

# Sistem Rekomendasi Produk Aplikasi Marketplace Berdasarkan Karakteristik Pembeli Menggunakan Metode User Based Collaborative Filtering

Fajar Rohman Hariri<sup>1\*</sup>, Lingga Wahyu Rochim<sup>2</sup>

<sup>1,2</sup> Program Studi Teknik Informatika, Universitas Islam Negeri Maulana Malik Ibrahim, Malang, Jawa Timur  
Email: <sup>1\*</sup>fajar@ti.uin-malang.ac.id, <sup>2</sup>linggawahyurochim@gmail.com

(Naskah masuk: 25 Sep 2022, direvisi: 12 Okt 2022, 26 Okt 2022, diterima: 27 Okt 2022)

## Abstrak

Sistem rekomendasi produk merupakan sebuah sistem yang dapat memberikan prediksi produk yang relevan terhadap perilaku atau karakteristik *user*, sehingga dapat mempengaruhi *user* dalam mengambil keputusan untuk membeli suatu produk. Penelitian ini dilakukan untuk dapat memberikan rekomendasi kepada pembeli pada aplikasi *marketplace* Sindomall dengan menggunakan metode *User Based Collaborative Filtering* dikolaborasikan dengan algoritma *Improved Triangle Similarity Complemented with User Rating Preferences* (ITR) untuk menghitung nilai *similarity* antar pembeli dan algoritma *Weighted Sum* untuk menghitung nilai prediksi produk. Karakteristik pembeli diambil dari data perilaku pembeli dalam memberikan *rating* pada produk. Dalam pengujian model yang dilakukan dengan menggunakan data nilai prediksi pada 20 *user* acak pada *database* aplikasi Sindomall pada bulan Desember 2021 didapatkan model optimal dengan nilai parameter presentase *user* sebesar 100%. Hasil dari pengujian *error* sistem menggunakan model terpilih mendapatkan nilai MAE dan RMSE masing-masing sebesar 0,006 dan 0,013, sedangkan pada tahap pengujian akurasi sistem didapatkan nilai akurasi sebesar 0,849, nilai presisi sebesar 0,923, nilai *recall* sebesar 0,869, dan nilai *F-score* (F1) sebesar 0,895.

**Kata Kunci:** *User Based Collaborative Filtering*, Sistem Rekomendasi Produk, *Marketplace*, Karakteristik Pembeli

## *Marketplace Application Product Recommendation System Based On Buyer Characteristics Using User Based Collaborative Filtering Method*

### *Abstract*

*The product recommendation system is a system that can provide product predictions that are relevant to the behavior or characteristics of the user, so that it can influence the user in making decisions to buy a product. This research was conducted to be able to provide recommendations to buyers on the Sindomall marketplace application using the User Based Collaborative Filtering method in collaboration with the Improved triangle similarity complemented with user rating preferences (ITR) algorithm to calculate the similarity value between buyers and the Weighted Sum algorithm to calculate the predictive value of the product. Buyer characteristics are taken from buyer behavior data in rating the product. In model testing conducted using predictive value data on 20 random users on the Sindomall application database in December 2021, the optimal model was obtained with a user percentage parameter value of 100%. The results of the system error testing using the selected model get the MAE and RMSE values of 0.006 and 0.013 respectively, while at the system accuracy testing stage the accuracy value is 0.849, the precision value is 0.923, the recall value is 0.869, and the F-score (F1) value is 0.895.*

**Keywords:** *User Based Collaborative Filtering*, *Product Recommendation System*, *Marketplace*, *Buyer Characteristics*

## I. PENDAHULUAN

Dikarenakannya perkembangan era digital yang semakin *modern* telah mengakibatkan perubahan perilaku konsumen yang sebelumnya melakukan transaksi melalui toko secara langsung beralih melalui *online shop* atau *e-commerce*. Tidak hanya berupa *traditional e-commerce* melainkan juga mulai bermunculan banyak konsep *e-commerce* yang mulai bermunculan salah satu contohnya adalah *online marketplace*. Perubahan tersebut mampu merubah pemikiran konsumen dalam menentukan keputusan untuk membeli sebuah produk. Keputusan pembeli dalam membeli sebuah produk diakibatkan oleh banyak faktor seperti harga, kualitas produk, promosi bahkan perusahaan tempat mereka membeli sebuah produk [1].

Mudahnya berjualan di *online shop* dan antusias pelanggan yang tinggi menjadi salah satu alasan para pedagang beralih ke penjualan secara *online*. Hal itulah yang membuat pertumbuhan *marketplace* di Indonesia adalah yang tertinggi jika disandingkan dengan bangsa lain. Pernyataan tersebut dipengaruhi oleh peran pemerintah yang senantiasa mendukung perkembangan dan pembangunan *marketplace* secara terus-menerus.

Salah satu hal yang mendukung *marketplace* mengalami pertumbuhan yang pesat adalah fitur-fitur yang disediakan oleh perusahaan sangat inovatif dan dapat menarik minat pelanggan untuk terus membeli produk pada *marketplace* tersebut, baik fitur diskon, *voucher*, *games*, maupun fitur rekomendasi tampilan produk yang disediakan. Diantara banyak fitur yang disediakan *marketplace*, fitur sistem rekomendasi menjadi salah satu fitur yang sangat berpengaruh terhadap pembeli dalam mengambil keputusan untuk membeli suatu produk [2]. Selain itu fitur rekomendasi merupakan salah satu bentuk promosi produk yang dimiliki sehingga dapat meningkatkan penjualan.

Konsep sistem rekomendasi sebenarnya telah diterapkan pada beberapa aplikasi *marketplace* yang sudah ada seperti Amazon, eBay, Tokopedia, Bukalapak, dan lainnya sebagai alat bisnis. Hasilnya beberapa *marketplace* tersebut mengalami peningkatan dalam hal penjualan produk dan membangun loyalitas pembeli [3]. Sistem rekomendasi diperlukan untuk mempengaruhi *user* dalam membuat keputusan hingga keinginan dalam membeli sebuah produk.

Terdapat berbagai macam metode dalam sistem rekomendasi produk, salah satunya *Collaborative Filtering*. Terdapat dua metode pada *Collaborative Filtering* yakni *Item Based* dan *User Based*.

*User Based Collaborative Filtering* sendiri merupakan metode yang dihasilkan dari kesamaan karakteristik dan minat antar *user*. Dimana tiap-tiap *user*-nya merupakan bagian dari sekelompok *user* yang mempunyai kecocokan dengan *user* yang lain. Salah satu hal yang terpenting dari metode ini adalah rekomendasi produk yang dihasilkan akan diperoleh berdasarkan produk yang dinilai oleh *user*. Produk yang direkomendasikan adalah hasil dari produk yang dinilai oleh *user* lainnya ataupun *user* yang memiliki kesamaan produk dinilai.

Penerapan metode *User Based Collaborative Filtering* akan diimplementasikan untuk penentuan rekomendasi produk dalam *marketplace* yang sesuai dengan perilaku pelanggan dalam memberikan penilaian terhadap sebuah produk. Terdapat masalah ketika sistem rekomendasi produk tanpa menerapkan metode yang dapat menyesuaikan dengan karakteristik pembeli, karena hal tersebut tidak akan mempengaruhi pembeli dalam membuat keputusan hingga keinginan dalam membeli sebuah produk.

Metode *User Based Collaborative Filtering* pernah dilakukan oleh Fakhri et al., pada tahun 2019 menggunakan 2 aspek penilaian berupa *rating* dan atribut *user* untuk menghitung nilai *similarity* antar *user* untuk sistem rekomendasi restoran di Bandung Raya. *Weighted Coefficient* merupakan algoritma yang digunakan untuk melakukan perhitungan nilai *similarity* berdasarkan 2 aspek penilaian. Metode *K-Nearest Neighbor* digunakan untuk perhitungan nilai prediksi dengan mengambil sejumlah pengguna dengan kemiripan tertinggi sampai yang terendah. Pengujian performa sistem dilakukan dengan menghitung nilai MAE dengan hasil yang didapat sebesar 2,5 [4].

Berat Ujkani et al. melakukan penelitian dengan menerapkan sistem rekomendasi dengan mengimplementasikan algoritma *Item Based Collaborative Filtering* untuk pemilihan tema didalam fitur *WordPress*. Metode tersebut digunakan untuk membuat sebuah sistem rekomendasi yaitu dengan mengimplementasikan algoritma *Similarity Weight* guna mendapatkan nilai *similarity* antar tema serta menggunakan metode *Weighted Sum* untuk mendapatkan nilai bobot prediksi. Hasil dari beberapa uji coba dari penelitian tersebut mendapatkan nilai MAE sebesar 0,945-0,975 dengan jumlah 10-50 *item* [5].

Berbeda dengan beberapa penelitian sebelumnya, pada penelitian ini menggunakan metode *User Based Collaborative Filtering* untuk menyelesaikan masalah penentuan sistem rekomendasi produk berdasarkan keterkaitan karakteristik antar pembeli dalam hal pemberian *rating* pada suatu produk tertentu. Rekomendasi produk menggunakan metode *User Based Collaborative Filtering* pada aplikasi *marketplace* berdasarkan karakteristik pembeli perlu dinilai akurasi dan *error* yang terukur sehingga dapat diketahui apakah rekomendasi yang dihasilkan baik atau tidak. Pengukuran nilai *similarity* antar *user* akan dilakukan dengan menggunakan persamaan *Improved triangle similarity complemented with user rating preferences* (ITR) sedangkan untuk perhitungan nilai prediksi untuk setiap produk akan dilakukan dengan menggunakan persamaan *Weighted Sum*, sedangkan skenario pengujian dilakukan dengan menghitung nilai *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), Akurasi, Presisi, *Recall* dan *F-score* (F1). Semakin kecil nilai MAE dan RMSE maka rekomendasi yang dihasilkan lebih baik.

## II. LANDASAN TEORI

### A. Sistem Rekomendasi

Sistem rekomendasi adalah galat sistem pengambilan informasi (*Information Retrieval*) dan cabang dari bidang

kecerdasan buatan (*Artificial Intelligence*). Sistem rekomendasi dapat dikatakan sebagai sebuah perangkat lunak yang memberikan prediksi berupa pilihan *item* kepada pengguna sesuai dengan karakteristik dan preferensi pengguna dengan tujuan sebagai sistem pendukung keputusan untuk pengguna dapat menyukai suatu *item*. Kriteria yang digunakan untuk penentuan sistem rekomendasi produk tidak hanya terbatas pada *single* kriteria seperti *rating* yang diberikan oleh pengguna, tetapi dapat menggunakan multi kriteria. Multi kriteria dalam rekomendasi sistem mengambil beberapa karakteristik penting dalam *item*, misal dalam sistem rekomendasi musik dapat mengambil beberapa kriteria seperti lirik, visual, audio, suara, ketukan, genre, dll. Dengan multi kriteria dalam rekomendasi sistem dapat menentukan preferensi subjektif dari pengguna pada beberapa atribut dalam *item* sehingga untuk meningkatkan akurasi prediksi sistem [6].

### B. User Based Collaborative Filtering

*Collaborative Filtering* atau biasa disingkat CF merupakan metode yang menghasilkan rekomendasi khusus kepada pengguna tentang *item* berdasarkan pola penilaian atau penggunaan (misalnya, pembelian) tanpa memerlukan informasi eksogen tentang *item* atau pengguna tersebut. Metode ini menghasilkan prediksi atau rekomendasi untuk pengguna tertentu untuk satu atau banyak *item*. *Item* yang dikatakan mungkin terdiri dari *item-item* yang dapat diberikan *rating* oleh pengguna, nilai *rating* dalam *Collaborative Filtering* dapat mengambil dari berbagai bentuk bisa berupa *voting* ataupun *binary*.

*User Based Collaborative Filtering* sendiri merupakan penyaringan yang dihasilkan dari ekuualitas karakteristik dan minat antar *user*. Dimana setiap entitas merupakan elemen dari sekelompok *user* yang memiliki kesamaan dengan *user* lainnya. Yang harus di garis bawahi dari metode ini yaitu rekomendasi produk yang dihasilkan akan diperoleh berdasarkan *item* yang dinilai *rating* oleh *user*. *Item* yang dijadikan sebagai rekomendasi merupakan hasil dari *item* yang diberi *rating* oleh *user* lainnya ataupun *user* yang memiliki kesamaan produk yang diberi *rating*.

### C. Improved Triangle Similarity Complemented With User Rating Preferences (ITR)

*Improved triangle similarity complemented with user rating preferences* (ITR) merupakan algoritma yang dapat diimplementasikan untuk mendapatkan nilai *similarity* terhadap dua buah titik, didalam algoritma ITR terdapat dua istilah *improved triangle similarity* ( $sim^{TRIANGLE}$ ) dan *user rating preferences* (URP) (Iftikhar et al., 2020). Dalam praktiknya,  $sim^{TRIANGLE}$  dianggap sebagai peningkatan dari kesamaan *Triangle Multiplying Jaccard* (TMJ), sehingga *Improved Triangle Similarity* (ITR) tidak hanya berfokus pada peringkat umum, seperti ukuran TMJ, tetapi juga memperhitungkan peringkat pengguna yang tidak umum [7].

### D. Weighted Sum

Untuk mendapatkan perhitungan (bobot) prediksi untuk setiap *item* yang direkomendasikan kepada *user*, dapat diuji dengan menggunakan persamaan *Weighted Sum*. *Weighted*

*Sum* sendiri merupakan metode untuk pengambilan keputusan dengan multi-kriteria sederhana guna mengevaluasi sejumlah alternatif yang kaitannya dengan sebanyak jumlah kriteria keputusan [8].

### E. Evaluasi Sistem Rekomendasi

Evaluasi bertujuan untuk mengetahui seberapa tinggi tingkat akurasi ataupun tingkat *error* dari hasil sistem rekomendasi pada setiap *user*. Terdapat berbagai persamaan dan algoritma yang dapat digunakan untuk mendapatkan tingkat akurasi dan tingkat *error* pada hasil sistem rekomendasi dapat menggunakan persamaan *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), Akurasi, Presisi, *Recall*, serta *F-score* ( $F_1$ ).

## III. METODOLOGI PENELITIAN

### A. Pengumpulan Data

Data yang digunakan peneliti sebagai objek dalam melakukan penelitian merupakan data sekunder yang diperoleh dari data *rating* pembeli dan data produk yang belum laku terjual pada aplikasi Sindomall di perusahaan Sindo Lah Pte Ltd pada bulan Desember 2021.

Data *rating* pembeli yang didapat memiliki sebanyak 264 baris terhadap 420 produk (terdapat nilai atribut yang kosong untuk setiap pembeli) dengan kisaran nilai *rating* 1 sampai 5 (untuk pembeli yang sudah melakukan pembelian pada suatu produk) yang disusun secara berurutan berdasarkan nama pembeli dan nama produk.

### B. User Based Collaborative Filtering

Metode *User Based Collaborative Filtering* digunakan karena sesuai dengan visi yang digunakan pada aplikasi *marketplace* Sindomall, dimana lebih mengedepankan terbentuknya suatu kelompok komunitas pada *user-nya* terhadap beberapa produk tertentu. Dengan menggunakan metode *User Based Collaborative Filtering* yang mendefinisikan setiap *user-nya* merupakan bagian dari sekelompok *user* yang memiliki kecocokan dengan *user* lainnya, akan dapat membentuk suatu kelompok *user* dengan minat yang sama pada beberapa produk. Berikut ini akan dijelaskan bagaimana proses untuk mendapatkan sebuah rekomendasi produk dengan mengimplementasikan metode *User Based Collaborative Filtering*.

#### 1. Data Rating

Tahapan pertama dalam pengimplementasian *User Based Collaborative Filtering* yaitu dengan pemberian *rating* yang dilakukan oleh pembeli terhadap sebuah produk pada aplikasi Sindomall, adapun format data yang akan disajikan adalah seperti pada Tabel 1.

Tabel 1. Contoh Tabel Skenario *Rating*

Produk User	(a) <i>Philodendron Gloriosum</i>	(b) <i>Alocasia Macrorrhiza</i>	(c) <i>Labisia sp.</i>	(d) <i>Aglao nema Pictum</i>	(e) <i>Bego nia sp.</i>
(1) Cedric	5	4	3	3	4
(2) Daniel	4	3	5	5	4
(3) Yuni	2	3	3	5	5
(4) Hebe	5	5	5	4	4
(5) Giva	-	-	4	-	-

2. Perhitungan *Similarity*

Persamaan yang digunakan untuk mengukur nilai kemiripan antar dua *user* yaitu dengan menggunakan persamaan *Improved triangle similarity complemented with user rating preferences* (ITR). Dimana menurut penelitian yang dilakukan Fkih pada tahun 2021, peneliti *me-review* dan membandingkan 13 algoritma perhitungan *similarity measure* pada skenario uji coba menggunakan tiga *dataset* standar (*MovieLens100k*, *MovieLens1M* dan *Jester*), dimana ITR merupakan algoritma yang paling cocok dan efisien untuk metode *User Based Collaborative Filtering* [9]. Adapun rumus algoritma ITR adalah pada Persamaan 1.

$$sim^{ITR}(u, v) = sim^{TRIANGLE}(u, v) * sim^{URP}(u, v) \tag{1}$$

Keterangan:

- $sim^{ITR}(u, v)$  = Nilai *similarity* ITR antar *user* u dengan *user* v
- $sim^{TRIANGLE}(u, v)$  = Nilai *similarity triangle* antar *user* u dengan *user* v
- $sim^{URP}(u, v)$  = Nilai *similarity* URP antar *user* u dengan *user* v

Adapun untuk rumus persamaan  $sim^{TRIANGLE}(u, v)$  dan  $sim^{URP}(u, v)$  adalah pada Persamaan 2 dan 3.

$$sim^{TRIANGLE}(u, v) = 1 - \frac{\sqrt{\sum_{i \in I_{uv}} (r_{ui} - r_{vi})^2}}{\sqrt{\sum_{i \in I_{uv}} r_{ui}^2} + \sqrt{\sum_{i \in I_{uv}} r_{vi}^2}} \tag{2}$$

Keterangan:

- $I_{uv}$  = Himpunan nilai *rating* oleh *user* u atau *user* v
- $r_{ui}$  = Nilai *user* u pada *item* ke-i
- $r_{vi}$  = Nilai *user* v pada *item* ke-i

$$sim^{URP}(u, v) = 1 - \frac{1}{1 + \exp(-|\bar{r}_u - \bar{r}_v| * |\sigma_u - \sigma_v|)} \tag{3}$$

Keterangan:

- $\bar{r}_u$  = Rata-rata nilai *rating* *user* u

- $\bar{r}_v$  = Rata-rata nilai *rating* *user* v
- $\sigma_u$  = Simpangan baku nilai *rating* *user* u
- $\sigma_v$  = Simpangan baku nilai *rating* *user* v

Dengan rumus simpangan baku pada Persamaan 4

$$\sigma_u = \sqrt{\frac{\sum_{i \in I_u} (r_{ui} - \bar{r}_u)^2}{|I_u|}} \tag{4}$$

Keterangan:

- $r_{ui}$  = Nilai *user* u pada *item* ke-i
- $\bar{r}_u$  = Rata-rata nilai *rating* *user* u
- $I_u$  = Panjang himpunan nilai *rating* *user* u

Nilai *similarity* memiliki nilai antara 0 dan 1. Jika bernilai 0 maka antar dua *user* tersebut tidak memiliki kemiripan, sedangkan bernilai 1 maka antar dua *user* tersebut dapat dikatakan identik [10].

3. Perhitungan Prediksi

Perhitungan nilai prediksi produk dilakukan setelah berhasil mendapatkan nilai *similarity* antar *user*. Persamaan yang digunakan untuk menghitung nilai prediksi produk adalah dengan menggunakan *Weighted Sum*. Dimana menurut penelitian yang dilakukan oleh Fkih pada tahun 2021, peneliti membandingkan 3 metode perhitungan nilai prediksi *rating* pada 13 algoritma perhitungan *similarity measure*, metode *Weighted Sum* menghasilkan nilai akurasi terbesar dibandingkan 2 metode lainnya yaitu *Z-Score* dan *Mean-Centered Aggregation* [9]. Adapun rumus persamaan *Weighted Sum* adalah pada Persamaan 5.

$$P(a, i) = \frac{\sum_{i \in I_1} (R_{a,i} * S_{i,j})}{\sum_{i \in I_1} |S_{i,j}|} \tag{5}$$

Keterangan:

- $P(a, i)$  = Nilai prediksi antar *item* a terhadap objek i
- $R(a, i)$  = Penilaian untuk *item* a oleh objek i
- $S(i, j)$  = Nilai *similarity* antar objek i dengan objek j

4. Hasil Rekomendasi

Dari hasil perhitungan nilai prediksi *user* terhadap suatu produk, semua data prediksi tersebut khususnya pada suatu *user* akan diurutkan dari nilai prediksi tertinggi hingga ke yang terendah. Ternyata nilai prediksi *rating* tertinggi akan berada pada peringkat pertama. Penampilan hasil rekomendasi *Top-N* produk pada sistem akan diambil dari urutan *top ranking* 10 dari total keseluruhan produk. Hal tersebut didasarkan pada hasil penelitian yang dilakukan oleh Faizn dan Surjandari pada tahun 2020, menyimpulkan bahwa nilai *Top-N* rekomendasi produk dengan faktor *learning rate*, *batch size*, dan jumlah faktor prediktif pada *marketplace* di Indonesia sebesar 10 produk [11].

Pada Tabel 5 merupakan hasil dari skenario kasus perhitungan nilai prediksi dimana *user* Giva memiliki nilai prediksi pada setiap produk yang belum pernah diberikan *rating*. Data tersebut masih belum diurutkan dari nilai yang

tertinggi ke terendah, adapun jika diurutkan maka hasil rekomendasi produk pada *user* Giva adalah dengan urutan sebagai berikut:

1. *Aglaonema Pictum* (Produk d)
2. *Begonia sp.* (Produk e)
3. *Alocasia Macrorrhiza* (Produk b)
4. *Philodendron Gloriosum* (Produk a)

**C. Pengujian**

Data yang akan diuji pada tahap pengujian merupakan data *real* hasil dari perhitungan manual nilai prediksi dengan menggunakan MATLAB dibandingkan dengan data *real* hasil perhitungan sistem dengan menggunakan *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE) untuk menghitung perbedaan besaran *error* dalam perhitungan nilai prediksi. Terdapat 2 matrik evaluasi yang terkenal untuk melakukan penilaian pada kinerja sistem rekomendasi yaitu MAE dan RMSE [12]. Adapun nilai MAE dan RMSE akan dihitung dengan Persamaan 6 dan Persamaan 7.

$$MAE = \frac{\sum_{i=1}^N |p_{i,j} - q_{i,j}|}{N} \tag{6}$$

Keterangan:

- MAE = Nilai rata-rata *error*
- N = Total *user*
- $p_{i,j}$  = Nilai prediksi *rating user* i terhadap produk j
- $q_{i,j}$  = Nilai real *rating user* i terhadap produk j

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\tilde{r}_i - r_i)^2}{n}} \tag{7}$$

Keterangan:

- n = Jumlah *user*
- $\tilde{r}_i$  = Nilai prediksi *rating user* i terhadap produk j
- $r_i$  = Nilai real *rating user* i terhadap produk j

Karena sistem rekomendasi dianggap sebagai jenis sistem pengambilan informasi (*information retrieval*) tertentu, presisi dan *recall* matrik telah digunakan oleh banyak peneliti untuk evaluasi sistem rekomendasi, presisi diartikan sebagai rasio rekomendasi yang relevan dengan jumlah total *item* yang direkomendasikan. *Recall* diartikan sebagai rasio rekomendasi yang relevan dengan jumlah produk yang relevan [13].

$$Presisi = \frac{\text{Produk yang direkomendasikan dengan benar}}{\text{Jumlah produk yang direkomendasikan}} \tag{8}$$

$$Recall = \frac{\text{Produk yang relevan}}{\text{Jumlah produk yang direkomendasikan}} \tag{9}$$

Hasil nilai presisi dan *recall* dapat digabungkan menjadi satu metrik, yang disebut *F-score* ( $F_1$ ), seperti yang didefinisikan dalam Persamaan 10.

$$F_1 = \frac{2 * Presisi * Recall}{Presisi + Recall} \tag{10}$$

Untuk perhitungan nilai akurasi akan dihitung menggunakan Persamaan 11.

$$Akurasi = \frac{TP}{(TP + TN + FP + FN)} \tag{11}$$

Dimana,

1. *True Positive* (TP) adalah produk yang direkomendasikan benar dan produk tersebut relevan.
2. *True Negative* (TN) adalah produk yang direkomendasikan benar tetapi produk tersebut tidak relevan.
3. *False Positive* (FP) adalah produk yang direkomendasikan salah tetapi produk tersebut relevan.
4. *False Negative* (FN) adalah produk yang direkomendasikan salah dan produk tersebut tidak relevan.

Untuk setiap skenario pengujian, perhitungan nilai prediksi akan dimulai dengan nilai persentase *user* sebesar 10% dengan nilai kemiripan terdekat. Kemudian, akan diperluas setiap kali pengujian sebesar 10% hingga 100% (keseluruhan *user*), dari total *user* dengan nilai kemiripan tersebut akan dicari pada nilai persentase ke berapa sistem memiliki nilai akurasi yang paling tinggi.

**IV. HASIL DAN PEMBAHASAN**

Tahapan pertama dalam melakukan pengujian adalah memperoleh data *groundtruth* dengan tujuan sebagai pembanding dari hasil perhitungan nilai prediksi yang diambil dalam sistem yang sudah diimplementasikan. Adapun data *groundtruth* pada riset ini adalah berupa nilai prediksi pada 20 *user* secara acak dari total 264 *user* yang telah memberikan *rating* pada produk, data *rating user* tersebut diambil dari *database* aplikasi Sindomall pada bulan Desember 2021, dari data *rating* yang didapat akan di-generate dengan aplikasi Matlab untuk mendapatkan *groundtruth* nilai prediksi produk. *Groundtruth* yang diperoleh berupa tabel dengan data atribut “*UserLoggedInID*” dengan tipe data *integer*, atribut “*ProductID*” dengan tipe berupa *integer* serta *index* “*WeightedSum*” yang merupakan nilai prediksi dengan tipe data *float*. Tabel 2 merupakan salah satu sampel dari *user* yang diimplementasikan untuk data *groundtruth*.

Tabel 2. Data Sampel *Groundtruth*

<i>UserLoggedInID</i>	<i>ProductID</i>	Nilai Prediksi
6624	2815	0,0845069849469635
6624	387	0,107536762168715
6624	1844	0,165158582296050
6624	2157	0,282859712238740
6624	872	0,0422399885213950

...	...	...
6624	836	0,0351995093802755
6624	231	0,354197832863011
6624	507	0,0360600121009402
6624	679	0,0634751132445122
6624	1499	0,156593303058968

Tahapan setelah memperoleh data *groundtruth* adalah melakukan tahap pelatihan untuk membangun model *User Based Collaborative Filtering* dan tahap pengujian dengan menggunakan data *groundtruth* yang selanjutnya dilakukan perbandingan pada hasil perhitungan nilai prediksi yang diperoleh dalam sistem pada nilai atribut “*UserLoggedInID*” dan nilai atribut “*ProductID*” yang sama.

Pada tahap pelatihan ini, yang dilakukan pertama kali adalah pengambilan data *rating user*, semua data *rating user* berdasarkan “*UserID*” yang diinputkan pada keseluruhan produk diambil, jika suatu *user* merupakan *user* baru yang artinya belum memberikan data *rating* pada produk apapun maka nilai *rating*-nya akan diinisialisasikan menjadi 0,

Data *rating user* yang sebelumnya telah didapatkan akan dihitung kemiripannya dengan algoritma ITR terhadap data *rating user* yang lainnya yang telah memberikan *rating* pada produk. Tabel 3 contoh sampel dari perhitungan nilai *similarity* pada salah satu *user*.

Tabel 3. Nilai *Similarity* Pada Salah Satu *User*.

<i>UserLoggedIn ID</i>	<i>UserReviewed ID</i>	<i>Similarity</i>
6624	6037	0,5278646458073842
6624	6028	0,20362516398012426
6624	4544	0,023570980236055958
6624	6088	0,39716693627526856
6624	5679	0,07273191452649491
...	...	...
6624	6075	0,07273191452649491
6624	4972	0,3088548412757774
6624	4925	0,2587619719450473
6624	566	0,14219589186489656
6624	5776	0,05710594822922677

Setelah itu, diadakan pengukuran dengan merubah nilai parameter persentase *user* dengan nilai *similarity* antar *user* yang paling besar, dimana tingkat *error* dari setiap parameter akan dilakukan pemantauan pada perubahan nilai *error*-nya. Hal tersebut dilakukan untuk dapat mengetahui berapa jumlah persentase *user* yang memiliki nilai kemiripan terbesar dalam menghitung nilai prediksi yang memiliki nilai *error* paling minimal. Skenario perubahan parameter persentase *user* pada proses pelatihan dimulai pada persentase 10% dari total *user*, dan meningkat sebesar 10% dalam setiap pengujian sampai mencapai nilai 100% atau keseluruhan *user*. Berdasarkan skenario tersebut, maka model pelatihan yang diperoleh terdapat pada Tabel 4.

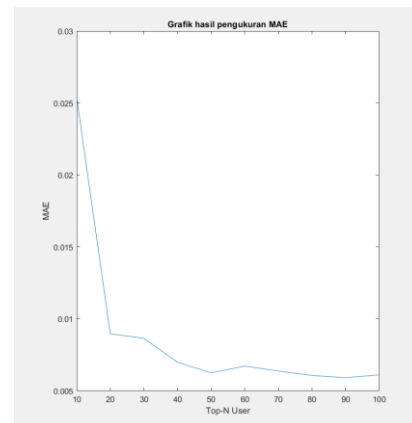
Tabel 4. Skenario Perubahan Parameter Persentase *User*

Model	Persentase <i>User</i> (%)
1	10
2	20
3	30
4	40
5	50
6	60
7	70
8	80
9	90
10	100

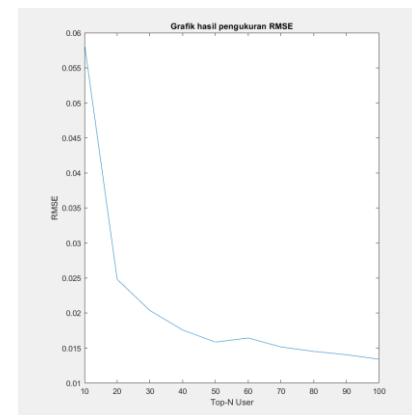
Setelah model skenario pengujian telah terbentuk, tahapan berikutnya adalah melakukan pengujian untuk dapat memilih model mana yang memiliki nilai *error* paling minimal dengan membandingkan hasil nilai prediksi dari sistem dengan data uji.

Setelah melakukan pengujian dengan skenario yang telah ditentukan, maka akan diperoleh tingkat *error* dari metode *User Based Collaborative Filtering* berdasarkan pengukuran *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE).

Adapun hasil pengukuran MAE dan RMSE pada masing-masing model adalah pada Gambar 1 dan Gambar 2.

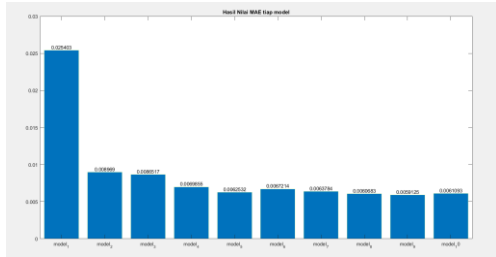


Gambar 1. Grafik Hasil Pengukuran MAE Pada Tiap Model

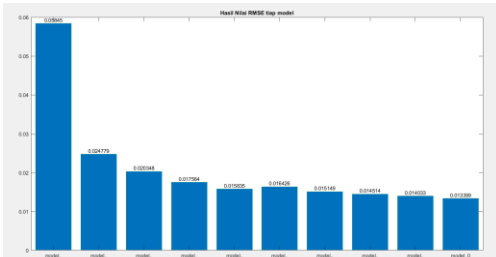


Gambar 2. Grafik Hasil Pengukuran RMSE Pada Tiap Model

Berdasarkan uji coba dalam memilih model dengan nilai *error* yang paling minimal pada Tabel 6, maka dapat ditarik kesimpulan pada Gambar 3 dan Gambar 4.



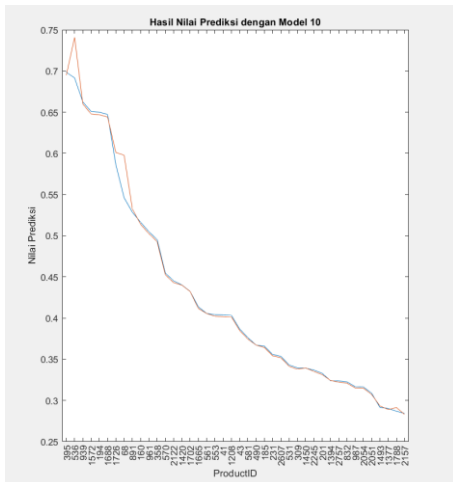
Gambar 3. Pengukuran MAE Pada Tiap Model Pelatihan



Gambar 4. Pengukuran RMSE Pada Tiap Model Pelatihan

Berdasarkan grafik pada Gambar 4 dan Gambar 5 perbandingan hasil nilai *error* pada setiap model pelatihan dengan variasi perubahan nilai parameter persentase *user*, model 10 memiliki nilai RMSE yang paling minimal dengan nilai 0,013399 dibandingkan dengan nilai model lainnya. Oleh karena itu, parameter yang akan digunakan pada tahap uji coba adalah sesuai dengan model 10 yaitu keseluruhan *user* atau 100% *user*.

Data yang akan digunakan pada tahap uji coba merupakan sama dengan data yang digunakan pada tahap pelatihan. Adapun hasil dari uji coba pengukuran nilai prediksi menggunakan model 10 terhadap data uji salah satu *user* dengan ID 6624 ada pada Gambar 5.



Gambar 5. Sampel Hasil Perbandingan Hasil Prediksi Produk Dengan Data *Groundtruth* Pada *User* Dengan ID 6624

Hasil pengukuran nilai prediksi untuk sample *user* ID 6624 kemudian dibandingkan dengan nilai aktual seperti yang ditunjukkan pada Tabel 5.

Tabel 5. Sampel Hasil Perhitungan Nilai Prediksi Pada *User* Dengan ID 6624

User ID	Product ID	Nilai Prediksi	Nilai Aktual
6624	395	0,698378618061853	0,695020703282443
6624	536	0,691609239954889	0,740901720318779
6624	939	0,662347837842095	0,659941575988402
6624	1572	0,650917320832368	0,647787607471637
6624	194	0,649987387519991	0,646862145425609
6624	1688	0,647253344019953	0,644141247638854
6624	1726	0,586157133227591	0,600997000908553
6624	68	0,546105868753410	0,597808602271507
6624	891	0,528583617369766	0,532655616707886
6624	160	0,516288957802623	0,513806558890321
6624	961	0,504922873013092	0,502495124032950
6624	358	0,495273396601240	0,492892043828855
6624	570	0,454633900779681	0,452447949126571
6624	2122	0,445166463615163	0,443026032896337
6624	1420	0,440215404392039	0,439990257781758
6624	1702	0,432388758442872	0,432201243583373
6624	1665	0,413596047870877	0,411607412700880
6624	561	0,405618914281552	0,405560113063947
6624	553	0,404059232857387	0,402116452201337
6624	41	0,403768673720325	0,401827290118532
6624	1208	0,403394133172160	0,401454550420521
6624	43	0,386782707656397	0,384922995264193
6624	581	0,376212453964920	0,374403565023170
6624	490	0,367082811637095	0,367209298091913
6624	185	0,365902158385215	0,364142842974228
6624	231	0,355909100070144	0,354197832863011
6624	2607	0,353478404923001	0,351778824882324
6624	531	0,343051052825519	0,341401609141845
6624	309	0,339473842601996	0,337841598710481
6624	1450	0,339178209668113	0,339438865859722
6624	2245	0,336808447481514	0,335189019231022
6624	201	0,332903534704676	0,331302881892015
6624	1394	0,324033440210626	0,324366914847526
6624	2757	0,323674210936016	0,322117934170796
6624	832	0,322579733516727	0,321028719171930
6624	987	0,316753342725811	0,315230342588839
6624	2054	0,316342434813483	0,314821410386700
6624	2051	0,308685543186407	0,307201334304794
6624	1493	0,291560100900166	0,293335573763094
6624	1377	0,290280098230495	0,288884385637343
6624	1788	0,286871889316137	0,291369091989498
6624	2157	0,284226317296316	0,282859712238740

Dari hasil tabel pengukuran nilai prediksi pada Tabel 5, selanjutnya data-data tersebut dihitung nilai MAE dan RMSE pada seluruh data uji yang ditampilkan pada Tabel 6.

Tabel 6. Hasil Perhitungan Nilai MAE Dan RMSE Pada Data Uji

User ID	MAE	RMSE
481	0,005491635088760	0,012411641149291
644	0,006237039453667	0,014122151817566

703	0,015334498182913	0,023582001972058
1001	0,005871580756359	0,013233925891327
1011	0,005471407229314	0,012371107929966
1224	0,006048661127732	0,012696901956723
2669	0,005879764269807	0,013248615900930
3721	0,006018842716870	0,013438112554318
3876	0,005245937176106	0,011846455485858
3973	0,003618292769555	0,009913514777409
4517	0,005978100232473	0,013502852095531
4811	0,006571222684326	0,014789041877031
5035	0,005681433461807	0,014828236666735
5127	0,005262750417987	0,011880887938167
5344	0,005861303726034	0,013217358258767
5451	0,006194328025977	0,014026922473032
6037	0,007427294663996	0,015358599741547
6561	0,004175673769774	0,010530450148534
6624	0,004468373926925	0,011436925646037
6905	0,005348047212621	0,011550455281337

6624	1665	-	-
6624	561	-	-
6624	553	v	v
6624	41	v	v
6624	1208	v	v
6624	43	v	v
6624	581	v	v
6624	490	v	v
6624	185	v	v
6624	231	v	v
6624	2607	v	v
6624	531	v	v
6624	309	-	-
6624	1450	v	v
6624	2245	v	v
6624	201	v	v
6624	1394	v	v
6624	2757	v	v
6624	832	v	v
6624	987	v	v
6624	2054	v	v
6624	2051	v	v
6624	1493	v	v
6624	1377	v	v
6624	1788	v	v
6624	2157	v	v

Dari perhitungan nilai MAE dan RMSE diatas selanjutnya akan dihitung nilai rata-ratanya. Adapun perhitungan rata-rata nilai MAE dan RMSE adalah sebagai berikut.

$$MAE = \frac{0,122186186893004}{20} = 0,006109309344650$$

$$RMSE = \frac{0,267986159562163}{20} = 0,013399307978108$$

Setelah nilai prediksi dari produk yang direkomendasikan sesuai dengan "ProductID" yang telah didapatkan, selanjutnya akan diimplementasikan tahapan selanjutnya dengan menghitung nilai akurasi, presisi, recall, dan F-score (F1) berdasarkan output produk yang didapatkan. Output produk dikatakan benar apabila memiliki "ProductID" yang sama dengan hasil perhitungan nilai prediksi, dan output produk dikatakan relevan apabila memiliki "ProductID" yang sama dengan hasil perhitungan nilai prediksi dan berstatus belum terjual. Adapun contoh dari output produk yang direkomendasikan pada salah satu user ada pada Tabel 7.

Tabel 71. Sampel Hasil Analisis Output Rekomendasi Produk Pada User Dengan ID 6624

User ID	Product ID	Produk yang benar	Produk yang relevan
6624	395	v	v
6624	536	v	v
6624	939	v	v
6624	1572	v	v
6624	194	v	v
6624	1688	v	v
6624	1726	v	v
6624	68	v	v
6624	891	v	v
6624	160	v	v
6624	961	v	v
6624	358	v	v
6624	570	v	-
6624	2122	v	-
6624	1420	v	v
6624	1702	v	-

Berdasarkan Tabel 7 di atas, maka selanjutnya akan dihitung nilai presisi dan recall-nya. Dari nilai Presisi dan Recall yang telah didapatkan, maka selanjutnya akan dihitung nilai F-score (F1). Hasil perhitungan nilai akurasi, presisi, recall, dan F-score (F1) pada data uji yang ditunjukkan pada Tabel 8.

Tabel 8. Hasil Perhitungan Nilai Akurasi, Presisi, Recall, Dan F-Score (F1) Pada Data Uji

User ID	Akurasi	Presisi	Recall	F-score (F1)
481	0,857	0,928	0,857	0,891
644	0,857	0,928	0,857	0,891
703	0,833	0,904	0,857	0,879
1001	0,857	0,928	0,857	0,891
1011	0,880	0,928	0,904	0,915
1224	0,833	0,857	0,857	0,857
2669	0,928	0,952	0,952	0,952
3721	0,857	0,928	0,857	0,891
3876	0,809	0,904	0,904	0,904
3973	0,857	0,928	0,857	0,891
4517	0,857	0,928	0,857	0,891
4811	0,809	0,952	0,880	0,914
5035	0,857	0,928	0,857	0,891
5127	0,857	0,928	0,857	0,891
5344	0,809	0,952	0,880	0,914
5451	0,857	0,928	0,857	0,891
6037	0,857	0,928	0,857	0,891
6561	0,809	0,880	0,880	0,880
6624	0,857	0,928	0,857	0,891
6905	0,857	0,928	0,857	0,891

Berdasarkan output hasil nilai akurasi, presisi, recall, dan F-score (F1) di atas selanjutnya akan dihitung nilai rata-



ratanya. Adapun perhitungan nilai rata-rata akurasi, presisi, *recall*, dan *F-score* ( $F_1$ ) adalah sebagai berikut.

$$\text{Akurasi} = \frac{16,994}{20} = 0,849$$

$$\text{Presisi} = \frac{18,465}{20} = 0,923$$

$$\text{Recall} = \frac{17,398}{20} = 0,869$$

$$F_1 = \frac{17,907}{20} = 0,895$$

Mengacu pada proses uji coba yang selesai dilakukan, maka diperoleh nilai akurasi dan *error* dari perhitungan nilai prediksi pada model yang memiliki nilai *error* paling minimal pada data *rating* 20 *user* yang dipilih secara acak sebelumnya. Pada proses pelatihan parameter persentase *user* yaitu untuk mendapatkan nilai *similarity* pada *user* yang memiliki nilai terdekat, parameter tersebut memiliki pengaruh yang berbeda pada hasil akhir. Jika nilai persentase terlalu kecil maka nilai prediksi yang didapatkan tidak terlalu optimal sehingga mendapatkan nilai *error* yang tinggi. Sedangkan jika nilai persentase tinggi maka data yang digunakan untuk menghitung nilai prediksi akan semakin lengkap sehingga akan mendapatkan nilai *error* yang rendah.

Pada perhitungan nilai *Mean Absolute Error* (MAE) dan *Root Mean Square Error* (RMSE) mendapatkan kesimpulan bahwa metode *User Based Collaborative Filtering* dengan menggunakan algoritma ITR dalam perhitungan nilai *similarity* dan metode *Weighted Sum* dalam perhitungan nilai prediksi dengan menggunakan parameter persentase *user* sebesar 100% pada data uji memperoleh hasil akurasi RMSE sebesar 0,013 dan *error* MAE sebesar 0,006. Pada pengevaluasian hasil rekomendasi produk yang diberikan oleh sistem hasil dari menghitung besaran akurasi, presisi, *recall*, dan *F-score* ( $F_1$ ), *output* produk dikatakan benar apabila memiliki "*ProductID*" yang sama dengan hasil perhitungan *top 10* nilai prediksi produk, dan *output* produk dikatakan relevan apabila memiliki "*ProductID*" yang sama dengan hasil perhitungan *top 10* nilai prediksi dan berstatus belum terjual. Adapun hasil dari perhitungan diperoleh nilai Akurasi sebesar 0,849, nilai Presisi sebesar 0,923, nilai *Recall* sebesar 0,869, dan nilai *F-score* ( $F_1$ ) sebesar 0,895.

Walaupun nilai MAE dan RMSE yang didapatkan kecil sedangkan apabila dilakukan perhitungan berdasarkan persentase absolut *error* rata-ratanya didapatkan hasil yang tidaklah kecil. Hal ini disebabkan karena dalam perhitungan RMSE tidak dapat terlepas dari konsep variansi, sehingga nilainya akan sangat rentan terhadap selisih antar rata-rata data yang digunakan. Oleh karena itu, nilai MAE dan RMSE yang kecil dikarenakan data nilai prediksi yang digunakan merupakan nilai satuan sehingga selisih nilai prediksi dengan data nilai pada data *groundtruth* akan rentan menjadi nilai yang lebih kecil.

Berdasarkan grafik yang ditunjukkan pada Gambar 5, masih dapat memungkinkan untuk menghasilkan rekomendasi

produk pada *user* tertentu dengan melakukan perangkingan dari produk yang memiliki nilai prediksi tertinggi ke terendah. Masih terdapatnya nilai *error* yang dapat ditoleransi dikarenakan masih dominannya produk yang belum diberi oleh para *user*, sehingga perhitungan nilai *similarity* antar *user* tidak dapat dilakukan secara maksimal

## V. KESIMPULAN DAN SARAN

Dari penelitian yang telah selesai dilakukan untuk membangun sistem rekomendasi produk pada aplikasi *marketplace* dengan menggunakan metode *User Based Collaborative Filtering* dikolaborasi dengan algoritma ITR untuk menghitung nilai *similarity* dan algoritma *Weighted Sum* untuk menghitung nilai prediksi serta dilakukan pengujian dengan data uji yang didapatkan dari *database* aplikasi *marketplace* Sindomall pada bulan Desember tahun 2021 yang disajikan dalam bentuk nilai prediksi pada setiap *user* yang telah memberikan *rating* pada suatu produk.

Pembuatan beberapa model yang ditujukan untuk mencari nilai parameter persentase *user* dengan nilai *error* paling minimal, yang selanjutnya diperoleh model terpilih dengan nilai parameter persentase *user* sebesar 100% yang akan digunakan sebagai parameter dalam pengujian untuk mengukur nilai *error* dan akurasi dengan data uji berupa data *rating* 20 *user* yang diambil secara acak. Pada tahap pengujian *error* sistem didapatkan nilai MAE sebesar 0,006 dan nilai RMSE sebesar 0,013, sedangkan pada tahap pengujian akurasi sistem didapatkan nilai akurasi sebesar 0,849, nilai presisi sebesar 0,923, nilai *recall* sebesar 0,869, dan nilai *F-score* ( $F_1$ ) sebesar 0,895.

Dengan nilai *error* dan akurasi sistem yang telah diperoleh dapat disimpulkan bahwa sistem dapat memberikan rekomendasi produk pada *user* dengan baik, hal tersebut dikarenakan nilai *error* yang didapat terbilang rendah karena mendekati nilai 0 dan nilai akurasi sistem yang dapat dikatakan tinggi.

Selain itu, peneliti menyadari bahwa penelitian yang dilakukan masih memiliki banyak sekali kekurangan sehingga peneliti mengharapkan adanya perbaikan untuk dapat meningkatkan performa sistem. Berikut merupakan *point-point* yang perlu dipertimbangkan pada penelitian selanjutnya:

1. Penambahan kriteria lain dalam penambihan data karakteristik pembeli tidak hanya terbatas pada perilaku pembeli dalam memberi *rating* pada produk sehingga akurasi sistem menjadi lebih baik.
2. Menggunakan algoritma perhitungan nilai *similarity* yang dapat mengatasi *sparsity problem*, yaitu permasalahan terkait nilai *rating* produk yang kosong sehingga dapat meningkatkan akurasi sistem.

## REFERENSI

- [1] Kotler and Keller, "Marketing Management," *Soldering & Surface Mount Technology*, vol. 13, no. 3, 2012.
- [2] P. Huang, N. H. Lurie, and S. Mitra, "Searching for Experience on The Web: An Empirical Examination of

- Consumer Behavior for Search and Experience Goods,” *Journal of Marketing*, vol. 73, no. 2, pp. 55–69, 2009.
- [3] C. S. Prasetya, “Sistem Rekomendasi Pada E-commerce Menggunakan K-nearest Neighbor,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 4, no. 3, p. 194, 2017.
- [4] A. A. Fakhri, Z. K. Baizal, and E. B. Setiawan, “Restaurant Recommender System Using User-based Collaborative Filtering Approach: A Case Study at Bandung Raya Region,” *Journal of Physics: Conference Series*, vol. 1192, p. 012023, 2019.
- [5] B. Ujkani, D. Minkovska, and L. Stoyanova, “A Recommender System for WordPress Themes Using Item-based Collaborative Filtering Technique,” *2020 XXIX International Scientific Conference Electronics (ET)*, 2020.
- [6] Y. Arif, H. Nurhayati, F. Kurniawan, S. Nugroho, and M. hariadi, “Blockchain-based Data Sharing for Decentralized Tourism Destinations Recommendation System,” *International Journal of Intelligent Engineering and Systems*, vol. 13, no. 6, pp. 472–486, 2020.
- [7] A. Iftikhar, M. A. Ghazanfar, M. Ayub, Z. Mehmood, and M. Maqsood, “An Improved Product Recommendation Method for Collaborative Filtering,” *IEEE Access*, vol. 8, pp. 123841–123857, 2020.
- [8] D. Handoko, Mesran, S. D. Nasution, Yuhandri, and H. Nurdiyanto, “Application of Weight Sum Model (WSM) in Determining Special Allocation Funds Recipients,” *The IJICS (International Journal of Informatics and Computer Science)*, vol. 1, no. 2, pp. 31–35, 2017.
- [9] F. Fkih, “Similarity Measures for Collaborative Filtering-Based Recommender Systems: Review and Experimental Comparison,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 9, pp. 7645–7669, 2022.
- [10] A. A. Firmansyah, “Pengembangan Pencarian Produk Terkait Menggunakan Euclidean Distance dan Cosine Similarity Pada Aplikasi Halal Nutrition Food,” Thesis, Institut Teknologi Sepuluh Nopember, Surabaya, 2018.
- [11] A. Faizin and I. Surjandari, “Product Recommender System Using Neural Collaborative Filtering for Marketplace in Indonesia,” *IOP Conference Series: Materials Science and Engineering*, vol. 909, no. 1, p. 012072, 2020.
- [12] T. Silveira, M. Zhang, X. Lin, Y. Liu, and S. Ma, “How Good Your Recommender System Is? A Survey On Evaluations in Recommendation,” *International Journal of Machine Learning and Cybernetics*, vol. 10, no. 5, pp. 813–831, 2017.
- [13] N. Polatidis and C. K. Georgiadis, “A Dynamic Multi-Level Collaborative Filtering Method for Improved Recommendations,” *Computer Standards & Interfaces*, vol. 51, pp. 14–21, 2017.