

### **BAB III. Landasan Teori**

Bab ini akan menjelaskan mengenai dasar teori yang digunakan oleh penulis sebagai acuan dalam membuat sistem.

#### **3.1. Sistem Rekomendasi**

Menurut McGinty dan Smyth, Sistem Rekomendasi merupakan model aplikasi dari hasil observasi terhadap keadaan dan keinginan pelanggan. Sistem Rekomendasi memanfaatkan opini seseorang terhadap suatu barang dalam domain atau kategori tertentu, untuk membantu seseorang dalam memilih produk. Karena itu sistem rekomendasi memerlukan model rekomendasi yang tepat agar apa yang direkomendasikan sesuai dengan keinginan pelanggan, serta mempermudah pelanggan mengambil keputusan yang tepat dalam menentukan produk yang akan dibelinya (Purwanto, 2009). Sistem ini sifatnya hanya membantu user dalam menentukan atau mempertimbangkan suatu produk. Untuk keputusan yang sebenarnya akan kembali kepada user yang menentukannya. Terdapat variable input dan output dalam suatu sistem rekomedasi. Input adalah jawaban dari pertanyaan yang diajukan, sedangkan output nya adalah hasil rekomendasinya. Menurut (Burke,2007) Model sistem rekomendasi dikelompokkan beberapa pendekatan sistem rekomendasi menjadi enam, yaitu (Handayani, 2012):

- a. *Content-based* akan merekomendasikan suatu item dengan cara mencari tingkat kesamaan antara item yang sebelumnya.

- b. *Collaborative-filtering* akan merekomendasikan suatu item berdasarkan *rating* yang diberikan oleh pengguna lainnya.
- c. *Demographic* akan merekomendasikan item berdasarkan data sebaran pengguna pada wilayah tertentu.
- d. *Knowledge-based* akan merekomendasikan berdasarkan domain pengetahuan yang spesifik yang memerlukan seorang domain *expert* (pakar) untuk menentukan kualitas data-datanya.
- e. *Community-based* akan merekomendasikan suatu item berdasarkan preferensi yang dimiliki oleh orang lain dilingkungannya.
- f. *Hybrid* akan merekomendasikan dengan cara mengkombinasikan teknik-teknik rekomendasi yang sudah disebutkan sebelumnya.

### **3.2. Content-Based Filtering**

*Content-Based Filtering* merupakan sebuah metode untuk membuat rekomendasi dengan cara menganalisis deskripsi item yang telah dinilai oleh pengguna lain dan deskripsi item untuk direkomendasikan (Pazzani, 1999). Berbagai algoritma telah diusulkan untuk menganalisis dan menemukan beberapa konten untuk mendapatkan nilai numerik dan masalah regresi yang berfungsi sebagai dasar untuk membuat rekomendasi.

### **3.3. Collaborative Filtering**

*Collaborative Filtering* merupakan sebuah metode untuk membuat rekomendasi berdasarkan korelasi antara pengguna lain dengan pengguna sistem untuk memberikan

rating terhadap suatu item (Pazzani, 1999). Ada 2 macam algoritma *Collaborative* yaitu :

a. *User Based Collaborative Filtering*

Algoritma ini menggunakan teknik statistika untuk menemukan sekumpulan pengguna yang memiliki sejarah dan akan menjadi sasaran. Setelah sekumpulan tetangga terbentuk, sistem menggunakan algoritma yang berbeda untuk menggabungkan kesukaan dan akan menghasilkan predikis atau rekomendasi N-teratas kepada *active user*

b. *Item Based Collaborative Filtering*

Algoritma ini merupakan metode rekomendasi yang didasari atas adanya kesamaan antara pemberian rating terhadap suatu produk dengan produk yang menjadi sasaran. Tingkat kesamaan produk ini, kemudian dibagi dengan parameter kebutuhan untuk memperoleh nilai kegunaan produk. Nilai kegunaan produk tertinggi lah yang kemudian dijadikan rekomendasi (Purwanto, 2009).

Untuk menghitung metode *Item Based Collaborative Filtering* menggunakan rumus *Adjusted cosine similarity* dan *weighted sum* (Kurniawan, 2016).

*Adjusted cosine similarity* digunakan untuk mencari nilai kemiripan antar produk yang dibandingkan. Dengan tahapan sebagai berikut :

1. Mengecek antara nilai rating yang dimiliki oleh 2 produk
2. Mengecek kolom dan baris kedua produk. Jika salah satu diantara produk yang memiliki nilai rating maka tidak akan dapat dihitng kemiripannya.

3. cek kolom dan baris kedua produk, maka akan memperoleh nilai ratingnya
4. Pengecek kolom dan baris sampai baris terakhir
5. setelah didapatkan nilai rating antar produk selanjutnya adalah menghitung nilai kemiripan dari nilai rating yang sudah didapat.

Persamannya

$$S(i, j) = \frac{\sum_{u \in U} (R_{u,i} - \bar{R}_u) (R_{u,j} - \bar{R}_u)}{\sqrt{\sum_{u \in U} (R_{u,i} - \bar{R}_u)^2} \sqrt{\sum_{u \in U} (R_{u,j} - \bar{R}_u)^2}} \dots (3.1)$$

Keterangan:

- $S(i, j)$  = Nilai kemiripan antar item i dengan item j
- $u \in U$  = Nilai user yang mereteng baik item i maupun item j
- $R_{u, i}$  = Rating user u pada item i
- $R_{u, j}$  = Rating user u pada item j
- $\bar{R}_u$  = Nilai rating rata-rata user u

*Weighted sum* digunakan untuk mencari nilai prediksi produk yang akan di rekomendasikan kepada user. Tahap perhitungannya adalah sebagai berikut :

1. Kolom pertama dicari nilai rating yang tidak kosong. Setelah mendapatkan nilai rating maka dihitung dengan nilai rating pada produk selanjutnya
2. Sebelum dibandingkan dengan *similarity* dari kolom yang kosong apakah memenuhi syarat perhitungan
3. Selanjutnya pada kolom yang kosong dicari nilai prediksinya dengan persamaan sebagai berikut

$$P(u, j) = \frac{\sum_{i \in I} (R_{u,i} * S_{i,j})}{\sum_{i \in I} |S_{i,j}|} \quad \dots (3.2)$$

Keterangan :

$P(u, j)$  =Prediksi untuk *user*  $u$  pada item  $j$

$R_{(u,i)}$  =Rating *user*  $u$  pada item  $i$

$i \in I$  =Himpunan item yang mirip dengan item  $j$

$S_{I,j}$  =Nilai *Adjusted cosine similarity* yang akan dibandingkan

$|S_{I,j}|$  =Harga mutlak nilai *Adjusted cosine similarity* yang akan dibandingkan

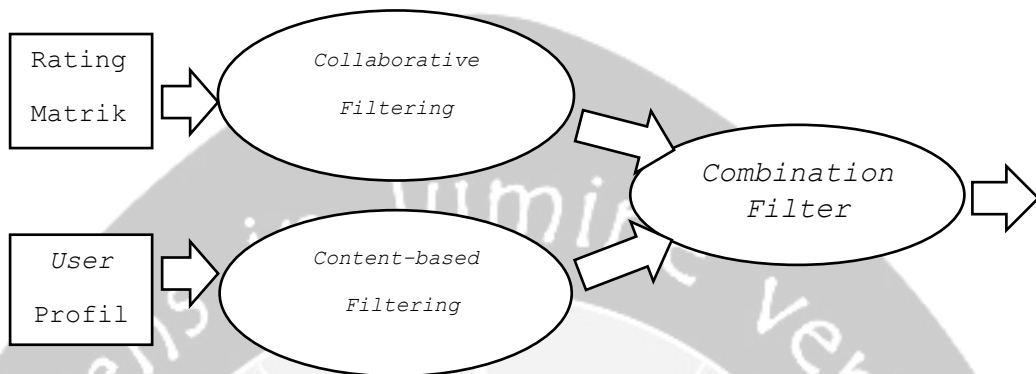
### 3.4. Sistem Rekomendasi Hybrid

Sistem rekomendasi *hybrid* merupakan suatu pendekatan yang menggabungkan lebih dari satu pendekatan rekomendasi untuk mengatasi masalah kekurangan masing-masing pendekatan yang nantinya diharapkan akan menghasilkan sebuah rekomendasi yang lebih akurat dan lebih baik lagi. Metode ini masih dibagi lagi menjadi beberapa bagian yaitu :

- a. Penggabungan secara linier (*Hybrid Linear Combination*)

Penggabungan ini menggabungkan hasil prediksi (rating) dari metode *Content-Based* dan *Collaborative*. Penggabungan ini dilakukan dengan cara pemberian ranking atau rating. Gambar 3.1 merupakan alur penggabungan secara linear. Penggabungan ini menggunakan rating matrik dan user profil yang didapat dari database kemudian di olah menggunakan metode *Collaborative* dan *Content-Based*

*Filtering*. Hasil masing-masing dari metode di gabungkan menjadi satu menjadi hasil sebuah *combination filter*



Gambar 3.1. *Hybrid Linear Combination* (Qing Li, 2003)

b. Penggabungan secara sekuensial (*Sequential Combination*)

Penggabungan ini adalah melakukan perhitungan pada salah satu metode (misalkan *Content-Based*) kemudian hasilnya digabungkan dengan metode lainnya (misalkan *Collaborative*). Penggabungan *Hybrid* ini mengintegrasikan informasi item dan rating pengguna untuk menghitung kemiripan item-item. Metode *Hybrid* ini merupakan sebuah metode yang menerapkan penggabungan *Hybrid Recommender System* dengan tujuan untuk meningkatkan akurasi prediksi pendekatan *Collaborative Filtering* dan menangani masalah item baru yang belum dirating (*Cold-Star Problem*).

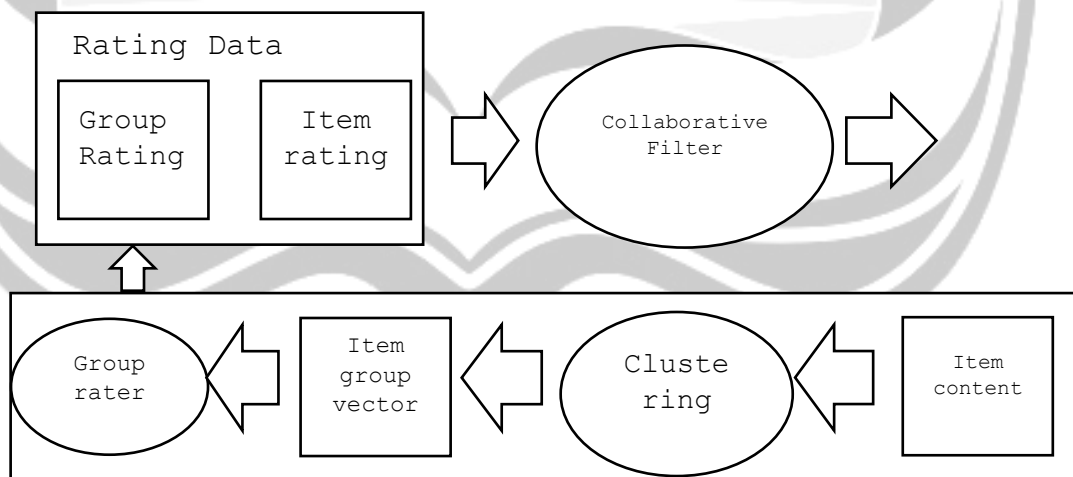
Gambar 3.2 merupakan sequential combination. Data user profil didapat dari database kemudian diolah menggunakan *Content-Based Filtering*, hasilnya kemudian menjadi input metode *Collaborative Filtering* maka akan didapat hasil kombinasinya.



Gambar 3.2. Sequential Combination (Qing Li, 2003)

c. Penggabungan Metode Item Based Clustering Hybrid

Penggabungan ini mengintegrasikan informasi item dan rating pengguna untuk menghitung kemiripan item-item. Penggabungan ini digambarkan pada gambar 3.3



Gambar 3.3. Item Based Clustering Hybrid (Qing Li, 2003)

Burk, R. (2002) memperkenalkan taksonomi untuk sistem rekomendasi *hybrid*.

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

### **3.5. Algoritma Nearest Neighbor**

Menurut Widarsana, Algoritma *Nearest Neighbor* merupakan salah satu metode penyelesaian masalah dengan melakukan klasifikasi terhadap objek berdasarkan jarak



paling dekat dengan objek tersebut (Hayati, 2011). Menurut Kusrini dan Emha (2009) algoritma *Nearest Neighbor* adalah pendekatan untuk mencari kasus dengan menghitung kedekatan antara kasus baru dengan kasus lama dengan berdasarkan pada pencocokan bobot dari sejumlah fitur yang ada yang memiliki kesamaan (*similarity*) (Hayati, 2011). Tujuan dari algoritma ini adalah untuk mengklasifikasikan objek baru berdasarkan atribut dengan training sample. *Classifier* tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Adapun rumus untuk menghitung algoritma ini yaitu :

$$\text{Similarity (T, S)} = \frac{\sum_{i=1}^n F(T_i, S_i)' * W_i}{W_i} \dots (3.3)$$

Keterangan:

T =kasus baru

S =kasus yang ada dalam memori (penyimpanan)

n =jumlah atribut dalam setiap kasus

I =atribut individu antara 1 s/d n

F = fungsi similiary atribut I antara kasus T dan kasus S

W =bobot yang akan diberikan pada atribut ke-i

### 3.6. Pariwisata

Wisata adalah kegiatan perjalanan yang dilakukan oleh seseorang atau sekelompok orang dengan mengunjungi tempat tertentu untuk tujuan rekreasi, pengembangan pribadi, atau mempelajari keunikan daya tarik wisata yang dikunjungi dalam jangka waktu sementara (UU RI NO 10 Pasal 1, 2009). Sedangkan pariwisata adalah berbagai

macam kegiatan wisata dan didukung berbagai fasilitas serta layanan yang disediakan oleh masyarakat, pengusaha, Pemerintah, dan Pemerintah Daerah. Kawasan Strategis Pariwisata adalah kawasan yang memiliki fungsi utama pariwisata atau memiliki potensi untuk pengembangan pariwisata yang mempunyai pengaruh penting dalam satu atau lebih aspek, seperti pertumbuhan ekonomi, sosial dan budaya, pemberdayaan sumber daya alam, daya dukung lingkungan hidup, serta pertahanan dan keamanan. (UU RI No 10 pasal 1, 2009). Wisatawan adalah orang yang melakukan wisata (UU RI no 10 tahun 2009). Setiap orang dari kecil, tua maupun muda serta laki-laki maupun perempuan yang melakukan kegiatan wisata termasuk dalam kategori wisatawan.

### **3.7. PHP**

*PHP Hypertext Preprocessor* (PHP) adalah bahasa pemrograman *server-side* yang digunakan untuk menyusun dokumen HTML, sehingga website yang dihasilkan menjadi dinamis (menyesuaikan dengan masukan yang didapat) (Peranginangin, 2006). Kode-kode PHP dapat disisipkan diantara kode HTML, dan sebelum dokumen diminta untuk dikirimkan, maka kode-kode akan dieksekusi, dan hasilnya akan disatukan dengan dokumen HTML, sehingga kodenya tidak akan ikut ditampilkan. Implementasi dari aplikasi ini yang dapat berjalan pada browser dengan bantuan internet, pada semua platform bisa mengakses.

### 3.8. MySQL

Merupakan database sebagai penyimpanan data dan basis pengetahuan yang akan digunakan dalam sistem. Database MySQL sangat cocok berpasangan dengan PHP (Peranginangin, 2006). Kelebihan MySQL antara lain sebagai berikut (Miftakhul Huda):

a. *Portability*

MySQL dapat berjalan stabil pada berbagai sistem operasi seperti Windows, Linux, FreeBSD, Mac OS X Server, Solaris, Amiga, dan masih banyak lagi.

b. *Open Source*

MySQL didistribusikan secara *open source* (gratis), dibawah lisensi GPL sehingga dapat digunakan secara cuma-cuma.

c. *Multiuser*

MySQL dapat digunakan oleh beberapa *user* dalam waktu yang bersamaan tanpa mengalami masalah atau konflik

d. *Performance tuning*

MySQL memiliki kecepatan yang menakjubkan dalam menangani *query* sederhana, dengan kata lain dapat memproses lebih banyak SQL per satuan waktu.

e. *Column types*

MySQL memiliki tipe kolom yang sangat kompleks, seperti *signed / unsigned integer, float, double, char, text, date, timestamp*, dan lain-lain.

f. *Command dan functions*

MySQL memiliki operator dan fungsi secara penuh yang mendukung perintah *Select* dan *Where* dalam *query*.

g. Security

MySQL memiliki beberapa lapisan keamanan seperti *level subnetmask*, nama *host*, dan izin akses *user* dengan sistem perizinan yang mendetail serta *password* terenkripsi.

h. Scalability dan limits

MySQL mampu menangani database dalam skala besar, dengan jumlah *records* lebih dari 50 juta dan 60 ribu tabel serta 5 milyar baris. Selain itu batas indeks yang dapat ditampung mencapai 32 indeks pada tiap tabelnya.

i. Connectivity

MySQL dapat melakukan koneksi dengan *client* menggunakan protokol TCP/IP, *Unix socket* (UNIX), atau *Named Pipes* (NT).

j. Localization

MySQL dapat mendeteksi pesan kesalahan pada *client* dengan menggunakan lebih dari dua puluh bahasa. Meskipun demikian, bahasa Indonesia belum termasuk di dalamnya.

k. Interface

MySQL memiliki *interface* (antar muka) terhadap berbagai aplikasi dan bahasa pemrograman dengan menggunakan fungsi API (*Application Programming Interface*).

l. Clients dan tools

MySQL dilengkapi dengan berbagai *tool* yang dapat digunakan untuk administrasi *database*, dan pada setiap *tool* yang ada disertakan petunjuk *online*.

m. Struktur tabel

MySQL memiliki struktur tabel yang lebih fleksibel dalam menangani *ALTER TABLE*, dibandingkan database lainnya semacam PostgreSQL ataupun Oracle.

### **3.9. Google Maps Application Programming Interface (GM API)**

Menurut Sirenden & Dachi, *Google Maps Application Programming Interface* merupakan framefork yang merupakan bagian dari Google. API bisa di artikan sebagai kode program yang merupakan antar muka atau penghubung antara aplikasi yang dibuat oleh user dengan fungsi-fungsi yang dikerjakan. Pada GM, titik koordinat suatu tempat ditunjukkan dengan sistem koordinat geografis (Sibuea, 2015). Ada dua cara untuk mengakses data Google Maps (GM), tergantung dari data yang ingin kita ambil dan parsing (uraikan) dari GM yaitu mengakses data GM tanpa menggunakan API Key dan mengakses data GM menggunakan API Key.