

## Machine Learning Rekomendasi Produk dalam Penjualan Menggunakan Metode Item-Based Collaborative Filtering

Daniel Theodorus<sup>1✉</sup>, Sarjon Defit<sup>2</sup>, Gunadi Widi Nurchayo<sup>3</sup>

<sup>1</sup>Sentral Tukang Indonesia

<sup>2,3</sup>Universitas Putra Indonesia YPTK Padang

[dtheo0690@gmail.com](mailto:dtheo0690@gmail.com)

### Abstract

The shift towards Industry 4.0 has pushed many companies to adopt a digital system. With the sheer amount of data available today, companies start to face difficulties with providing product recommendation to their customers. As a result, data analysis has become increasingly important in the pursuit of providing the best service (user experience) to customers. The location appointed in this research is PT. Sentral Tukang Indonesia which is engaged in the sale of building materials and carpentry tools such as: paint, plywood, aluminum, ceramics, and hpl. Machine Learning has emerged as a possible solution in the field of data analysis. The recommendation system emerged as a solution in providing product recommendation based on interactions between customers in historical sales data. The purpose of this study is to assist companies in providing product recommendation to increase sales, to make it easier for customers to find the products they need, providing the best service (user experience) to customers. The data used is customer, item, and historical sales at PT. Sentral Tukang Indonesia over a time span of 1 period. Data historical sales divide to dataset training 80% and dataset testing 20%. The Item-based Collaborative Filtering method used in this study uses Cosine Similarity algorithm to calculate the level of similarity between products. Score prediction uses Weighted Sum formula while computation of error rate uses the Root Mean Squared Error formula. The result of this study shows top 10 product recommendations per customer. The products displayed are products with the highest score from the individual customer. This research can be used as a reference by companies looking to provide product recommendations needed by their customers.

**Keywords:** Machine Learning, Item-based Collaborative Filtering, Cosine Similarity, Recommendation System, Sentral Tukang Indonesia.

### Abstrak

Industri 4.0 mendorong banyak perusahaan bertransformasi ke sistem digital. *Machine Learning* merupakan salah satu solusi dalam analisa data. Analisa data menjadi poin penting dalam memberikan layanan yang terbaik (*user experience*) kepada pelanggan. Lokasi yang diangkat dalam penelitian ini adalah PT. Sentral Tukang Indonesia yang bergerak dalam bidang penjualan bahan bangunan dan alat pertukangan seperti: cat, tripleks, aluminium, keramik, dan hpl. Dengan banyaknya data yang tersedia, menyebabkan perusahaan mengalami kesulitan dalam memberikan rekomendasi produk kepada pelanggan. Sistem rekomendasi muncul sebagai solusi dalam memberikan rekomendasi produk, berdasarkan interaksi antara pelanggan dengan pelanggan lainnya yang terdapat di dalam data histori penjualan. Tujuan dari penelitian ini adalah Membantu perusahaan dalam memberikan rekomendasi produk sehingga dapat meningkatkan penjualan, memudahkan pelanggan untuk menemukan produk yang dibutuhkan, dan meningkatkan layanan yang terbaik kepada pelanggan. Data yang digunakan adalah data histori penjualan dalam 1 periode (Q1 2021), data pelanggan, dan data produk pada PT. Sentral Tukang Indonesia. Data histori penjualan tersebut akan dibagi menjadi 80% untuk dataset *training* dan 20% untuk dataset *testing*. Metode *Item-based Collaborative Filtering* pada penelitian ini memakai algoritma *Cosine Similarity* untuk menghitung tingkat kemiripan antar produk. Prediksi *score* memakai rumus *Weighted Sum* dan dalam menghitung tingkat *error* memakai rumus *Root Mean Squared Error*. Hasil dari penelitian ini memperlihatkan rekomendasi top 10 produk per pelanggan. Produk yang tampil merupakan produk yang memiliki *score* tertinggi dari pelanggan tersebut. Penelitian ini dapat menjadi referensi dan acuan bagi perusahaan dalam memberikan rekomendasi produk yang dibutuhkan oleh pelanggan.

**Kata kunci:** *Machine Learning, Item-based Collaborative Filtering, Cosine Similarity, Sistem Rekomendasi, Sentral Tukang Indonesia.*

© 2021 JiDT

### 1. Pendahuluan

PT. Sentral Tukang Indonesia merupakan salah satu toko retail yang bergerak dalam distribusi bahan bangunan. Memiliki cabang di kota Padang dan Pekanbaru. Banyaknya data pelanggan, produk dan data histori penjualan merupakan tantangan tersendiri pada perusahaan tersebut dalam memanfaatkan data. Data tersebut belum dimanfaatkan secara maksimal dalam menunjang pengambilan keputusan perusahaan.

Pergantian karyawan terutama bagian sales - admin menjadi salah satu permasalahan yang ada pada perusahaan tersebut sehingga menimbulkan masalah dalam memberikan pelayanan kepada pelanggan.

*Knowledge Discovery in Database (KDD)* merupakan sebuah proses dalam melihat suatu hubungan, pola baru dalam suatu data ataupun database yang besar dengan menggunakan pola statistik dan matematika [1]. *Data Mining* adalah bagian dalam proses dari KDD. *Data*

*Mining* merupakan penemuan data dalam basis data yang besar menggunakan pola statistik, matematik dan *Machine Learning* dalam mengambil sebuah kesimpulan dari basis data yang akan diproses. Pada masa sekarang, *Machine Learning* memakai rumus statistika dan aljabar linier.

Salah satu penerapan *Machine Learning* ini adalah dalam sistem rekomendasi [2]. Metode sistem Rekomendasi dapat dikelompokkan menjadi 2 bagian yaitu: metode *Based On Content* dan metode *Collaborative Filtering*. Metode *Collaborative Filtering* dibagi menjadi 2 macam yaitu metode *User-based Collaborative Filtering* dan metode *Item-based Collaborative Filtering*.

Sistem rekomendasi menggunakan metode *Based Collaborative Filtering* digunakan dalam bidang pariwisata. Pariwisata di Kota Malang menghasilkan nilai prediksi menunjukkan kemiripan dengan pelanggan yang lain [3]. Sistem rekomendasi agenda wisata berbasis *Android* di Kota Bengkulu berdasarkan preferensi dari pengguna sebelumnya [4].

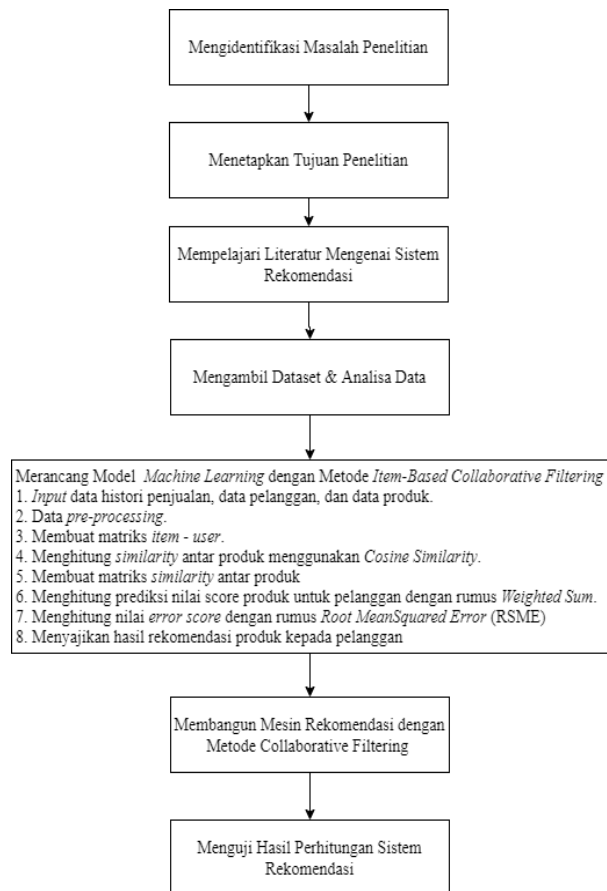
Implementasi sistem rekomendasi dengan menggunakan metode *Based Collaborative Filtering* dapat digunakan di bidang lain. Penelitian menggunakan analisa sentimen dalam pencarian hotel berdasarkan popularitas. Menghasilkan rekomendasi berkualitas tinggi dan mengurangi waktu dalam eksekusi sistem [5]. Dalam pemilihan konsentrasi mata kuliah dengan memakai data nilai mahasiswa dan data mata kuliah [6]. Rekomendasi produk kepada calon pembeli aksesoris *smartphone* berdasarkan *rating* dari setiap produk [7]. Sistem rekomendasi berbasis *Machine Learning* dalam memberikan rekomendasi musik berdasarkan preferensi pengguna menghasilkan tingkat kemiripan hingga 0,6684, serta nilai *precision* mencapai 0,125, nilai *recall* 0,200, waktu *response* mencapai 3,5 detik [8].

Penelitian membahas mengenai *improvisasi Collaborative Filtering*. *Item* yang tidak memiliki nilai dari pengguna, maka sistem rekomendasi dapat memakai top-N *items* sebagai hasil rekomendasi [9]. Penelitian mengungkapkan bahwa pendekatan dengan memakai peringkat *multi* kriteria lebih baik daripada tradisional *Item-based Collaborative Filtering* [10]. Sistem rekomendasi dapat diterapkan dalam rekomendasi peringkat buku. Semakin besar jumlah pengguna, maka semakin tinggi indeks dan akan muncul paling atas memakai algoritma *Jaccard Similarity* [11]. Perbandingan antara tradisional *Collaborative Filtering* menunjukkan bahwa algoritma yang dilakukan *Machine Learning* menghabiskan waktu yang lebih sedikit dan dapat diandalkan [12]. Algoritma dengan memakai *Item-based Collaborative Filtering* lebih dapat menangani mode pembelajaran mooc saat ini dan memberikan layanan rekomendasi lebih baik dalam pembelajaran mooc. Algoritma sistem rekomendasi mencapai hasil yang bagus [13].

Peneliti membangun sistem rekomendasi berbasis *Machine Learning* dengan menggunakan metode *Item-based Collaborative Filtering* dengan memakai algoritma *Cosine Similarity* dalam menghitung kemiripan produk menggunakan histori data penjualan. Tujuan dari penelitian ini agar pelanggan mendapatkan rekomendasi barang yang dibutuhkan sewaktu akan melakukan order sehingga dapat meningkatkan pengalaman (*user experience*) yang baik.

## 2. Metodologi Penelitian

Pada metodologi penelitian ini terdapat langkah-langkah yang dilakukan untuk menyelesaikan permasalahan yang terjadi yang berhubungan dalam sistem rekomendasi menggunakan metode *Item-Based Collaborative Filtering*. Kerangka kerja penelitian ini yang digambarkan pada Gambar 1.



Gambar 1. Kerangka Kerja Penelitian

Berdasarkan kerangka kerja penelitian di atas, maka dapat diuraikan sebagai berikut:

### a. Mengidentifikasi Masalah Penelitian.

Permasalahan yang terjadi adalah sales admin yang ada di PT. Sentral Tukang Indonesia sering mengalami pergantian, sehingga dalam memberikan rekomendasi produk kepada pelanggan tidak efektif, karena belum ada ketentuan yang baku dalam mengatur sistem rekomendasi ini.

b. Menetapkan Tujuan Penelitian.

Tujuan penelitian ini adalah untuk memberikan rekomendasi produk kepada pelanggan berdasarkan histori data penjualan.

c. Mempelajari Literatur Mengenai Sistem Rekomendasi.

Literatur yang dipakai berasal dari artikel, jurnal dan buku yang membahas mengenai sistem rekomendasi ini.

d. Mengambil Dataset & Analisa Data.

Peneliti melakukan pengamatan langsung pada perusahaan tersebut. Selain itu, peneliti juga melakukan analisa terhadap data yang digunakan dalam penelitian ini. Peneliti memperoleh data produk, data pelanggan, dan data histori penjualan pada PT. Sentral Tukang Indonesia cabang Pekanbaru selama periode Q1 2021.

e. Merancang Model *Machine Learning* dengan Metode *Item-based Collaborative Filtering*.

Tahapan merancang model dilakukan dengan melihat konfigurasi kerangka kerja yang tersusun atas desain masukan, alur proses metode *Item-based Collaborative Filtering*, desain hasil atau keluaran.

f. Membangun Sistem Rekomendasi dengan Metode *Item-based Collaborative Filtering*.

Tahapan ini dataset *training* dan *testing* terbagi atas 80%-20%. Pembagian dataset tersebut dilakukan secara *random*. Model yang dibangun berbasis *Machine Learning* dengan menggunakan *Python-Jupyter Notebook*.

g. Menguji Hasil Perhitungan Sistem Rekomendasi.

Pada Tahapan ini menghasilkan hasil rekomendasi berupa top 10 produk yang memiliki *score* tertinggi per pelanggan. Tahapan ini dilakukan *testing* untuk menganalisis hasil perhitungan yang dilakukan dengan sistem rekomendasi metode *Item-based Collaborative Filtering* dengan *Cosine Similarity*.

2.1. Machine Learning

*Machine Learning* merupakan suatu teknik dalam melakukan inferensi terhadap data dengan memakai pendekatan secara matematis. Poin utama *Machine Learning* yaitu dalam pembuatan suatu model (matematis) yang mereferensikan pola-pola data. *Machine Learning* banyak memanfaatkan statistika dan aljabar linear [2]. Dataset terdiri *Machine Learning* dari dataset *training* dan dataset *testing*. Dataset *training* digunakan dalam mencari model yang cocok dalam melatih algoritma, sementara dataset *testing* akan dipakai untuk evaluasi dari kinerja dari model yang didapat [14]. Kegunaan dari *Machine Learning* sebagai berikut [15]:

a. *Classification* adalah metode yang terdapat dalam *Machine Learning* yang dipakai untuk memprediksi suatu nilai/kelas dari individu dalam sebuah populasi.

b. *Similarity matching* adalah metode yang terdapat dalam *Machine Learning* yang digunakan untuk mengidentifikasi kemiripan antar individu berdasarkan data yang ada.

c. *Clustering* adalah metode yang terdapat dalam *Machine Learning* yang digunakan untuk membagi data - data ke dalam kelompok yang berdasarkan kriteria.

2.2. Sistem Rekomendasi

Sistem rekomendasi merupakan sebuah sistem dan teknik dalam memberikan saran baik berupa produk yang kemungkinan besar disukai / menarik bagi pengguna tertentu. Sistem rekomendasi diharapkan dapat membantu pengguna dalam mengambil sebuah keputusan, seperti rekomendasi daftar lagu yang diprediksi untuk didengarkan oleh user [8]. Secara umum terdapat dua teknik dalam sistem rekomendasi yaitu: *Content-based Filtering* dan *Collaborative Filtering*. *Content-based Filtering* berfokus pada karakteristik pengguna secara spesifik. *Collaborative Filtering* berfokus pada selera terhadap suatu produk berdasarkan opini dari orang lain [16].

Dalam *Item-based Collaborative Filtering*, preferensi pelanggan dapat dikumpulkan dengan dua cara. Pertama, secara *explicit* dengan memakai perhitungan peringkat untuk produk dalam skala tertentu. Kedua, secara *implicit* dengan memakai data histori pembelian pengguna. Rekomendasi memakai *User-based Collaborative Filtering* menunjukkan hasil yang relatif lebih akurat, sedangkan rekomendasi dengan *Item-based Collaborative Filtering* memperlihatkan hasil yang relatif lebih stabil [17].

*Cosine Similarity* merupakan suatu teknik dalam mengukur kesamaan antara dua vektor dalam suatu produk [18]. Rumus *Cosine Similarity* dapat dilihat pada Persamaan 1.

$$\text{sim}(c, d) = \frac{c \cdot d}{|c| \cdot |d|} = \frac{\sum_{i=1}^n C_i D_i}{\sqrt{\sum_{i=1}^n C_i^2} \sqrt{\sum_{i=1}^n D_i^2}} \quad (1)$$

Dimana  $\text{sim}(c, d)$  merupakan tingkat kesamaan vektor C dan vektor D,  $C_i$  dan  $D_i$  merupakan komponen dari vektor C dan D.

*Weighted Sum* merupakan penjumlahan dari perkalian rating atribut dengan bobot atribut. Untuk menghitung prediksi *score* dari produk - produk tersebut dengan membandingkan *rating* / nilai penjualan dari produk dengan kemiripan antara produk tersebut dengan produk lainnya. Rumus

Weighted Sum dapat dilihat pada Persamaan 2 [3][4].

$$S_{(C,D)} = \frac{\sum_{i \in I} (N_{C,i} * S_{i,D})}{\sum_{i \in I} |S_{i,D}|} \quad (2)$$

Dimana  $S_{(C,D)}$  merupakan prediksi nilai score produk D oleh pelanggan C,  $i \in I$  merupakan himpunan produk,  $N_{C,i}$  merupakan nilai penjualan pelanggan C pada penjualan i,  $S_{i,D}$  merupakan nilai kemiripan produk i dan produk D.

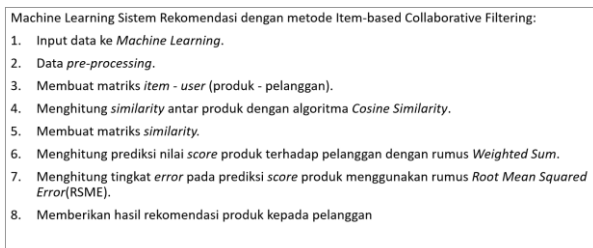
Root Mean Squared Error merupakan suatu metode pengukuran dimana mengukur perbedaan nilai dari prediksi sebuah model sebagai estimasi / nilai yang diamati. Semakin kecil nilai RSME, maka akan semakin baik rekomendasi yang dihasilkan. Nilai RSME menggunakan rumus dalam Persamaan 3.

$$\text{nilai\_me} = \sqrt{\frac{\sum_{i=1}^N (C_i - \hat{C}_i)^2}{N}} \quad (3)$$

Dimana nilai\_me merupakan nilai error,  $C_i$  merupakan nilai aktual,  $\hat{C}_i$  merupakan nilai hasil peramalan, N merupakan nilai banyak data.

### 3. Hasil dan Pembahasan

Pada tahap pembahasan ini menggunakan data histori penjualan pada PT Sentral Tukang Indonesia cabang Pekanbaru periode Q1 2021. Berikut proses metode Item-based Collaborative Filtering yang dijelaskan dalam Gambar 2.



Gambar 2. Langkah - langkah Item-based Collaborative Filtering

#### 3.1. Input Data & Data pre-processing

Data histori penjualan akan dikelompokkan berdasarkan customer\_no dan item\_no, kemudian dihitung data per baris. Hal ini bertujuan agar rentang qty tidak terlalu besar. Data - data tersebut akan dibersihkan terlebih dahulu agar tidak ada data yang bernilai kosong, NULL, maupun duplikat.

Data yang diinput dan diolah adalah data training dan data testing. Data - data tersebut akan disusun berdasarkan kebutuhan sistem. Sistem rekomendasi membutuhkan 2 komponen utama yaitu item dan user. Komponen user dalam data tersebut adalah kolom customer\_no. Komponen item dalam data tersebut adalah item\_no. Komponen value adalah qty\_sales.

#### 3.2. Membuat Matriks Item - User

Data training dan data testing akan ditransformasi menjadi matriks item - user menggunakan pivot tabel. Matriks item - user yang berasal dari data training memiliki 346 baris dan 169 kolom yang diperlihatkan pada Tabel 1. Matriks item - user yang berasal dari data testing memiliki 134 baris dan 103 kolom yang diperlihatkan pada Tabel 1 dan Tabel 2.

Tabel 1. Sampel Matriks Item - User Data Training

Item_No	100076	100109	100111	100156	100188
100473	2	0	0	0	0
102574	1	0	0	0	1
102576	0	0	0	0	1
102578	0	0	1	1	0
102580	0	0	0	0	0
102582	0	0	2	0	0

Tabel 2. Sampel Matriks Item - User Data Testing

Item_No	100076	100608	100654	103862	103872
100473	0	0	0	0	0
100803	0	2	1	3	0
100807	0	1	0	0	0
100808	0	0	0	3	0
102576	0	0	0	0	0
102578	0	0	0	0	0

#### 3.3. Menghitung Similarity Antar Produk dengan Menggunakan Algoritma Cosine similarity.

Data training dan data testing akan dihitung tingkat similarity antar produk menggunakan Cosine Similarity.

Selanjutnya menghitung tingkat similarity antar produk dengan memakai data training. Menggunakan beberapa contoh perhitungan diperoleh nilai sebagai berikut:

a. Nilai similarity produk 100808, 100807

$$C.D = (1 * 0) + (1 * 0) + (2 * 2) + (2 * 2) + (2 * 0) + (0 * 1) = 0 + 0 + 4 + 4 + 0 + 0 = 8$$

$$|C| = \sqrt{1^2 + 1^2 + 2^2 + 2^2 + 2^2 + 0^2} =$$

$$\sqrt{1 + 1 + 4 + 4 + 4 + 0} = \sqrt{14} = 3,74165$$

$$|D| = \sqrt{0^2 + 0^2 + 2^2 + 2^2 + 0^2 + 1^2} =$$

$$\sqrt{0 + 0 + 4 + 4 + 1} = \sqrt{9} = 3$$

$$\text{sim}(100808,100807) = \frac{8}{(3,74165 * 3)} = 0,712698$$

Nilai similarity produk 100808, 100809

$$C.D = (1 * 0) + (1 * 0) + (0 * 1) + (2 * 0) + (2 * 0) + (2 * 1) + (0 * 1) = 0 + 0 + 0 + 0 + 0 + 2 + 0 = 2$$

$$|C| = \sqrt{1^2 + 1^2 + 2^2 + 2^2 + 0^2} =$$

$$\sqrt{1 + 1 + 4 + 4 + 0} = \sqrt{14} = 3,74165$$

$$|D| = \sqrt{0^2 + 0^2 + 1^2 + 0^2 + 0^2 + 1^2 + 1^2} =$$

$$\sqrt{0 + 0 + 1 + 0 + 0 + 1 + 1} = \sqrt{3} = 1,73205$$

$$sim(100808,100809) = \frac{2}{(3,74165*1,73205)} = 0,308607$$

b. Nilai *similarity* produk 100808, 100804

$$C.D = (0 * 4) + (1 * 2) + (1 * 0) + (2 * 2) + (2 * 0) + (2 * 0) + (0 * 3) = 0 + 2 + 0 + 4 + 0 + 0 + 0 = 6$$

$$|C| = \sqrt{1^2 + 1^2 + 2^2 + 2^2 + 2^2 + 0^2} = \sqrt{1 + 1 + 4 + 4 + 4 + 0} = \sqrt{14} = 3,74165$$

$$|D| = \sqrt{4^2 + 2^2 + 0^2 + 2^2 + 0^2 + 0^2 + 3^2} = \sqrt{16 + 4 + 0 + 4 + 0 + 0 + 9} = \sqrt{33} = 5,77456$$

$$sim(100808,100804) = \frac{6}{(3,74165*5,77456)} = 0,279145$$

c. Nilai *similarity* produk 100804, 100807

$$C.D = (4 * 0) + (2 * 0) + (2 * 2) + (0 * 2) + (3 * 1) = 0 + 0 + 4 + 0 + 3 = 7$$

$$|C| = \sqrt{4^2 + 2^2 + 2^2 + 0^2 + 3^2} = \sqrt{16 + 4 + 4 + 0 + 9} = \sqrt{33} = 5,77456$$

$$|D| = \sqrt{0^2 + 0^2 + 2^2 + 2^2 + 1^2} = \sqrt{0 + 0 + 4 + 4 + 1} = \sqrt{9} = 3$$

$$sim(100804,100807) = \frac{7}{(5,77456*3)} = 0,40$$

d. Nilai *similarity* produk 100804, 100809

$$C.D = (4 * 0) + (2 * 0) + (0 * 1) + (2 * 0) + (0 * 1) + (3 * 0) + (0 * 1) = 0 + 0 + 0 + 0 + 0 + 0 + 0 = 0$$

$$sim(100804,100809) = 0$$

e. Nilai *similarity* produk 100807, 100809

$$C.D = (0 * 1) + (2 * 0) + (2 * 0) + (0 * 1) + (1 * 0) + (0 * 1) = 0 + 0 + 0 + 0 + 0 + 0 = 0$$

$$sim(100807,100809) = 0$$

### 3.4. Membuat Matriks Similarity Antar Produk.

Setelah menghitung nilai *similarity* antar produk, selanjutnya akan dibuat matriks *item - item*. Matriks ini bertujuan untuk melihat hubungan dan kemiripan antar 2 produk. Data *training* matrik *item - item* memiliki 346 baris dan 346 kolom Data *testing* matriks *item - item* memiliki 134 baris dan 134 kolom. Berdasarkan perhitungan *similarity* antar produk, maka matriks *similarity* pada data *training* dapat dilihat pada Tabel 3.

Tabel 3. Sampel Matriks Similarity Antar Produk Data Training

Item_no	100804	100807	100808	100809
100804	1	0,4	0,279145	0
100807	0,4	1	0,712698	0
100808	0,279145	0,712698	1	0,308607
100809	0	0	0,308607	1

Tabel 3 memperlihatkan beberapa perhitungan *similarity* antar produk. Produk 100804 memiliki tingkat kemiripan dengan produk 100807 sebesar 0,4,

produk 100807 memiliki tingkat kemiripan dengan produk 100808 sebesar 0,712698.

### 3.5. Menghitung Prediksi Score Produk dan Error per Pelanggan.

Perhitungan *score* produk dengan memakai rumus *Weighted Sum* dan tingkat *error* dengan rumus *Root Mean Squared Error* dalam memberikan rekomendasi kepada pelanggan. Perhitungan nilai *score* untuk masing - masing pelanggan dilakukan dengan menghitung tingkat *similarity* dari produk yang dipilih oleh pelanggan dengan produk histori data penjualan. Selanjutnya, nilai *similarity* tersebut akan dikalikan dengan *qty\_sales* dan ditotalkan. Kemudian, akan dibagi dengan total dari nilai *similarity* dari produk histori data penjualan. Berikut beberapa perhitungan *score* dan *error* untuk pelanggan 100654 data *training*.

a. Nilai prediksi *score* dan *error* pelanggan 100654 terhadap produk 100808 yang disajikan pada Tabel 4.

Tabel 4. Score dan Error Pelanggan 100654 Terhadap Produk 100808

Cust_no 100654	Sim terhadap produk 100808	Score	Error
Item_no	qty	qty * sim	(qty - score) <sup>2</sup>
100804	4	0,2714526	1,1165810
112609	3	0	0
110831	2	0,06952347	0,1390469
102580	1	0	0
109688	1	0	0
112217	1	0	0
112390	1	0	0
112610	1	0	0
Total		0,34866873	1,25562798
Score(100654,100808) = $\frac{\sum Score}{\sum Sim}$			3,60120616
Total			25,777251
RSME = $\sqrt{\frac{\sum Error}{count(Qty)}}$			1,7950366

b. Nilai prediksi *score* dan *error* pelanggan 100654 terhadap produk 103082 disajikan pada Tabel 5.

Tabel 5 Score dan Error Pelanggan 100654 Terhadap Produk 103082

Cust_no 100654	Sim terhadap produk 103082	Score	Error
Item_no	qty	qty * sim	(qty - score) <sup>2</sup>
100804	4	0,348155311	1,39262124
112609	3	0	0
110831	2	0	0
102580	1	0	0
109688	1	0	0
112217	1	0	0
112390	1	0	0
112610	1	0	0
Total		0,348155311	1,39262124
Score(100654,103082) = $\frac{\sum Score}{\sum Sim}$			4
Total			24,798423
RSME = $\sqrt{\frac{\sum Error}{count(Qty)}}$			1,7606257

3.6. Memberikan Hasil Rekomendasi Produk kepada Pelanggan.

Dalam memberikan rekomendasi produk kepada pelanggan dilakukan dengan menghitung nilai *score* produk untuk pelanggan tertentu. Rekomendasi yang diberikan kepada pelanggan tersebut adalah 10 produk teratas. Nilai *score* yang dipakai dalam rekomendasi ini adalah nilai *score* 10 produk tertinggi. Produk - produk yang dijadikan rekomendasi kepada pelanggan, antara 1 pelanggan dengan pelanggan lain tidaklah sama. Rekomendasi produk kepada pelanggan 100654 dengan memakai data *training* dapat dilihat pada Tabel 6.

Tabel 6. 10 Rekomendasi Produk kepada Pelanggan 100654

Score	RSME	Item_no	Item_name
4	1,7606	103082	PLAT SS 04X1X2
4	1,7190	100807	CANANG MDF KODE 6
4	1,9765	100811	CANANG PARTICLE BOARD KODE
3,715209	1,7280	100805	CANANG MDF KODE 18
3,614710	1,7510	103057	PLAT AL 04X1X2
3,601206	1,7950	100808	CANANG MDF KODE 9
3	1,8660	120240	TRIPLEX MICA OKE 3 DOF
3	1,8660	112167	CARTA HPL 1727 T DELIGHT LINEN
3	1,8660	111511	CARTA HPL 0248 T STAIN BLUE
3	1,8660	116945	EWIGE LEBEN LEM KACA STD WHITE

Rekomendasi produk menggunakan *Machine Learning* untuk pelanggan 100654 dengan menggunakan data *training* dapat dilihat pada Gambar 3.

```
get_recommendation_Training(100654)
```

100654

	predicted	error_rsme	item_no	item_name
0	4.000000	1.760626	103082	PLAT SS 04X1X2
1	4.000000	1.719082	100807	CANANG MDF KODE 6
2	4.000000	1.976509	100811	CANANG PARTICLE BOARD KODE 15
3	3.715209	1.728061	100805	CANANG MDF KODE 18
4	3.614710	1.751024	103057	PLAT AL 04X1X2
5	3.601206	1.795037	100808	CANANG MDF KODE 9
6	3.000000	1.866094	120240	TRIPLEX MICA OKE 3 DOF
7	3.000000	1.866094	112167	CARTA HPL 1727 T DELIGHT LINEN
8	3.000000	1.866094	111511	CARTA HPL 0248 T STAIN BLUE
9	3.000000	1.866094	116945	EWIGE LEBEN LEM KACA STD WHITE

Gambar 3. Rekomendasi Produk untuk Pelanggan 100654

Pada Gambar 3 dan Tabel 6 memperlihatkan perhitungan rekomendasi pada data *training* secara manual dan secara sistem memperlihatkan hasil yang sama. Rekomendasi produk kepada pelanggan 101230 dengan memakai data *testing* dapat dilihat pada Tabel 7.

Tabel 7. 10 Rekomendasi Produk kepada Pelanggan 101230

Score	RSME	Item_no	Item_name
3	1,3380	111546	CARTA HPL 0298 F CLEAR WHITE
3	1,3710	117166	BLOCK BOARD 18 U
3	1,2310	110493	KSK PIPA SS BULAT 1IN / METER
2	1,4832	110159	KLASSEN FW.9120 REL LMR AS 380
2	1,5433	116922	BLOCKMINT 1 SISI 15
2	1,5129	109225	HOTSU PAKU TEMBAK F 30
1,828427	1,4342	110830	TRIPLEX 9 IND
1,640754	1,3599	112609	BLOCK BOARD AFR 15
1	1,5743	103059	PLAT AL 06X1X2
1	1,5599	112396	CARTA HPL 7876 T SHANNON TEAK

Rekomendasi produk menggunakan *Machine Learning* untuk pelanggan 101230 dengan menggunakan data *testing* dapat dilihat pada Gambar 4.

```
get_recommendation_Testing(101230)
```

101230

	predicted	error_rsme	item_no	item_name
0	3.000000	1.338099	111546	CARTA HPL 0298 F CLEAR WHITE
1	3.000000	1.371051	117166	BLOCK BOARD 18 U
2	3.000000	1.231010	110493	KSK PIPA SS BULAT 1IN / METER
3	2.000000	1.483272	110159	KLASSEN FW.9120 REL LMR AS 380
4	2.000000	1.543315	116922	BLOCKMINT 1 SISI 15
5	2.000000	1.512988	109225	HOTSU PAKU TEMBAK F 30
6	1.828427	1.434223	110830	TRIPLEX 9 IND
7	1.640754	1.359954	112609	BLOCK BOARD AFR 15
8	1.000000	1.574356	103059	PLAT AL 06X1X2
9	1.000000	1.559930	112396	CARTA HPL 7876 T SHANNON TEAK

Gambar 4. Rekomendasi Produk untuk Pelanggan 101230

Pada gambar 4 dan Tabel 7 memperlihatkan perhitungan rekomendasi pada data *testing* secara manual dan secara sistem memperlihatkan hasil yang sama.

4. Kesimpulan

Berdasarkan hasil dan pembahasan dapat disimpulkan bahwa metode *Item-based Collaborative Filtering* dapat digunakan dalam memberikan rekomendasi produk kepada pelanggan, *Cosine Similarity* memberikan hasil yang baik dalam menghitung tingkat kemiripan antar produk, dan metode ini dapat menjadi referensi dan acuan perusahaan dalam memberikan rekomendasi kepada pelanggan.

Daftar Rujukan

[1] Sharon, Defit, S., & Nurcahyo, G. W. (2021). Tingkat Efisiensi Penggunaan Resep Dokter Spesialis Menggunakan Metode K-Means Clustering. *Jurnal Informasi Dan Teknologi*, 121–127. <https://doi.org/10.37034/jidt.v3i3.118>

[2] Jan Wira Gotama Putra, 2020. Pengenalan Konsep Pembelajaran Mesin dan Deep Learning, Tokyo.

- [3] Islamiyah, M., Subekti, P., & Andini, T. D. (2019). Utilization of CollaboratPemanfaatan Metode Based Collaborative Filtering Untuk Rekomendasi Wisata Di Kabupaten Malangive Filtering Method for Tourism Recommendations in Malang Regency. *Jurnal Ilmiah Teknologi Informasi Asia*, 13(2), 143. <https://doi.org/10.32815/jitika.v13i2.70>
- [4] Setiawan, Y., Nurwanto, A., & Erlansari, A. (2019). Implementasi Item Based Collaborative Filtering Dalam Pemberian Rekomendasi Agenda Wisata Berbasis Android. *Pseudocode*, 6(1), 13–20. <https://doi.org/10.33369/pseudocode.6.1.13-20>
- [5] Ramzan, B., Bajwa, I. S., Jamil, N., Amin, R. U., Ramzan, S., Mirza, F., & Sarwar, N. (2019). An Intelligent Data Analysis for Recommendation Systems Using Machine Learning. *Scientific Programming*, 2019, 1–20. <https://doi.org/10.1155/2019/5941096>
- [6] Jepriana I. W., Hanief S, (2020). Analisis dan Implementasi Metode Item- Based Collaborative Filtering Untuk Sistem Rekomendasi Stmik Stikom Bali. *Janapati*, 9(2), 171-180. <http://dx.doi.org/10.23887/janapati.v9i2.23218>
- [7] Prasetyo, B., Haryanto, H., Astuti, S., Astuti, E. Z., & Rahayu, Y. (2019). Implementasi Metode Item-Based Collaborative Filtering dalam Pemberian Rekomendasi Calon Pembeli Aksesoris Smartphone. *Eksplora Informatika*, 9(1), 17–27. <https://doi.org/10.30864/eksplora.v9i1.244>
- [8] Putra, A. I., & Santika, R. R. (2020). Implementasi Machine Learning dalam Penentuan Rekomendasi Musik dengan Metode Content-Based Filtering. *Edumatic : Jurnal Pendidikan Informatika*, 4(1), 121–130. <https://doi.org/10.29408/edumatic.v4i1.2162>
- [9] Feng, J., Fengs, X., Zhang, N., & Peng, J. (2018). An improved collaborative filtering method based on similarity. *PLOS ONE*, 13(9), e0204003. <https://doi.org/10.1371/journal.pone.0204003>
- [10] Kouadria, A., Nouali, O., & Al-Shamri, M. Y. H. (2019). A Multi-criteria Collaborative Filtering Recommender System Using Learning-to-Rank and Rank Aggregation. *Arabian Journal for Science and Engineering*, 45(4), 2835–2845. <https://doi.org/10.1007/s13369-019-04180-3>
- [11] Rana, A., & Deeba, K. (2019). Online Book Recommendation System using Collaborative Filtering (With Jaccard Similarity). *Journal of Physics: Conference Series*, 1362, 012130. <https://doi.org/10.1088/1742-6596/1362/1/012130>
- [12] Kumar Ojha, R., & Bhagirathi Nayak, D. (2018). Application of Machine Learning in Collaborative Filtering Recommender Systems. *International Journal of Engineering & Technology*, 7(4.38), 213. <https://doi.org/10.14419/ijet.v7i4.38.24445>
- [13] Lu, Q., & Xia, J. (2019). Research on the Application of Item-based Collaborative Filtering Algorithms in MOOC. *Journal of Physics: Conference Series*, 1302, 032020. <https://doi.org/10.1088/1742-6596/1302/3/032020>
- [14] Sinaga, A. S. R. (2020). Machine Learning Prediksi Karakter Pengguna Hastag (#) Bahasa Generasi Milenial Di Sosial Media. *Indonesian Journal of Applied Informatics*, 4(2), 165. <https://doi.org/10.20961/ijai.v4i2.41764>
- [15] Pambudi, H. K., Kusuma, P. G. A., Yulianti, F., & Julian, K. A. (2020). PREDIKSI STATUS PENGIRIMAN BARANG MENGGUNAKAN METODE MACHINE LEARNING. *Jurnal Ilmiah Teknologi Infomasi Terapan*, 6(2), 100–109. <https://doi.org/10.33197/jitter.vol6.iss2.2020.396>
- [16] Sari, K. R., Suharso, W., & Azhar, Y. (2020). Pembuatan Sistem Rekomendasi Film dengan Menggunakan Metode Item Based Collaborative Filtering pada Apache Mahout. *Jurnal Repositor*, 2(6), 767. <https://doi.org/10.22219/repositor.v2i6.936>
- [17] Wiputra, M. M., & Shandi, Y. J. (2021). Perancangan Sistem Rekomendasi Menggunakan Metode Collaborative Filtering dengan Studi Kasus Perancangan Website Rekomendasi Film. *Media Informatika*, 20(1), 1–18. <https://doi.org/10.37595/mediainfo.v20i1.54>
- [18] Jaja, V. L., Susanto, B., & Sasongko, L. R. (2020). Penerapan Metode Item-Based Collaborative Filtering Untuk Sistem Rekomendasi Data MovieLens. *d’CARTESIAN*, 9(2), 78. <https://doi.org/10.35799/dc.9.2.2020.28274>