BAB 2

TINJAUAN PUSTAKA

2.1 Sejarah Sistem Rekomendasi

Kemampuan komputer untuk memberikan rekomendasi diakui cukup awal dalam sejarah komputasi. Perpustakaan berbasis komputer yang menjadi langkah awal terbentuknya sistem rekomendasi otomatis. Sistem ini cukup primitif karena pengelompokan data pengguna berdasarkan hasil dari wawancara dan hard-code dari system untuk mendapatkan daftar rekomendasi buku, hal tersebut merupakan awal yang baik bagi ruang sistem rekomendasi.

Pada Awal 1990-an, sistem rekomendasi dengan *collaborative-filtering* mulai banyak digunakan dan menjadi solusi pengelolaan data yang cukup banyak. *Collaborative filtering system* dapat membantu pengguna untuk melakukan *query* terhadapa beberapa barang yang terdapat dalam satu *domain*, data-data tersebut seperti email perusahaan, berdasarkan opini dari pengguna lain (contoh: “berikan saya data *email* yang di *forward* oleh jhon”. Meski membutuhkan upaya dari pengguna, namun sistem tersebut memungkinkan pengguna untuk dapat memanfaatkan reaksi dari pengguna lain sehingga dapat pengguna dapat menentukan relevansinya.

2.2 Collaborative-Filtering

Metode *collaborative-filtering* merupakan satu pendekatan untuk design *recommender system* yang sudah sering digunakan. Metode penyaringan kolaboratif didasarkan pada mengumpulkan dan menganalisis sejumlah besar informasi tentang perilaku pengguna, kegiatan atau preferensi serta melakukan prediksi apa yang dibutuhkan oleh pengguna berdasarkan apa yang dipilih oleh pengguna lain (Breese, Heckerman, and Kadie, 1998).

Asusmsi yang mendasari *collaborative-filtering* tersebut adalah pendekatan bahwa jika seseorang A memiliki pendapat yang sama dengan B pada suatu masalah, Maka A cenderung mengikuti opini B dalam memandang masalah, misalnya X dari pada memilih pendapat lain dari orang lain yang dipilih secara acak. Sebagai contoh gambar di bawah ini menunjukan pendapat beberapa individu terhadap suatu barang.



Gambar 2.1 Collaborative-Filtering Recommender System
(Moshanin, 2013)

Untuk mengisi tanda tanya tersebut pada teknik *Collaborative-Filtering* asumsi dasar yang digunakan yaitu kesamaan antara orang yang dimaksud dengan orang lain yang pendapatnya sama. Contoh pad bagian foto dan buku, orang yang dimaksud memilih untuk menyukai sedangkan console/game orang tersebut tidak menyukainya. Pendapat tersebut sebagian sama dengan orang ke 2 dan ke 3 (dari atas). Sehingga jawaban untuk tanda tanya tersebut adalah ***tidak menyukai***.

Berikut merupakan kelemahan teknik *Collaborative-Filtering* antara lain :

1) Kelemahan utama pada teknik ini yaitu sistem tidak dapat memberikan

rekomendasi apabila belum adanya penilaian pada object yang di rekomendasikan.

2) *Collaborative-Filtering* akan menghasilkan data yang kurang akurat ketika penilaian pada satu data terlalu sedikit dan akan menjadi salah persepsi.
3) Teknik ini tidak memuat informasi / kegunaan dari barang yang direkomendasikan.

Contoh pengumpulan data eksplisit meliputi proses sebagai berikut ini:

1) Meminta pengguna untuk menilai barang pada skala geser (misalnya: 1 s/d 5).
2) Meminta pengguna untuk melakukan pencarian.
3) Meminta pengguna untuk menilai barang dari favorit ke paling favorit.
4) Menghadirkan dua barang ke pengguna dan meminta pengguna untuk memilih yang lebih baik salah satu dari barang tersebut.
5) Meminta pengguna untuk membuat daftar barang yang disukai.

Contoh pengumpulan data implisit meliputi berikut ini:

1) Mengamati barang yang dilihat pengguna di sebuah toko online.
2) Mengenali barang dari jumlah pengunjung halaman *website* tersebut
3) Mengumpulkan data dari transaksi barang yang dijual.

4) Memperoleh daftar barang yang pengguna telah kunjungi di halaman *website* tersebut.

5) Menganalisis media sosisal mengenai *trending* apakah banyak yang menyukai atau tidak menyukai.

2.3 Content-based

Pendekatan *Content-based* merupakan teknik pembuatan *recommender system* yang berdasarkan penjelasan item dan profil dari penggunanya (Brusilovsky, Kobsa, and Nejdl, 2007).

Content-based merupakan sistem rekomendasi yang cukup umum beberapa tahun terakhir dan sudah banyak diterapkan pada beberapa aplikasi / sistem terutama pada sistem penjualan online. Yang paling populer pendekatan ini diterapkan pada film, musik, berita, buku, artikel penelitian, *social tags*, dan produk lain pada secara umum. Beberapa penelitian pernah menerapkan pendekatan ini untuk melakukan pengangkatan angkatan bersenjata (Buettner, 2014).

Sebagai study kasus yaitu pada sistem rekomendasi yang terdapat pada rekomendasi film / *movie* dengan beberapa atribut seperti *genre*, artis, penulis dan sutradara. Sehingga barang atau jasa yang di rekomendasikan berdasarkan yang sudah populer sebelumnya. Seperti apabila *genre* komedi sedang populer maka apabila ada film baru yang ber genre komedi maka sistem rekomendasi akan menempatkan film

tersebut pada daftar film yang direkomendasikan, walaupun film tersebut termasuk baru dalam data pendukung sistem rekomendasi.

Kekurangan dari pendekatan *content-based* ini adalah :

1) Hanya dapat digunakan untuk fitur yang sesuai, seperti gambar, film dan *music*.

2) Barang / jasa yang di spesialisasi berlebihan, contoh pada *content-based* jarang sekali ditemukan rekomendasi yang kegunaannya berbeda padahal beberapa orang mungkin memiliki kepentingan yang berbeda. Misalkan seseorang yang baru saja membeli mobil sepertinya kurang pas apabila di berikan rekomendsasi barang yang serupa, namun akan cocok apabila barang yang direkomendasikan merupakan aksesoris dari mobil tersebut seperti (ban, lampu mobil, GPS dan akeseoris lain pelengkap mobil tersebut).

3) Tidak mampu menentukan profil dari user baru.

2.4 Cluster Models

Seperti namanya pendekatan ini melakukan pemisahan user/pengguna kedalam beberapa segmen dan melakukan klasifikasi. Tujuan dari algoritma ini yaitu untuk mengelompokan pengguna-pengguna yang mirip untuk kemudian dapat diambil barang mana yang di rekomendasikan untuk pengguna dengan segmen tertentu. (Linden, Smith, and York, 2003)

Evaluasi internal merupakan evaluasi yang terjadi pada pengelompokan

pertama dikarenakan hasil pengelompokan tersebut dievaluasi kembali sehingga mendapatkan hasil yang berbeda. Metode tersebut biasanya merupakan hasil terbaik dari *cluster model* karena algoritma tersebut menghasilkan cluster dengan kesamaan yang tinggi dalam *cluster* dan kesamaan yang renda antara *cluster.*

Namun skor tinggi pada ukuran internal tidak selalu menghasilkan pencarian informasi yang efektif. Selain itu, evaluasi yang dilakukan ini bias terhadap algoritma yang menggunakan model cluster yang sama. Misalnya *K-Means clustering* yang mengoptimalkan jarak objek dan kriteria internal yang berdasarkan jarak kemungkinan akan melebih-lebihkan pengelompokan yang dihasilkan.

2.5 Search Based Model

Metode pencarian atau yang berbasis konten akan menghubungkan satu barang dengan barang lain yang memiliki kesamaan seperti warna, rasa, genre, kategori dan attribute lain agar dapat dikelompokan. (Linden, Smith, and York, 2003).

Seperti halnya *clustering*, *search based* melakukan pengelompokan terhadap objek yang diteliti, namun karena pengelompokan tersebut berdasarkan aktifitas pencarian pada sistem maka, pengelompokan tersebut lebih bersifat general dan terkadang dalam suatu pengelompokan objek satu dengan yang lain dapat sangat berbeda.

2.6 Demographic

Merupakan pendekatan penyusun sistem rekomendasi berdasarkan data / parameter *demography* tiap subjek nya. Contoh kasus sistem rekomendasi akan memberikan rekomendasi yang sesuai dengan demografi orang yang akan direkomendasikan, Misalnya : Budi merupakan orang Bandung, Di Bandung sedang meningkat pembelian ponsel asal Tiongkok, maka Budi di rekomendasikan oleh sistem yang berkaitan dengan ponsel asal Tiongkok.

Keleibhan dari pendekatan ini yaitu :

1) Tidak terpengaruh genre atau jenis barang / jasa pada sistem.
2) Pengetahuan domain pada subjek atau barang yang terdapat pada sistem tidak terlalu di perlukan.
3) Adaptive, Kualitas akan meningkan berbading lurus dengan waktu. (Makin lama rentang data maka kualitas hasil pendekatan ini makin baik).

Kekurangan dari pendekatan ini yaitu :

1) Harus memiliki data demografi tiap orang atau barang.
2) Sulit menangani data dengan user yang memiliki opini tidak konsisten seperti memberikan penilaian yang kurang jelas antara menyukai atau tidak menyukai.

3) Kualitas data tergantung pada besarnya data yang sudah dikumpulkan.

Tiap pendekatan memiliki kelebihan serta keterbatasannya masing-masing, tergantung dari parameter pendukung dibentuknya sistem rekomendasi tersebut. Seperti halnya ilmu eksak lain, pendekatan sistem rekomendasi ini akan terus berkembang.

Penelitian metode demographic ini sudah pernah dilakukan sebelumnya oleh Laila Safoury dan Akram Salah dengan judul penelitian *Exploiting User Demographic Attributes for Solving Cold-Start Problem in Recommender System*, dengan menggunakan kesamaan *demographic* dan *item ratings* yang terdapat penelitian tersebut mendapatkan hasil bahwa masalah data awal yang minim dapat ditangani oleh pendekatan *demographic.* (Safoury and Salah, 2013)

Berdasarkan penelitian terkait mengenai beberapa pendekatan penyusun sistem rekomendasi dapat ditarik kesimpulan bahwa tiap pendekatan memiliki kelebihan serta kekurangannya masing-masing. Pada penelitian ini akan dilakukan analisis dengan menggunakan pendekatan *Hybdrid* yang diharapkan dapat menutupi kekurangan-kekurangan pendekatan lain.

2.7 User Rating System

Merupakan sistem penilaian suatu objek oleh user, biasanya sistem ini digunakan oleh hotel, film, TV series dan restoran. Biasanya rating sistem ini

memiliki nilai 1-5 dimana angka tertinggi menunjukan kualitas terbaik yang dimiliki objek tersebut.

Selain dapat membantu dalam pengambilan keputusan, sistem rating sendiri dapat menjadi *parameter* pendukung pembuatan sistem rekomendasi, dengan user rating diharapkan agar minimnya data tentang subjek dapat teratasi.

2.8 Hybrid Recommender System

Hybrid Recommender System yaitu metode yang menggabungkan dua atau lebih teknik untuk mendapatkan performa yang lebih baik dengan sedikit kelemahan dari setiap individu. Seringkali metode ini terdiri dari *collaborative-filtering* yang dikombinasikan dengan metode lain untuk menghindari *rump-up problem* (Burke 2002). Berikut merupakan beberapa kombinasi yang pernah dilakukan sebelumnya.

1) *Weighted*.

Metode penggabungan antara dua atau lebih metode yang menghasilkan nilai prediksi pada masing-masing barang / jasa yang direkomendasikan. Penelitian terkait yang pernah di lakukan oleh Huizhi Liang, Yue Xu, Yuefeng Li, Richi Nayak dan Gavin Shaw dengan judul *A Hybrid Recommender Systems based on Weighted Tags*. Pada penelitian tersebut mereka menggabungkan *collaborative filtering* dan *social tags*. (Liang et al., 2011)

2) *Switching*.

Metode penggabungan antara dua atau lebih metode yang penggunaannya tergantung dari keadaan parameter. Metode ini dapat dengan cerdas melakukan perpindahan antara pendekatan satu dengan pendekatan lain dengan berbagai criteria. Kelebihan dari pendekatan ini adalah merupakan metode yang *efficient* dalam penggunaan kekuatan dan kelemahan dari sistem rekomendasi konstitusional. (Ghazanfar and Prugel-Bennett, 2010)

3) *Mixed*.

Penggabungan metode *content-based* dan *collaborative-filtering* yang lebih menitik beratkan pada metode *content-based* dibanding dengan *collaborative-filtering.* Dimana *content-based* berperan sebagai pendekatan yang memberikan keunggulana akan suatu barang dan *collaborative-filtering* menambahkan apabila terdapat penilaian dari pengguna lain.

4) *Feature Combination*.

Memungkinkan sistem rekomendasi mendapatkan hasil dengan keterkaitan yang tidak terlalu ekslusif sehingga beberapa barang / jasa yang berkaitan dapat saling melengkapi dan tidak esklusif, sebalikan pendekatan hybrid dengan cara ini dapat menyebabkan buramnya collaborative-filtering yang digunakan.

Pada penelitian ini sisem rekomendasi yang akan dibangung diharapkan dapat memiliki kriteria seperti dibawah :

1. Dapat menghasilkan daftar barang yang direkomendasikan sebelum dan setelah adanya rating data dari pengguna.
2. Menambah tingkat akurasi dari daftar barang yang direkomendasikan oleh sistem melebihi pendekatan *Content-based* dan *Collaborative-Filtering*.
3. Daftar barang yang direkomendasikan dapat menyesuaikan tergantung dari parameter yang membangunnya. Misalnya pada saat user belum melakukan login pendekatan yang digunakan berdasarkan parameter *demographic* atau *Rating System*, sedangkan pada saat login berlaku pendekatan *Content-based* dan *Collaborative-Filtering*.

Data-data pendukung terbentuknya *recommender system* dengan metode *Hybrid* ini nantinya akan didapat dari penilaian pengguna lain terhadap iklan/posting, penilaian reputasi user yang membuat iklan/posting, barang yang paling dicari dari *search engine*, data transaksi pengguna, data komentar pengguna dan profile pengguna baik dari internal sistem maupun *social media connection*.

Recommender System yang akan dibangun merupakan penggabungan dari beberapa pendekatan yang disesuaikan dengan keadaan pengguna, Tabel 2.1 merupakan gambaran dari recommender system yang akan dibangun.

Tabel 2.1 *User Behavior* dan Pendekatan Pembentuk Sistem Rekomendasi

Parameter / Users	**User Baru**	**User Terdaftar Tanpa aktifitas**	**User Terdaftar dan sudah memiliki transaksi**
Collaborative-Filtering			✓
Content-based		✓	✓
Demographic	✓	✓	
User Ratings	✓	✓	
Similar product			✓

Dari tabel 2.1 terdapat beberapa segmen user :

1) *User* Baru : user / pengguna yang belum memiliki data pribadi maupun aktivitas.
2) *User* terdaftar namun belum memiliki aktifitas, sistem sudah memiliki data dari *user* tersebut namun *user* belum melakukan aktifitas seperti mencari, membeli atau memberi komentar terhadap suatu barang atau jasa.
3) User terdaftar dan sudah memiliki aktifitas.