan RSs adalah Amazon [1] dan CDNow [3].

**1.2 Rumusan Permasalahan**

Perumusan masalah yang ada untuk makalah ini adalah

1. Teknik kerja Web-Ecomerce dengan sistem recommender

**1.3 Tujuan Penulisan Makalah**

Tujuan dari penulisan untuk memenuhi tugas dari mata kuliah sistem temu kembali.

**1.4 Sistematika Penulisan Makalah**

**BAB I**

Berisikan tentang latar belakang, perumusan masalah, tujuan penulisan makalah, dan sistematika penulisan makalah.

**BAB II**

Berisikan tentang pembahasan materi sistem recommender.

**BAB III**

Membahas tentang hasil dan pembahasan tentang sistem recommender.

**BAB IV**

Akan membahas tentan kesimpulan dan saran tentang sistem recommender.

**BAB II**

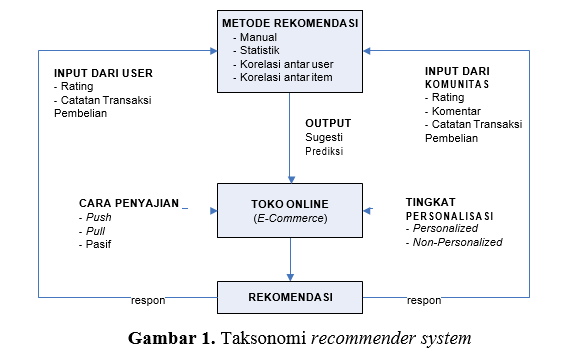
**PEMBAHASAN MATERI**

**2.1 Pengertian Web Ecomerce Recomender**

Personalisasi web merupakan suatu proses untuk menyesuaikan halaman web menurut selera, atau kebiasaan masing-masing user. Personalisasi dibutuhkan karena saat ini sedang terjadi luapan informasi di internet yang membuat user mengalami kesulitan dalam mencari informasi yang berguna dan sedang dibutuhkan. Sebuah cara untuk melakukan personalisasi dalam web adalah melalui recommender system (RSs). Tujuan utama RSs adalah memprediksi item atau suatu informasi yang kemungkinan menarik minat atau berguna bagi user. Dalam web e-commerce, RSs berbasis collaborative filtering (CF) dapat digunakan untuk membantu user dalam memilih item yang sesuai berdasarkan rating yang diberikan user lain. Makalah ini menjelaskan pembuatan web e-commerce yang mampu memberi rekomendasi secara otomatis kepada user. Metode yang digunakan adalah metode item-based CF dengan algoritma Slope One yang membuat model korelasi antar item dengan membandingkan rating antara item satu dengan yang lain dan menghitung rata- rata selisih ratingnya. Semakin kecil rata-rata selisih rating antara 2 item, maka semakin dekat korelasinya. Berdasarkan model korelasi antar item, diprediksi berapa rating yang diberikan user terhadap sebuah item. Rekomendasi ditampilkan kepada user dalam bentuk daftar item hasil prediksi yang telah diurutkan dari rating hasil prediksi yang terbesar. Dalam aplikasi yang dibuat, rekomendasi yang diberikan kepada user ada 2 macam, yaitu rekomendasi personalized bagi user yang terdaftar dan rekomendasi non-personalized bagi user yang tidak terdaftar atau pengunjung.

**2.2 RECOMMENDER SYSTEM**

Recommender system (RSs) adalah salah satu bentuk personalized information system yang digunakan dalam web e-commerce untuk menawarkan item kepada user dan memberi informasi yang dapat membantu



user dalam memilih atau membeli item [8]. Gambar 1 berikut ini menunjukkan taksonomi recommender system [5].a.)

Pada Gambar 1 terdapat tiga 3 komponen utama dalam sebuah recommender system [8], yaitu:

1. Input/ Output

Input yang dianalisa RSs didapat baik secara eksplisit maupun implisit dari user kemudian dikombinasikan dengan input dari user-user lain/ komunitas. Input yang didapat secara eksplisit, misalnya dengan cara meminta user untuk memberi rating terhadap suatu item. Sedangkan secara implisit, misalnya dari data transaksi pembelian item oleh user pada waktu lampau, atau bisa juga dengan memonitor item-item mana saja yang telah dilihat oleh user. Output yang dihasilkan oleh RSs dapat berupa sugesti (merekomendasikan sebuah item secara khusus) atau prediksi (bisa berupa prediksi per item atau beberapa item sekaligus dalam bentuk list/ daftar).

2. Metode Rekomendasi

Metode yang digunakan dalam memberikan rekomendasi dapat dilakukan dengan beberapa cara, antara lain secara manual, melalui pendekatan statistik, dengan berdasarkan korelasi antar user (user-to-user), atau juga dengan berdasarkan korelasi antar item (item-to-item).

3. Desain Rekomendasi

Desain rekomendasi terkait pada 2 hal, yakni bagaimana rekomendasi disajikan dan bagaimana sifat rekomendasi atau tingkat personalisasinya. Ada 3 cara untuk menyajikan rekomendasi kepada user, yaitu:

• Push : bentuk aktif pemberian rekomendasi, seperti mengirimkan kepada user melalui email.

• Pull : rekomendasi tidak ditampilkan jika user tidak meminta

• Pasif : menampilkan item lain yang berhubungan item yang sedang dilihat/ diakses user pada saat itu.

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

• Personalized : rekomendasi yang diberikan kepada user tidak sama antara satu dengan yang lain, bergantung pada masing-masing profil user.

• Non-personalized : bentuk rekomendasi ini tidak melihat profil masing-masing user, dengan kata lain rekomendasi bersifat umum sehingga dapat diberikan bagi semua pengunjung/ visitor.

Berbagai teknik data mining seperti association rule dan market basket analysis [9], nearest neighbor [12], hingga clustering telah digunakan dalam membangun sebuah RSs. Namun, secara garis besar teknik yang digunakan dalam RSs ada 2 macam, yakni content-based filtering dan collaborative filtering [8]. 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.

Penelitian tentang RSs berbasis CF telah banyak dilakukan. RSs berbasis CF yang ada pertama kali menggunakan metode nearest-neighbor [12]. Metode ini dikenal pula dengan nama memory-based atau user-based karena mencari user yang paling mirip dengan user target dalam hal merating item, lalu merekomendasikan item-item yang disukai user terdekat kepada user target. Kelemahan metode memory- based adalah masalah skalabilitas seiring dengan makin banyaknya jumlah user dan jumlah item. Metode model-based atau item-based muncul sebagai solusi metode user-based CF. Berbeda dengan metode user- based yang menghitung kemiripan antar user secara online, metode item-based melakukan pembuatan model korelasi antar item terlebih dahulu secara offline untuk kemudian digunakan dalam membuat rekomendasi secara online sehingga rekomendasi dapat diberikan secara real-time [7].

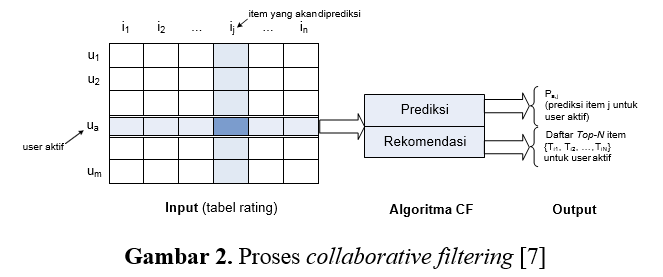
Pada makalah ini metode item-based yang digunakan adalah algoritma Slope One yang mencari korelasi antar item berdasarkan “beda popularitas” [4]. Dengan algoritma ini, akan dibuat 2 macam bentuk rekomendasi kepada user, rekomendasi personalized dan non-personalized.

2.1. **Collaborative Filtering**

Collaborative filtering (CF) adalah suatu metode dalam membuat 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 [7]. Secara umum, ada 2 proses yang dilakukan dalam CF, yaitu:

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

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



Sebagaimana ditunjukkan pada Gambar 2, algoritma CF menghasilkan 2 bentuk output, yaitu [7]:

• Prediksi (Pa,j) adalah suatu nilai yang menyatakan prediksi besarnya rating item j yang mungkin didapat dari user aktif, dimana item j belum pernah dirating oleh user aktif.

• Rekomendasi adalah sebuah daftar yang berisi N item yang mempunyai kemungkinan terbesar untuk disukai oleh user aktif, dimana item-item tersebut belum pernah dirating oleh user aktif. Bentuk ini disebut juga sebagai rekomendasi Top-N.

Ada perbedaan antara rekomendasi (dengan seorang user, disarankan item-item yang mungkin menarik bagi user tersebut) dan prediksi (dengan sebuah item, dilakukan prediksi berapa rating yang mungkin diberikan user pada item tersebut). Namun, pada dasarnya untuk memberi rekomendasi kepada user diperlukan perhitungan prediksi tingkat ketertarikan user terhadap item-item.

Secara umum, berbagai algoritma CF yang ada saat ini dapat dikelompokkan menjadi 2 kategori [7],

yaitu:

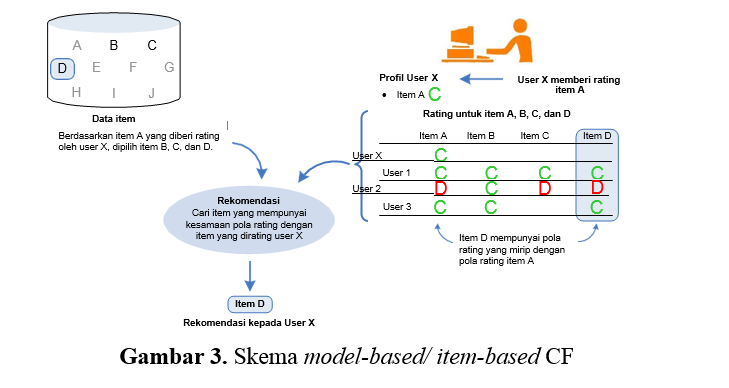
1. Memory-based CF

Algoritma memory-based mempergunakan seluruh data rating dalam membuat prediksi. Metode ini memakai teknik statistik dalam mencari sekumpulan user, disebut dengan nearest-neighbor, yang

mempunyai kemiripan dengan user target. Berdasarkan kemiripan antar user tersebut, dibuat prediksi dan rekomendasi Top-N untuk user target. Metode ini sering disebut juga dengan nearest-neighbor atau user- based collaborative filtering.

2. Model-based CF

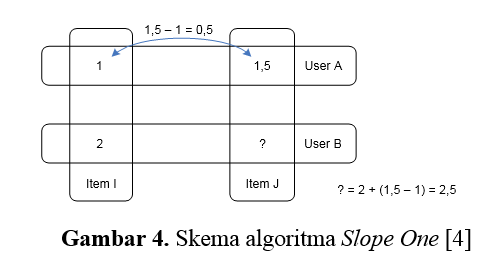
Berbeda dengan memory-based yang mencari hubungan antar user, model-based mencari hubungan antar item berdasarkan tabel rating untuk membuat rekomendasi. Sehingga metode ini disebut juga dengan item- based. Sebelum menghasilkan rekomendasi, algoritma model-based membuat sebuah model korelasi antar item terlebih dahulu untuk mengetahui hubungan antar item berdasarkan nilai rating yang didapat. Proses pembuatan model dapat dilakukan dengan menggunakan berbagai teknik, seperti association rule, classification, atau clustering. Pembuatan model korelasi antar item dilakukan secara offline. Berdasarkan model yang sudah dibentuk, maka perhitungan prediksi atau pemberian rekomendasi dapat dilakukan secara online. Skema model-based atau item-based dapat dilihat pada Gambar 4.



2.2. Algoritma Slope One

Prinsip utama algoritma Slope One adalah bekerja berdasarkan atas “perbedaan popularitas” antara item yang satu dengan yang lainnya dengan cara menghitung selisih rating antara 2 item [4]. Dari selisih rating yang didapat kemudian digunakan untuk memprediksi berapa besarnya nilai rating yang diberikan terhadap suatu item. Untuk dapat melakukan prediksi, algoritma Slope One membutuhkan 2 input, yaitu rating dari user target dan item mana yang akan diprediksi. Oleh karena itu, jika user tidak memberi satu rating pun terhadap item, maka prediksi tidak dapat dilakukan.

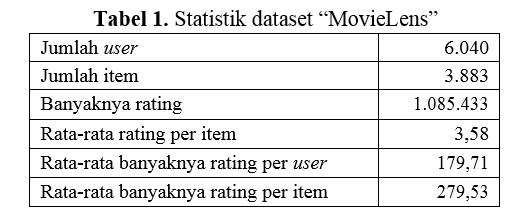
Dimisalkan terdapat 2 user, user A dan user B, dan 2 item, item I dan J. Sebagaimana diperlihatkan pada Gambar 5, user A memberi nilai rating 1 untuk item I dan 1,5 untuk item J, sedangkan user B memberi nilai rating 2 untuk item I saja. Dari sini dapat diketahui bahwa selisih rating antara item I dan J sebesar 1,5−1 = 0,5. Dengan selisih ini, maka dapat diprediksi bahwa user B akan memberi rating terhadap item J sebesar 2+0,5 = 2,5.



1. **METODE PENELITIAN**

* **Persiapan Data**

Untuk mengetahui bagaimana kinerja recommender system, digunakan dataset rating yang telah terisi dengan record rating item dari user. Dataset yang digunakan dalam aplikasi web ini adalah dataset “MovieLens” [6], sebuah dataset rating film hingga tahun 2001. Dataset ini berisi 6.040 data user, 3.883 data item film, dan lebih dari 1 juta data rating dengan skala rating antara 1 sampai 5.



Pembersihan data dilakukan dengan menghapus rating yang tidak relevan dengan data user dan data item. Kategori item disederhanakan menjadi satu kategori saja. Sebelumnya, sebuah item film, seperti misalnya Toy Story, masuk dalam 3 kategori sekaligus yakni “animation”, “children”, dan “comedy”. Untuk mempermudah dalam implementasi, kategori diambil berdasarkan kategori yang tercantum paling awal sehingga untuk contoh film Toy Story dimasukkan dalam kategori “animation”.

1. **Alur Kerja Recommender System**

1.Pemberian Rating Terhadap Item

Proses pemberian rating oleh user terhadap sebuah item dapat ditunjukkan melalui flowchart pada Gambar 6. Rating ui yang diberikan user u terhadap item i akan diperiksa apakah ui sudah ada pada tabel rating atau belum. Jika sudah, ui yang lama akan di-update dengan ui yang baru, sebaliknya jika belum maka ui akan dimasukkan ke dalam rating sebagai data rating baru. Direct\_slope adalah sebuah fungsi yang melakukan update tabel links secara langsung pada saat rating diberikan. Jika fungsi direct\_slope diaktifkan, maka setiap kali seorang user memberi rating terhadap sebuah item, maka data model item ter-update secara otomatis.

2. Pemberian Rekomendasi

Flowchart proses pemberian rekomendasi kepada user ditunjukkan pada Gambar 7. Pada saat user memilih item dalam sebuah kategori, 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 kategori yang sama. Jika kondisi benar, user akan diberi rekomendasi personalized. Namun jika belum, user akan diberi rekomendasi non-personalized. Bagi user yang tidak terdaftar atau visitor, akan selalu diberikanrekomendasi non-personalized.

**BAB III**

**HASIL DAN PEMBAHASAN**

1. **Pembuatan Model Item**

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

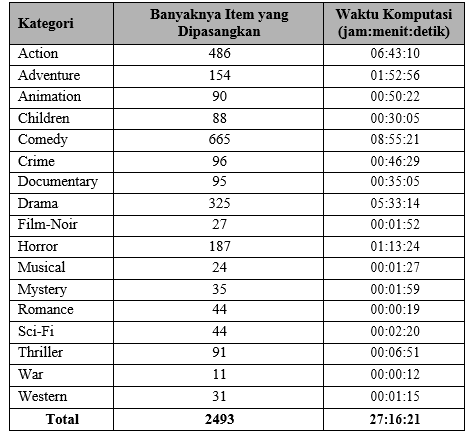
1. Untuk setiap user, dicari item-item dalam satu kategori yang sama.

2. Selanjutnya, untuk masing-masing item, dipasangkan satu sama lainnya lalu dihitung selisih ratingnya. Setiap pasangan item yang terbentuk juga menambahkan 1 nilai cnt\_slope untuk menunjukkan jumlah rating pasangan item.

3. Setelah diketahui pasangan item yang terbentuk dan selisih ratingnya, dilakukan penyimpanan ke dalam tabel rss\_links.

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 yang menghitung selisih rating dan jumlah rating masing-masing pasangan item dari tabel rss\_rating, kemudian menyimpannya dalam tabel rss\_links.

Berikut ini adalah tabel yang menunjukkan waktu yang diperlukan untuk memproses komputasi pada setiap kategori. Banyaknya item dapat diketahui dari tabel rss\_rating, yakni dengan melihat banyaknya item yang dipasangkan.



* 1. Pemberian Rating

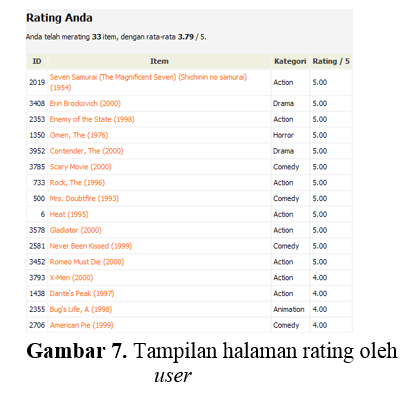
Pemberian rating adalah salah satu proses utama dalam sistem sebagai bentuk interaksi antara user dengan sistem. Pada proses ini, selain menyimpan data rating user ke dalam tabel rss\_ratings, juga dapat dilakukan proses update secara langsung (on-the-fly) terhadap tabel rss\_links. Dengan demikian, tabel model item selalu ter-update mengikuti pola rating dari para user sehingga tidak perlu lagi dilakukan pembuatan model item dari awal

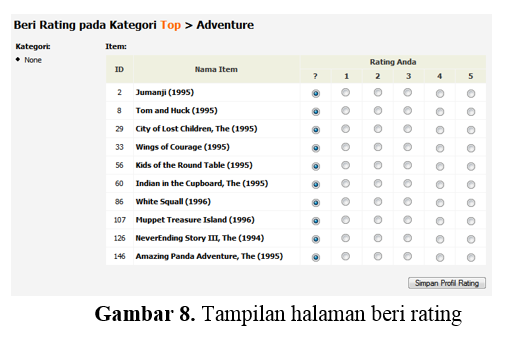
* 1. Fasilitas Recommender System

Fasilitas Recommender System diperuntukkan bagi user yang telah terdaftar dalam sistem. Kegunaan dari fasilitas ini adalah untuk mendapatkan input rating dari user sekaligus memudahkan user merating item- item dalam rangka meningkatkan kualitas rekomendasi bagi user yang bersangkutan.

* 1. Halaman Beri Rating

Halaman ini dikhususkan bagi user untuk melakukan rating terhadap beberapa item sekaligus dalam 1 kategori yang diketahuinya tanpa harus menuju halaman detail item terlebih dahulu. Item dapat ditampilkan dengan urutan kode atau nama item





1. **Kesimpulan**

 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.

 Keakurasian hasil perhitungan prediksi dicapai dengan nilai mean average error yang didapatkan sebesar 0,641.

 Hasil perhitungan prediksi secara online dalam sebuah halaman web dapat ditampilkan dalam kisaran waktu kurang dari 2 detik sehingga mampu memberikan rekomendasi real-time kepada user.

1. **Saran**

Untuk mempertajam hasil rekomendasi, dapat dilakukan pengklasifikasian user berdasarkan data demografinya, sehingga akan dapat diketahui pola seperti misalnya: user yang berusia antara 18–25 menyukai film-film aksi

**DAFTAR PUSTAKA**

[1] Amazon, http://www.amazon.com. Tanggal akses: 21-04-2006.

[2] Barneveld, J.J.F. User Interfaces for Personalized Information Systems. Telematica Instituut. 2003.

[3] CDNOW, http://www.cdnow.com. Tanggal akses: 21-04-2006.

[4] Lemire, Daniel, Anna Maclachlan. Slope one predictors for online rating-based collaborative filtering. SIAM Data Mining (SDM’05). 2005.

[5] Mladenic, Dunja. Text-Learning and Related Intelligent Agents: A Survey. IEEE. 1999.

[6] MovieLens, http://www.cs.umn.edu/GroupLens. Tanggal akses: 20-04-2006.