

## BAB III

### LANDASAN TEORI

#### 3.1 Sistem Rekomendasi

Sistem rekomendasi merupakan suatu aplikasi untuk menyediakan dan merekomendasikan suatu item dalam membuat suatu keputusan yang diinginkan oleh pengguna. Konsep sistem rekomendasi telah digunakan secara luas oleh hampir semua area bisnis dimana seorang konsumen memerlukan informasi untuk membuat suatu keputusan (Utomo & Anggriawan, 2015). Sistem rekomendasi banyak diterapkan pada aplikasi *e-commerce* untuk memberikan rekomendasi kepada pengguna dengan memprediksi item yang kemungkinan akan disukai oleh pengguna tersebut berdasarkan penilaian terhadap suatu item atau berdasarkan kriteria yang diinginkan oleh pengguna.

Sistem rekomendasi adalah sub kelas dari sistem penyaringan informasi yang berusaha untuk memprediksi rating atau preferensi yang akan pengguna berikan kepada suatu item seperti: musik, buku, film, atau elemen sosial seperti orang atau kelompok yang belum mereka pertimbangkan dengan menggunakan model yang dibangun dari karakteristik suatu item (*content-based approaches*) atau lingkungan sosial pengguna (*collaborative filtering approaches*). Meskipun banyak pendekatan yang berbeda untuk sistem rekomendasi yang telah dikembangkan dalam beberapa tahun lalu, minat terhadap pengembangan sistem rekomendasi sampai saat ini masih tetap tinggi dikarenakan permintaan pada aplikasi praktis yang dapat memberikan rekomendasi pribadi untuk memperoleh suatu informasi secara cepat didalam banyaknya kumpulan informasi yang ada. Semua itu merupakan tuntutan pertumbuhan yang menimbulkan beberapa tantangan kunci dalam pengembangan sistem rekomendasi sehingga muncul metode-metode yang dapat digunakan untuk menyelesaikan masalah tersebut seperti *content-based filtering*, *clustering-based filtering*, *collaborative filtering*, atau *hybrid*, dan lain-lain (Sharma & Gera, 2013).

Dalam pengembangan sistem rekomendasi, ada beberapa metode yang secara umum paling sering digunakan antara lain: *content-based filtering*,

*collaborative filtering*, dan *hybrid*. Dari ketiga metode tersebut, masing-masing memiliki perbedaan dalam memproses suatu masukan yang akan digunakan untuk memberikan suatu rekomendasi. Metode *content-based filtering* memberikan rekomendasi dengan cara membangun profil pengguna dan profil dari suatu item. Metode *collaborative filtering* menggunakan sistem rating, *feedback*, atau ulasan terhadap suatu item untuk memberikan rekomendasi. Metode *hybrid-based* menggabungkan beberapa metode menjadi satu dengan tujuan untuk melengkapi kekurangan dari metode-metode yang digunakan.

### **3.2 Content-based filtering**

*Content-based filtering* merupakan salah satu metode yang digunakan dalam sistem rekomendasi. Metode ini memberikan suatu rekomendasi berdasarkan korelasi antara isi (*content*) dari sebuah item dengan preferensi pengguna sebagai pembanding (Maharani & Gunawan, 2015). Dengan memberikan masukan berupa profil item yang diinginkan maka dapat dibentuk profil pengguna yang bersangkutan, dari profil pengguna inilah kemudian sistem dapat menganalisa karakteristik dan minat dari pengguna tersebut. Sistem akan memberikan rekomendasi kepada user dengan membandingkan karakteristik profil item yang ada dengan profil pengguna yang merupakan profil item yang disukai oleh pengguna tersebut. Misalkan, pengguna (*user*) A memberikan kriteria item yang disukainya dengan spesifikasi X, setelah dianalisa ternyata item Y mempunyai spesifikasi yang sama atau mirip dengan item yang memiliki spesifikasi X, maka item Y akan direkomendasikan kepada pengguna. Asumsi dasar yang digunakan adalah apabila pengguna menyukai spesifikasi X maka dia juga pasti akan menyukai item yang memiliki spesifikasi yang serupa, misalnya item Y.

Dalam penelitian ini, rekomendasi yang diberikan dari *content-based filtering* diperoleh berdasarkan dari kemiripan antara preferensi telepon genggam yang diinginkan pengguna dengan spesifikasi telepon genggam yang tersedia dalam sistem. Untuk menghitung nilai kemiripan atau *similarity* antara telepon genggam yang diinginkan pengguna dengan telepon genggam yang ada dalam

sistem, maka terlebih dahulu dihitung jarak (*distance*) antara profil telepon genggam yang diinginkan pengguna dengan telepon genggam lain yang ada dalam sistem dengan menggunakan persamaan sebagai berikut:

$$Dis(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Dimana:

$n$  = Jumlah elemen dalam vektor telepon genggam

$x_i$  = Nilai parameter ke- $i$  dari vektor profil pengguna

$y_i$  = Nilai parameter ke- $i$  dari vektor profil telepon genggam yang dihitung distancenya

Setelah nilai *distance* diperoleh, maka nilai kemiripan atau *similarity* dihitung berdasarkan rumus berikut:

$$sim(x, y) = \frac{1}{1 + Dis(x, y)}$$

Kriteria preferensi telepon genggam yang diinginkan (profil pengguna) diperoleh dari masukan pengguna berupa spesifikasi telepon genggam yang diinginkan. Dalam penelitian ini, telepon genggam dengan nilai *similarity* tertinggi akan direkomendasikan kepada user.

### 3.3 Structured Query Language

*Structured Query Language (SQL)* adalah bahasa pemrograman *database* komputer yang dirancang untuk mengelola data dalam basis data relasional. *Structured Query Language* merupakan bahasa pemrograman *database* komputer standar yang digunakan untuk mengubah dan mendefinisikan basis data relasional menggunakan pernyataan deklaratif (Groff, 2010).

### 3.4 Precision, Recall dan Accuracy

*Precision dan Recall* digunakan untuk mengukur kinerja sistem. *Precision* adalah kecocokan antara bagian data yang diambil dengan informasi yang

dibutuhkan. *Recall* merupakan tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. *Accuracy* adalah tingkat kedekatan antara nilai yang didapat terhadap nilai sebenarnya. *Precision*, *recall* dan *accuracy* dapat dihitung menggunakan *confusion matrix* dapat dilihat pada tabel 3.1.

**Tabel 3 1 Confusion Matrix**

	True	False
True (Positive)	TP (True Positive) <i>Correct result</i>	FP (False Positive) <i>Unexpected result</i>
False (Negative)	FN (False Negative) <i>Missing result</i>	TN (True Negative) <i>Correct absence of result</i>

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Keterangan :

TP = Banyak data dengan nilai sebenarnya positif dan nilai prediksi positif

FP = Banyak data dengan nilai sebenarnya negatif dan nilai prediksi positif

FN = Banyak data dengan nilai sebenarnya positif dan nilai prediksi negatif

TN = Banyak data dengan nilai sebenarnya negatif dan nilai prediksi negatif

Ukuran besaran *precision*, *recall*, dan *accuracy* biasanya diberi nilai dalam bentuk presentase antara 1 sampai 100%. Sebuah sistem akan dianggap baik jika tingkat *precision*, *recall*, dan *accuracy*-nya tinggi.