

## **BAB II**

### **LANDASAN TEORI**

#### **2.1 Jersey**

*Jersey* menurut kamus bahasa Inggris terjemahan Indonesia, arti kata *jersey* adalah kain rajutan atau baju kaos. *Jersey* merupakan sebuah seragam untuk pembeda tim dalam berbagai macam cabang olahraga. Olahraga sepak bola salah satu cabang olahraga populer yang semakin banyak peminatnya, bukan hanya pemain melainkan banyak orang menyukai tim sepak bola tersebut. Dalam perkembangannya sepak bola menjadikan sebuah tim identik dengan ciri khasnya masing-masing antara lain yang paling terlihat adalah desain *jersey* pada sebuah tim sepak bola. Tim sepak bola tertentu dengan ciri khas warna dapat membedakan dengan tim lain, disini mulai terlihat dari awal perkembangan *jersey* bola yang sudah banyak memiliki motif serta bahan dan warna yang menarik pencinta sepak bola atau penggemar sepak bola yang rela mengeluarkan uang banyak untuk memesan jersey bola tim kesayangannya (Agustyan & Soeprapto, 2018). Desain jersey menjadi pilihan pertama untuk digunakan dalam berpenampilan berolahraga, hanya saja ada beberapa yang sulit menentukan desain jersey yang diinginkan (Ritandza P, Zam, 2018).

#### **2.2 Sistem Rekomendasi**

Sistem rekomendasi adalah suatu sistem yang dirancang untuk memprediksi suatu item yang sesuai dengan minat user, yang mana item tersebut akan direkomendasikan pada user (Sanjung, 2011). Oleh karena itu sistem rekomendasi memerlukan model rekomendasi yang tepat agar yang direkomendasikan sesuai dengan keinginan pelanggan, serta mempermudah pelanggan mengambil keputusan yang tepat dalam menentukan produk yang akan digunakannya (Smyth & Lorraine, 2016). Sistem rekomendasi sendiri terbagi menjadi tiga metode, yaitu *content-based filtering*, *hybrid approached*, dan *collaborative filtering*.

Rekomendasi yang diberikan oleh *content-based filtering* membentuk profil penggunaannya berdasarkan atribut pembentuk suatu item. Sebagai contohnya untuk suatu dokumen, atribut pembentuknya adalah kata-kata atau *term* yang terdapat

pada dokumen tersebut. Parameter pembentuk profil pengguna ini juga diberikan nilai bobot berdasarkan kriteria tertentu, misalnya untuk kasus dokumen adalah frekuensi suatu kata atau *term* dalam dokumen tersebut (Adi, 2010). *Content-based filtering* juga sering disebut dengan nama *cluster model*, dimana *cluster model* menemukan pengguna yang sama dengan pengguna yang berada dalam setiap segment yang telah diklasifikasikan (Linden, 2003). Kelebihan sistem rekomendasi dengan pendekatan *content-based filtering* adalah memiliki kemampuan untuk merekomendasikan item (contoh: film, lagu, artikel dll) yang sifatnya baru bagi user, karena prinsip kerjanya yaitu dengan melihat diskripsi konten yang dikandung oleh item yang pernah diberi nilai rating tinggi sebelumnya oleh pengguna. *Content-based filtering* juga dapat menjadikan item baru untuk langsung direkomendasikan kepada pengguna tanpa harus menunggu pengguna-pengguna lain untuk melakukan rating pada item tersebut. Namun sistem rekomendasi *content based filtering* ini juga memiliki kelemahan yaitu penyaringan berbasis konten sulit untuk menghasilkan rekomendasi yang tidak teduga (*serendipitous recommendation*), karena semua informasi dipilih dan direkomendasikan berdasarkan konten dan yang kedua adalah tidak *reliable* (handal) tidak adanya ide, pendapat serta masukan dari pengguna sebelumnya yang sulit dianalisis oleh komputer sehingga sulit untuk pengguna pemula memakai sistem ini secara efektif (Li & Kim, 2002).

Sedangkan *hybrid approached* digunakan untuk menggambarkan setiap sistem rekomendasi yang menggabungkan beberapa teknik rekomendasi untuk menghasilkan sebuah *output*. *Hybrid approached* merupakan gabungan hasil prediksi dari *collaborative filtering* dan *content-based filtering* yang berfungsi untuk mengurangi batas dari rekomendasi yang dapat diberikan oleh kedua sistem rekomendasi dan meningkatkan *performance* dari rekomendasi tersebut. Kelebihan dari *hybrid approached* adalah metode ini dapat menggunakan yang terbaik dari masing-masing metode lain untuk mengatasi kelemahan. Misalnya, memanfaatkan *collaborative filtering* untuk membantu *content based filtering* dalam mengatasi atau menghindari *overspecialization* (dimana *content-based* harus selalu menyediakan tingkat batas baru untuk menyesuaikan dengan fitur

profil dan item), dan juga memanfaatkan *content-based filtering* untuk membantu *collaborative filtering* yang berhubungan dengan *cold-start* (apabila ada item atau user baru, item tersebut tidak dapat langsung dijadikan rekomendasi karena harus menunggu pengguna lain untuk memberikan rating terlebih dahulu). Sedangkan kelemahan *hybrid approached* adalah meningkatnya kompleksitas dari sistem rekomendasi tersebut dan juga menurunkan kecepatan dikarenakan lebih dari satu metode yang digunakan pada saat yang bersamaan (Devi, 2015).

Berikut adalah penjelasan mengenai metode *collaborative filtering* yang akan digunakan dalam penelitian.

### **2.3 Collaborative Filtering**

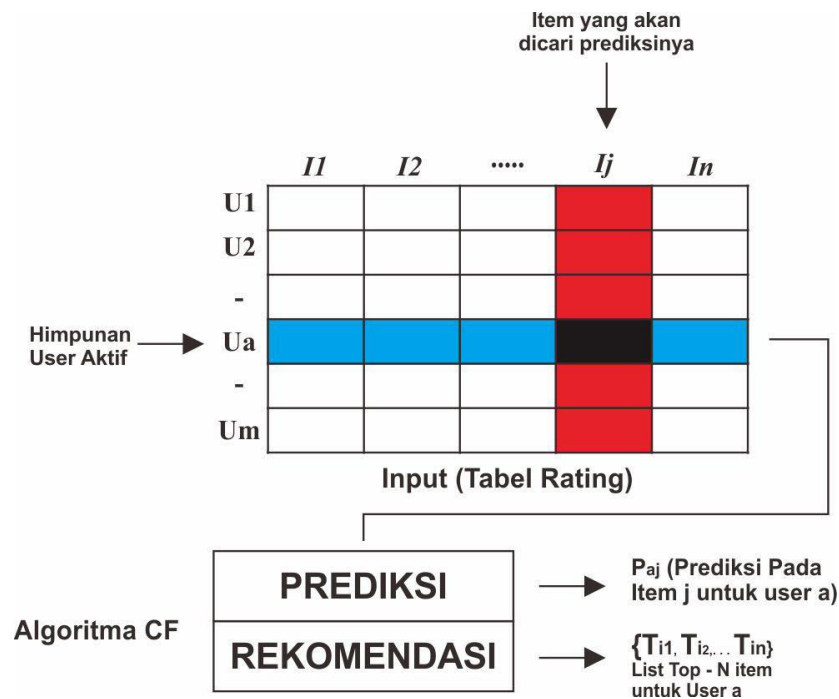
Sistem rekomendasi dengan metode *collaborative filtering* adalah metode yang digunakan untuk memprediksi kegunaan item berdasarkan penilaian pengguna sebelumnya. *Collaborative filtering* (CF) adalah suatu konsep dimana opini dari pengguna lain yang ada digunakan untuk memprediksi item yang mungkin disukai/diminati oleh seorang pengguna. Kualitas rekomendasi yang diberikan dengan menggunakan metode ini sangat bergantung dari opini pengguna lain (*neighbor*) terhadap suatu item. Belakangan diketahui bahwa melakukan reduksi *neighbor* (yaitu dengan memotong *neighbor* sehingga hanya beberapa pengguna yang memiliki kesamaan/*similarity* tertinggi sajalah yang akan digunakan dalam perhitungan) mampu meningkatkan kualitas rekomendasi yang diberikan (Hartijah, 2019).

*Collaborative filtering* memberikan rekomendasi berdasarkan kumpulan dari pendapat, minat dan ketertarikan beberapa user yang biasanya diberikan dalam bentuk rating pada suatu item. Setiap pengguna akan mendapatkan rekomendasi dengan melakukan dua hal utama dalam menyelesaikan algoritma *collaborative filtering*, yaitu:

- *Prediction*, melakukan prediksi opini yang akan diberikan oleh pengguna dalam skala bilangan yang sama.
- *Recommendation*, memberikan rekomendasi berupa daftar item dengan nilai prediksi tertinggi. Hal yang perlu dicatat adalah item-item yang

direkomendasikan belum pernah dibeli, dilihat atau di-rating oleh pengguna tersebut.

Berikut merupakan skema dari *Collaborative filtering* dalam pemberian rekomendasi kepada user aktif



Gambar 2.1 Skema rekomendasi *Collaborative filtering*

Keterangan:

$U_a$  : Himpunan *user* aktif

$I_j$  : *Item* yang akan dicari prediksinya

CF : *Collaborative Filtering*

Kelebihan metode *collaborative filtering* adalah metode ini dapat bekerja meskipun konten yang berhubungan dengan item atau *user* sangat sedikit atau bahkan tidak ada. Karena metode *collaborative filtering* tidak memberikan rekomendasi berdasarkan analisa dari konten-konten yang ada, yang biasanya berupa informasi item ataupun profil pengguna. Selain itu, metode *collaborative filtering* dapat mengatasi masalah *content based filtering* yang kesusahan dalam menganalisa gambar, video, ataupun musik, karena metode ini bekerja berdasarkan histori yang ada sebelumnya. Selain itu, kualitas prediksi lebih akurat

dan algoritma yang digunakan juga relatif sederhana untuk diterapkan. Namun, metode *collaborative filtering* juga mengalami masalah penting yang dinamakan *cold start problem*, yang dibagi menjadi dua jenis yaitu *cold start users* dan *cold start items*. (Hartijah, 2019) Dimana ketika terdapat item atau pengguna baru, item tersebut tidak dapat langsung dijadikan rekomendasi karena harus menunggu pengguna untuk memberikan rating terlebih dahulu, dan apabila ada pengguna baru, pengguna baru ini tidak dapat diberikan rekomendasi sebelum melakukan rating. *Collaborative filtering* memiliki dua pendekatan utama, yaitu *user-based collaborative filtering* dan *item-based collaborative filtering*.

### **2.3.1 User-Based Collaborative Filtering**

Pendekatan *user based collaborative filtering* sistem memberikan rekomendasi kepada user item-item yang disukai atau dirating oleh user – user lain yang memiliki banyak kemiripan dengannya. Misalnya, user a menyukai atau merating item 1,2 dan 3, kemudian user b menyukai item 1,2 dan 4 maka sistem akan merekomendasikan item 3 kepada user b dan item 4 kepada user a. Kelebihan dari pendekatan user based collaborative filtering adalah dapat menghasilkan rekomendasi yang berkualitas baik. Sedangkan kekurangannya adalah dapat menghasilkan rekomendasi yang berkualitas baik. Sedangkan kekurangannya adalah kompleksitas perhitungan akan semakin bertambah seiring dengan bertambahnya pengguna sistem, semakin banyak pengguna (user) yang menggunakan sistem maka proses perekomendasi akan semakin lama (Hartijah, 2019).

### **2.3.2 Item-Based Collaborative Filtering**

Pendekatan *item based collaborative filtering* memberikan rekomendasi berdasarkan kemiripan antar item. Metode ini merupakan metode rekomendasi yang didasari atas adanya kesamaan antara pemberian rating terhadap suatu item dengan item yang pernah dirating user lain. Item yang telah di rating oleh user akan menjadi patokan untuk mencari sejumlah item lainnya yang berkorelasi dengan item yang telah dirating user. Motivasi kunci dibalik metode ini adalah user akan cenderung menyukai item yang sejenis atau mempunyai korelasi dengan item yang telah disukainya (Hartijah, 2019).

Secara umum proses pemberian rekomendasi pada *collaborative filtering* terdiri atas 2 langkah, yaitu: Penemuan *similar* item yakni menghitung nilai

kemiripan antar item satu dengan item yang lainnya berdasarkan rating yang telah diberikan pengguna dan Penghitungan prediksi rating yang setelah mendapatkan semua nilai prediksinya akan dipilih nilai tertinggi yang akan dijadikan rekomendasi. Misalnya, terdapat item A dan item B yang telah diberi rating oleh pengguna X, kemudian pengguna Y melakukan rating pada item B,C, dan D. maka pengguna X akan mendapatkan rekomendasi diantara item C dan D yang memiliki nilai prediksi tertinggi.

Metode yang digunakan untuk penemuan *similarity* item adalah metode *adjusted cosine similarity*, sedangkan untuk perhitungan prediksinya menggunakan metode *simple weighted average*.

### 2.3.2.1 Adjusted Cosine Similarity

Selain *person corellation*, *adjusted cosine* adalah algoritma yang digunakan untuk menghitung dan menentukan nilai kemiripan (*similarity*) antara 2 item atau 2 user yang berbeda. Algoritma *Adjusted cosine similarity* adalah algoritma yang dibuat untuk mengatasi kelemahan algoritma dasarnya, yaitu algoritma *cosine similarity*. Algoritma *cosine similarity* adalah algoritma yang juga dikenal sebagai algoritma *vector-based*. Formula algoritma ini menunjukkan dua buah item yang berhubungan dan ratingnya, lalu mendefinisikan kemiripan diantaranya sebagai sebuah sudut diantara vektor-vektornya (Ramadhan, et al., 2017).

Algoritma *Adjusted cosine similarity* berusaha mengatasi kelemahan algoritma *cosine similarity*, dimana dalam hal merating, *user* memiliki perbedaan skema. Ada yang member rating tinggi untuk sebuah item, lalu memberi rating rendah pada item lainnya. Lalu, untuk item yang sama, item tersebut diberi rating biasa dan rendah. Untuk menyeimbangkan nilai ratingnya, maka dihitunglah rata-rata dari masing masing *user*. Persamaan 2.1 berikut digunakan dalam perhitungan rata-rata untuk tiap *user*.

$$\bar{R}_u = \frac{\sum \text{nilai all item } (R_u)}{\text{jumlah item yang dirating}} \dots\dots\dots (2.1)$$

Selanjutnya persamaan 2.2 berikut merupakan persamaan dari algoritma *Adjusted cosine similarity* yang digunakan untuk menemukan nilai *similarity* antar item:

$$Sim(i, j) = \frac{\sum_{u \in U} (R_{u,i} - \bar{R}_u)(R_{u,j} - \bar{R}_u)}{\sqrt{\sum_{u \in U} (R_{u,i} - \bar{R}_u)^2} \sqrt{\sum_{u \in U} (R_{u,j} - \bar{R}_u)^2}} \dots\dots\dots (2.2)$$

keterangan :

$Sim(i, j)$  : nilai kemiripan atau *similarity* antara nilai  $i$  dan  $j$

$R_{ui}$  dan  $R_{uj}$  : semua pengguna yang melakukan rating pada item  $i$  dan  $j$

$R_u$  : rata-rata dari rating item setiap pengguna

Dimana  $sim_{(i,j)}$  adalah nilai kemiripan antara  $i$  dan  $j$ . Sedangkan  $R_{u,i}$  dan  $R_{u,j}$  adalah semua pengguna yang melakukan rating pada item  $i$  dan  $j$ , sedangkan  $R_u$  adalah rata-rata dari rating item setiap pengguna. Dalam menghitung nilai kemiripan, nilai yang akan dihasilkan oleh algoritma ini berkisar antara 1 dengan -1. Semakin nilai kemiripan mendekati -1, maka nilai kemiripan kedua item bertolak belakang atau tidak memiliki kemiripan, dan apabila nilai kemiripan mendekati angka 1, maka kedua item tersebut cenderung mirip satu sama lainnya (Hartijah, 2019).

### 2.3.2.2 Simple Weighted Average

Kegunaan algoritma *simple weighted average* adalah untuk memprediksi nilai yang akan diberikan oleh pengguna pada suatu item yang belum pernah dirating sebelumnya dan kemudian nilai prediksi tersebut akan digunakan sebagai nilai rekomendasi untuk pengguna, dimana nilai yang diambil yaitu nilai tertinggi atau nilai yang lebih dari sama dengan nol. Berikut merupakan persamaan dari algoritma *simple weighted average* (Devi, 2015).

$$P(u, i) = \frac{\sum \text{all similar item}, N(S_{i,N} \times R_{u,N})}{\sum \text{all similar item}, N(|S_{i,N}|)} \dots\dots\dots 2.3)$$

keterangan :

$P_{u,i}$  : nilai prediksi (bobot) yang dihasilkan oleh pengguna  $u$  pada item  $i$

$S_{i,N}$  : nilai kemiripan yang telah dihitung sebelumnya

$R_{u,N}$  : rating dari pengguna  $u$  untuk item  $N$

Dimana  $P_{u,i}$  adalah nilai prediksi (bobot) yang dihasilkan oleh pengguna  $u$  pada item  $i$ . Sedangkan  $S_{i,N}$  adalah nilai kemiripan yang telah dihitung sebelumnya dan  $R_{u,N}$  adalah rating dari pengguna  $u$  untuk item  $N$ .

#### 2.4 Penelitian Sebelumnya

Penelitian sebelumnya dilakukan oleh (Sabani, 2020) yang telah melakukan penelitian “Sistem Rekomendasi Peminjaman Buku Perpustakaan Menggunakan Metode *Item-Based Collaborative Filtering*”. Pada penelitian ini disimpulkan bahwa hasil rekomendasi peminjaman buku tingkat error rating prediksi yang didapat sebesar 50% atau 0,5 berdasarkan 30 data buku dan 30 data mahasiswa UIN Maulana Malik Ibrahim Malang.

Penelitian kedua dilakukan oleh (Prasetyo, et al., 2019) yang telah melakukan penelitian “Implementasi Metode *Item-Based Collaborative Filtering* dalam Pemberian Rekomendasi Calon Pembeli Aksesoris *Smartphone*”. Pada penelitian ini disimpulkan bahwa hasil rekomendasi pada kasus penjualan aksesoris *smartphone* sebagai rekomendasi untuk user cukup baik, meskipun pemberian rating minim. Dibuktikan dengan kecilnya rata-rata nilai *Mean Absolute Error* (MAE) yaitu 0,572039 namun untuk proses eksekusi. Dari segi waktu yang dibutuhkan cukup lama yaitu 6,4 detik atau 6,4 *second*.

Penelitian ketiga dilakukan oleh (Sari, et al., 2020) yang telah melakukan penelitian “Pembuatan Sistem Rekomendasi Film dengan Menggunakan Metode *Item Based Collaborative Filtering* pada *Apache Mahout*”. Pada penelitian ini disimpulkan bahwa hasil rekomendasi menghasilkan nilai akurasi yang sangat memuaskan dengan besar 97%. Hal ini menunjukkan bahwa metode ini merupakan metode yang sangat tepat untuk kasus rekomendasi film. Akan tetapi pada pengembangannya masih memiliki kekurangan yaitu dalam pemrosesan data berukuran besar masih memakan waktu yang cukup lama.

Penelitian keempat dilakukan oleh (Setiawan, et al., 2019) yang telah melakukan penelitian “Implementasi *Item-Based Collaborative Filtering* dalam Pemberian Rekomendasi Agenda Wisata Berbasis Android”. Pada penelitian ini disimpulkan bahwa hasil rekomendasi pariwisata Bengkulu berbasis Android



dapat dibangun dengan tingkat kepuasan user experiences sangat baik yaitu sebesar 90,752% pada pengujian usability testing.

Penelitian kelima dilakukan oleh (Indriawan, et al., 2020) yang telah melakukan penelitian “Sistem Rekomendasi Penjualan Produk Pertanian Menggunakan Metode *Item Based Collaborative Filtering*”. Pada penelitian ini disimpulkan bahwa hasil rekomendasi yang diperoleh lebih akurat dan hasil rekomendasi kepada tiap *customer* akan berbeda-beda sesuai dengan nilai rating yang diberikan oleh *customer* tersebut terhadap sebuah produk. Dengan menggunakan rumus perhitungan *adjust cosine similarity* dan *weighted sum* sistem dapat menghitung nilai kemiripan produk sekaligus memprediksikan nilai rating seorang *customer* terhadap sebuah produk. Nilai *similarity* yang diperoleh memiliki *range* nilai dari -1 sampai dengan +1, semakin besar nilai *similarity* maka semakin besar juga nilai kemiripan antara produk tersebut. Setelah menentukan batas *minimum* nilai *similarity* atau kemiripan produk yaitu  $\geq 0,7$  maka didapatkan beberapa prediksi rating sebuah produk, nilai kemiripan produk yang mempunyai nilai paling tinggi akan direkomendasikan kepada *customer*.

Penelitian keenam dilakukan oleh (Hartijah, 2019) yang telah melakukan penelitian “Sistem Rekomendasi Kuliner Khas Gresik Menggunakan Metode *Item Based Collaborative Filtering*”. Pada penelitian ini disimpulkan bahwa hasil rekomendasi Pada penelitian ini disimpulkan bahwa hasil rekomendasi yang membantu pengguna mengambil keputusan tempat kuliner yang sesuai preferensi.

