

# Pengembangan Model Rekomendasi Produk UMKM Albis Menggunakan Item Based Collaborative Filtering

## *Development of Albis MSME Product Recommendation Model Using Item Based Collaborative Filtering*

Muhammad Hafizh Al Mustofa<sup>a,1,\*</sup>, Nurmalitasari<sup>a,2</sup>, Nurchim<sup>b,3</sup>

<sup>a</sup>Sistem Informasi, Universitas Duta Bangsa Surakarta, Surakarta, Indonesia

<sup>b</sup>Teknik Informatika, Universitas Duta Bangsa Surakarta, Surakarta, Indonesia

<sup>1</sup>202021120@mhs.udb.ac.id; <sup>2</sup>Nurmalitasari@udb.ac.id; <sup>3</sup>Nurchim@udb.ac.id  
\*corresponding author

Informasi Artikel	ABSTRAK
<p>Diserahkan : 10 Juni 2024 Diterima : 27 Juni 2024 Direvisi : 27 Juni 2024 Diterbitkan : 12 Juli 2024</p> <p><b>Kata Kunci:</b> Rekomendasi Item based collaborative filtering Cosine Similarity</p>	<p>UMKM atau Usaha Mikro, Kecil, dan Menengah adalah salah satu komponen penting yang mendukung perekonomian Indonesia, dijalankan oleh perorangan atau badan usaha dengan penghasilan tertentu setiap tahunnya. UMKM Albis, yang bergerak di bidang penjualan produk frozen food sejak tahun 2020 dan menawarkan 40 produk dengan berbagai merek. Banyaknya pilihan membuat pelanggan kesulitan dalam menentukan produk yang akan dibeli. Penelitian ini bertujuan memudahkan konsumen dalam memilih produk menggunakan model <i>item-based collaborative filtering</i>. Metode ini memberikan rekomendasi berdasarkan kemiripan antar produk menggunakan cosine similarity dan prediksi rating pengguna. Model ini menunjukkan kinerja yang baik dengan nilai <i>MSE</i> terendah sebesar 0,05416 dan <i>RMSE</i> sebesar 0,232724, Evaluasi tersebut menunjukkan bahwa model memiliki tingkat kesalahan yang rendah.</p>
<p><b>Keywords:</b> Recommendation Item based collaborative filtering Cosine similarity.</p> <p>This is an open access article under the <a href="https://creativecommons.org/licenses/by-sa/4.0/">CC-BY-SA</a> license.</p> 	<p><b>ABSTRACT</b></p> <p><i>UMKM, or Micro, Small, and Medium Enterprises, are a vital component supporting Indonesia's economy, run by individuals or business entities with specific annual incomes. UMKM Albis, operating in frozen food sales since 2020 and offers around 40 products with various brands. The wide range of choices makes it difficult for customers to decide which products to buy. This study aims to simplify the selection process for customers using an item-based collaborative filtering model. This method provides recommendations based on item similarity using cosine similarity and user rating predictions. The model demonstrates good performance with the lowest MSE of 0.05416 and RMSE of 0.232724, The evaluations shows that model has indicating a low error rate.</i></p>

### I. Pendahuluan

UMKM menjadi salah satu komponen dalam perekonomian negara Indonesia, usaha UMKM dapat dijalankan perorangan atau badan usaha dengan modal bersih dan penghasilan tertentu. Pada tahun 1997-1998 UMKM pelaku usaha mikro mengalami pasca krisis kenaikan ekonomi, karena pada tahun tersebut banyak perusahaan yang bangkrut, akan tetapi UMKM tetap dapat bertahan. Kerugian pada UMKM tidak terlalu berpengaruh pada pelaku UMKM karena usaha ini tidak memerlukan modal yang besar dan mata uang asing. UMKM menjadi peluang bagi masyarakat kecil untuk memperbaiki taraf hidup ekonomi mereka, karena UMKM tidak memerlukan modal yang besar untuk memulai sebuah usaha[1]. Selain itu, UMKM dapat meningkatkan kesejahteraan masyarakat dan menyumbang devisa yang besar bagi negara, sehingga perkembangan UMKM dapat memasuki pangsa nasional maupun internasional.

Albis adalah salah satu UMKM yang berdiri sejak tahun 2020, yang didirikan oleh bu Mahari dan menjual berbagai macam produk salah satunya adalah frozen food. Usaha ini berlokasi di Dusun III, Jaten, Kec. Jaten, Kabupaten Karanganyar, Jawa Tengah 57731. UMKM tersebut menjual 50 produk. Banyaknya produk yang dijual dengan berbagai macam jenis dan merk, membuat konsumen atau pelanggan sulit untuk menentukan barang yang dibeli sesuai dengan kebutuhan. Kesulitan tersebut dapat mengurangi minat belanja dan mengakibatkan berkurangnya kepuasan pelanggan terhadap toko.

Sistem rekomendasi termasuk ke dalam sistem pendukung keputusan karena dapat meningkatkan keefektifan dalam memilih keputusan dan membantu konsumen untuk memilih produk yang dibutuhkan[2]. Sistem pendukung keputusan adalah sistem yang memiliki kemampuan untuk mengkomunikasikan dan menyelesaikan masalah yang bersifat semi terstruktur[3]. Banyaknya penelitian yang terkait dengan sistem

rekomendasi untuk mempermudah dalam pengelompokkan produk yang diminati oleh para konsumen, salah satunya adalah pada penelitian implementasi metode *item based collaborative filtering* dalam pemberian rekomendasi calon pembeli smartphone pada toko flazzstore. Data pada penelitian ini menggunakan 3 dataframe user, produk, dan rating. Hasil menunjukkan bahwa model atau metode perhitungan *collaborative filtering* dapat berjalan dengan baik pada rekomendasi produk, menggunakan dataframe rating dengan nilai *MSE* yaitu sebesar 0,572039 dengan eksekusi program 6,4 detik [4].

Implementasi metode ini juga dapat digunakan pada data movieleense yang berupa kumpulan data berisi 100.000 rating dengan skala 1-5 dari 943 pengguna pada 1664 film, dengan 3 dataframe yaitu user, movies, dan ratings. Hasil penelitian tersebut adalah film Crash (1996) dan Lost Highway (1997) menempati urutan paling atas sebagai film populer dengan rekomendasi rating yang tinggi. Hasil model *collaborative filtering* menggunakan nilai Cosine pada film ini mencapai kurang lebih 0.98 dengan nilai *MSE* 0,9561 dan mendapat rekomendasi tertinggi dibandingkan dengan film yang lain [5]. Hasil tersebut dapat diterima dengan baik karena nilai cosine cukup tinggi dan memiliki nilai eror yang rendah.

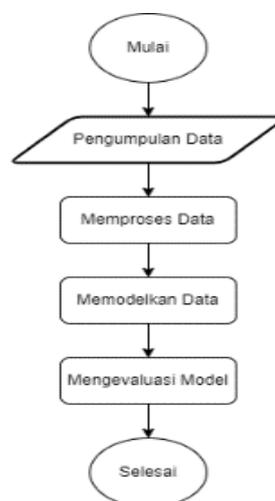
Penggunaan metode *item based collaborative filtering* dibandingkan dengan euclidan distance pada model rekomendasi buku menggunakan data yang berasal dari Kaggle dengan 3 dataframe yaitu user, ratings, dan books. Pada data ratings terdapat variabel user id, isbn, dan rating. Pada hasil evaluasi menunjukkan bahwa metode cosine similiarity mendapatkan nilai MAE yang rendah yaitu 0.4968783248 dibandingkan dengan euclidan distance dengan nilai MAE sebesar 0.5219753823 [6]. Dari hasil tersebut membuktikan bahwa cosine similiarty menghasilkan nilai yang lebih cocok karena memiliki nilai MAE yang lebih rendah dan dapat digunakan pada sistem rekomendasi.

Penelitian ini menggunakan metode analisis *item based collaboration filtering* untuk mengoptimalkan pemasaran sebuah produk dalam perusahaan untuk membantu menemukan produk yang diinginkan oleh pelanggan. Metode ini bertujuan untuk menghitung kemiripan antar item berdasarkan preferensi user atau pengguna sebelumnya, sehingga item dapat direkomendasikan kepada user kembali dan memiliki proses prediksi yang lebih baik dari *user based collaborative filtering* dalam menentukan kemiripan antar item[7]. Penelitian ini bertujuan untuk memudahkan para konsumen memilih produk yang akan dibeli karena pada UMKM albis memiliki 50 produk dengan jenis dan merk yang berbeda-beda dan model *item collaborative filtering* dapat mengatasi permasalahan tersebut menggunakan perhitungan *cosine similiarity*. *Cosine Similiarty* digunakan untuk menentukan nilai kemiripan antar produk, sehingga memunculkan produk serupa yang dicari oleh pelanggan [8].

## II. Metode

### A. Pengumpulan Data

Data dalam penelitian berasal dari catatan produsen albis dan bersifat primer yaitu data yang diambil secara langsung dari objek penelitian [9]. Terdapat 3 dataframe yang berbeda yaitu data user dengan variabel uid, profesi, dan umur dan berjumlah 27 data. Data produk dengan variabel idp, nama produk, harga, berat gram, ukuran, merk dan berjumlah 336 data. Data rating dengan variabel uid, idp dan rating berjumlah 1068 data. Maka jumlah data yang digunakan adalah 1434 data. Data diambil dari bulan januari-mei 2023. Analisa data menggunakan metode *Item Collaborative Filtering*, dengan tahapan sebagai berikut :



Gambar 1. Tahapan Penelitian

### B. Memproses Data

Pada tahap ini dimulai dengan menentukan nilai matrix dan menormalisasi data agar data dapat berjalan dengan baik pada model yang digunakan. Nilai matrix berfungsi untuk mengidentifikasi nilai kemiripan antar item dalam perhitungan *cosine similarity*.

### C. Memodelkan Data

Tahap pemodelan data menggunakan *item collaborative filtering* dilakukan dengan menghitung nilai kemiripan antar dua item atau lebih pada data menggunakan *cosine similarity*, setelah mendapatkan nilai dari *cosine similarity* maka dapat menghitung prediksi rating milik para pelanggan. Tujuan dari perhitungan ini adalah untuk mengetahui seberapa mirip dua item yang memiliki sudut kosinus yang sama. Jika nilainya 0 maka dua item tersebut sama sekali tidak memiliki kemiripan [10]

### D. Mengevaluasi Model

Mengevaluasi model dari hasil yang diperoleh menggunakan metode *Mean Squared Error (MSE)* dan *Root Mean Squared Error (RMSE)*. Tujuan evaluasi metode *MSE* dan *RMSE* adalah mengukur error pada model yang digunakan. Jika nilai error mendekati 0 maka error semakin rendah dan model bekerja dengan baik.

## III. Hasil dan Pembahasan

### A. Memproses Data

Nilai matrix adalah susunan bilangan yang membentuk persegi panjang yang berisi kolom dan baris yang sesuai dengan data yang digunakan [11]. Salah satu fungsi matrix adalah sebagai operasi dalam analisis data dan perhitungan statistik menggunakan machine learning untuk menormalisasi data.

Tabel 1. Matrix Data Rating dan Produk

idp	Uid				
	1	2	3	4	5
1	4.580556	4.587778	4.472973	4.574074	4.537755
2	4.900000	4.700000	4.400000	4.574074	4.800000
3	4.580556	5.000000	4.900000	5.000000	4.800000
4	4.580556	4.587778	4.472973	4.574074	4.600000
5	5.000000	5.000000	4.472973	5.000000	5.000000

Perhitungan nilai matrix digunakan untuk mengecek data terdapat nilai NaN atau tidak. Nilai NaN adalah nilai yang bersifat tidak angka atau Not a Number sehingga tidak dapat di hitung dan menyebabkan error pada model yang digunakan. Setelah nilai NaN dihilangkan maka data dapat di normalisasi. Menormalisasi data digunakan untuk menguji apakah data data dapat berdistribusi secara normal, dengan menghilangkan nilai yang bersifat ambiguity dan bersifat NaN [12]. Setelah nilai NaN dihilangkan maka data dapat di normalisasi agar menghilangkan data yang bersifat ambiguity, contoh adalah data yang memiliki nilai yang sama atau terduplikasi.

Tabel 2. Normalisasi Data

idp	Uid				
	1	2	3	4	5
1	1	0.580556	0.656481	0.560811	0.574074
2	2	0.900000	0.750000	0.500000	0.574074
3	3	0.580556	1.000000	0.916667	1.000000
4	4	0.580556	0.656481	0.560811	0.574074
5	5	1.000000	1.000000	0.560811	1.000000

Dengan menormalisasi data rating dan produk maka nilai dalam tabel dapat diperkecil dan meminimilisir eror dalam model yang digunakan pada tahap selanjutnya.

### B. Memodelkan Data

Pemodelan dilakukan menggunakan metode *Collaborative Filtering*, metode ini mengevaluasi atau menyaring item dengan menggunakan persepsi atau pendapat orang lain. Hasil penyaringan semua pengguna digunakan untuk membuat rekomendasi [13]. Salah cara untuk mendapatkan hasil tersebut adalah menggunakan *cosine similarity*.

*Cosine similarity* digunakan untuk perhitungan mencari nilai kemiripan antar dua item atau teks pada sebuah dokumen. Item pada dokumen tersebut dianggap sebagai sebuah vector dan nilai dari vector tersebut bersifat term frequency [14].

101	1.00	0.98	0.97	0.99	0.93
102	0.98	1.00	0.98	0.99	0.96
103	0.97	0.98	1.00	0.98	0.96
104	0.99	0.99	0.98	1.00	0.94
105	0.93	0.96	0.96	0.94	1.00

Gambar 2. Hasil Cosine Similiarty pada 5 Produk Teratas

Hasil cosine similarity atau persamaan antar item pada pada 5 produk teratas dapat di simpulkan bahwa item ke 4 atau item dengan idp 104 (roti burger bernandi) memiliki kemiripan tertinggi dengan item pertama dan kedua dengan nilai 0,99 dan kemiripan terendah terdapat pada item kelima dengan nilai 0,94.

Tabel 3. Hasil Cosine Similiarity Roti Burger Bernandi

Nama_produk	Kemiripan_produk
Sosis Kanzler Sapi Cocktail	0.991667
Cireng Rujak Bonju	0.990470
Bakso Sapi Cap Semar	0.990395
Sosis Ayam Champ	0.989992
Dada Ayam Nindya Utama	0.989722

Pada tabel diatas menunjukkan hasil cosine similarity tertinggi dari produk roti burger bernandi dengan nilai paling tinggi pada produk sosis kanzler sapi cocktail dengan nilai 0,991667. Nilai tersebut tergolong tinggi, karena pada cosine similarity jika nilai yang dihasilkan mendekati angka 1 maka nilai persamaan semakin tinggi [15].

### C. Mengevaluasi Model

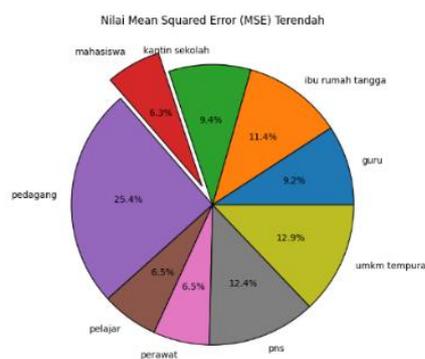
Prediksi atau peramalan adalah metode yang digunakan dalam machine learning dan data mining [16]. Prediksi berfungsi memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu[17]. Adanya hasil perhitungan cosine similarity maka dapat menghasilkan prediksi rating agar dapat mengetahui rekomendasi produk berdasarkan user dengan variabel profesi.

Tabel 4. Hasil Evaluasi Prediksi Rating

Profesi	Nama_produk	Prediksi_rating	MSE	RMSE
Guru	Cireng original family group	4.609430	0.079591	0.282118
Ibu rumah tangga	Sosis ayam mister max	4.621521	0.098326	0.313570
Kantin sekolah	Sosis ulir cocktail umami	4.496042	0.080856	0.284352

Mahasiswa	Cireng original family group	4.597828	0.054161	0.232724
Pedagang	Sosis ayam umiami	4.570625	0.219102	0.468083
Pelajar	Cireng original family group	4.504017	0.055996	0.236634
Perawat	Sosis ayam mister max	4.582421	0.055627	0.235855
Pns	Sosis ayam mister max	4.571109	0.107122	0.327294
Umkm tempura	Chicken katzu san	4.53496	0.111544	0.333982
Guru	Cireng original family group	4.609430	0.079591	0.282118

Hasil rekomendasi dari model menunjukkan bahwa mahasiswa merekomendasikan produk cireng original family group dengan prediksi rating 4.5 dan menjadi prediksi rating tertinggi karena memiliki nilai  $MSE$  0,054161 dan  $RMSE$  0,232724 terendah. Hasil tersebut menunjukkan bahwa error memiliki kategori yang rendah, karena jika nilai eror semakin mendekati nilai 0 maka eror semakin rendah [18].



Gambar 3. Nilai MSE Terendah Menggunakan Pie Chart

Berdasarkan hasil rekomendasi berdasarkan user, dapat divisualisasikan ke dalam bentuk persentase, diagram tersebut menunjukkan bahwa mahasiswa hanya memiliki persentase error terendah dengan 6,3% dan disusul dengan pelajar 6,5%, perawat 6,5%, dan guru sebanyak 9,2%.

#### IV. Kesimpulan dan saran

Metode *item-based collaborative filtering* terbukti efektif memudahkan pengguna dengan menampilkan produk serupa. Karena pada roti Bernardi dengan sosis Kanzler sapi cocktail, cireng rujak Bonju, dan bakso sapi cap Semar mempunyai nilai kemiripan 0,99 dan evaluasi model menunjukkan nilai  $MSE$  terendah 0,05416 dan  $RMSE$  0,232724, nilai tersebut menunjukkan tingkat error rendah. Penggunaan rekomendasi diterapkan pada berbagai e-commerce seperti Shopee, Amazon, dan Tokopedia, sehingga menegaskan pentingnya rekomendasi dalam dunia digital. Penelitian ini dapat menjadi landasan bagi penelitian selanjutnya menggunakan metode *content-based*, *user-based*, dan *hybrid collaborative filtering*.

#### Daftar Pustaka

- [1] Kadani dan N. Srijani, "Peran UMKM (Usaha Mikro Kecil Menengah) Dalam Meningkatkan Kesejahteraan Masyarakat," *Equilibrium*, vol. 8, no. 2, hlm. 191–201, Jun 2020, doi: <http://doi.org/10.25273/equilibrium.v8i2.7118>.

- [2] M. Faulah dan E. Irawadi Alwi, "Sistem Informasi dan Teknologi Islam Sistem Pendukung Keputusan Guru Honorer Terbaik Menggunakan Metode Weighted Product," *BUSITI*, vol. 4, no. 4, hlm. 382–389, Nov 2023, doi: <https://doi.org/10.33096/busiti.v4i4.1924>.
- [3] I. As dan N. Kurniati, "Sistem Informasi dan Teknologi Islam Sistem Pendukung Keputusan Pemilihan Bibit Padi Unggul menggunakan Metode Topsis Technique for Order Preference by Similarity to Ideal," *BUSITI*, vol. 3, no. 4, hlm. 258–267, Nov 2022, doi: <https://doi.org/10.33096/busiti.v3i4.1446>.
- [4] B. Prasetyo, H. Haryanto, S. Astuti, E. Z. Astuti, dan Y. Rahayu, "Implementasi Metode Item-Based Collaborative Filtering dalam Pemberian Rekomendasi Calon Pembeli Aksesoris Smartphone," *Eksplora Informatika*, vol. 9, no. 1, hlm. 17–27, Sep 2019, doi: 10.30864/eksplora.v9i1.244.
- [5] Y. Visher Laja Jaja, B. Susanto, L. Ricky Sasongko, dan K. Kunci, "Penerapan Metode Item-Based Collaborative Filtering Untuk Sistem Rekomendasi Data MovieLens," *Jurnal Matematika dan Aplikasi*, vol. 9, no. 2, hlm. 78–83, Sep 2020, doi: <https://doi.org/10.35799/dc.9.2.2020.28274>.
- [6] M. I. Wardah dan S. D. Putra, "Implementasi Machine Learning Untuk Rekomendasi Film Di Imdb Menggunakan Collaborative Filtering Berdasarkan Analisa Sentimen IMDB," *Jurnal Manajemen Informatika Jayakarta*, vol. 2, no. 3, hlm. 243, Jul 2022, doi: 10.52362/jmijayakarta.v2i3.868.
- [7] W. Jepriana dan S. Hanief, "Analisis dan Implementasi Metode Item-Based Collaborative Filtering untuk Sistem Rekomendasi Konsentrasi di Stmik Stikom Bali," *JANAPATI*, vol. 9, no. 2, hlm. 1–10, Mei 2020, doi: 10.23887/janapati.v9i2.23218.
- [8] A. Azhari, E. Buulolo, dan N. Silalahi, "Sistem Rekomendasi Dosen Pendamping Skripsi Berbasis Text Rank menggunakan Metode Cosine Similarity," *Pelita Informatika : Informasi dan Informatika*, vol. 10, no. 3, 2022.
- [9] P. Rifka Agustianti dkk., *Metode Penelitian Kuantitatif dan Kualitatif Tohar Media*, 3 ed. Gowa: Tohar Media, 2022.
- [10] Y. Setiawan, A. Nurwanto, dan A. Erlansari, "Implementasi Item Based Collaborative Filtering Dalam Pemberian Rekomendasi Agenda Wisata Berbasis Android," *Jurnal Pseudocode*, vol. 4, no. 1, hlm. 13–20, Feb 2019, doi: <https://doi.org/10.33369/pseudocode.6.1.13-20>.
- [11] M. pd. Prof.Dr.Hj.Rahayu Karidinata, *Aljabar Matriks Elementer*, 2 ed. Bandung: CV.PUSTAKA SETIA, 2019.
- [12] U. Azmi, Z. N. Hadi, dan S. Soraya, "ARDL METHOD: Forecasting Data Curah Hujan Harian NTB," *Jurnal Varian*, vol. 3, no. 2, hlm. 73–82, Mei 2020, doi: 10.30812/varian.v3i2.627.
- [13] H. Februariyanti, A. Dwi Laksono, J. Sasongko Wibowo, dan M. Siswo Utomo, "Implementasi Metode Collaborative Filtering untuk Sistem Rekomendasi Penjualan pada Toko Mebel," *Jurnal Khatulistiwa Informatika*, vol. 9, no. 1, hlm. 43–50, Jun 2021, doi: <https://doi.org/10.31294/jki.v9i1.9859.g4873>.
- [14] A. Lia Hananto, S. Sulaiman, S. Widiyanto, dan A. Yuniar Rahman, "Evaluation comparison of wave amount measurement results in brass-plated tire steel cord using RMSE and cosine similarity," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 1, hlm. 207–214, Apr 2021, doi: 10.11591/ijeecs.v22.i1.pp207-214.
- [15] D. Jurafsky dan J. H. Martin, *Speech and Language Processing An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition Third Edition*, 3 ed. California: Stanford University, 2024.
- [16] Nurmalitasari dan Eko Purwanto, "Prediksi Performa Mahasiswa Menggunakan Model Regresi Logistik," *Jurnal Derivat*, vol. 9, no. 2, hlm. 145–152, Des 2022, doi: <https://doi.org/10.31316/jderivat.v9i2.2639>.
- [17] P. B. Lestari Lokapitasari dan dan Herdianti, "Sistem Informasi dan Teknologi Islam Implementasi Algoritma Neural Network untuk Memprediksi Harga Bawang Merah di Kabupaten Bima," *BUSITI*, vol. 4, no. 2, hlm. 112–119, Mei 2023, doi: <https://doi.org/10.33096/busiti.v4i2.1661>.

- 
- [18] M. N. Faruqhy, D. Andreswari, dan J. P. Sari, "Prediksi Prestasi Nilai Akademik Mahasiswa Berdasarkan Jalur Masuk Perguruan Tinggi Menggunakan Metode Multiple Linear Regression (Studi Kasus: Fakultas Teknik Universitas Bengkulu)," *Jurnal Rekursif*, vol. 9, no. 2, hlm. 172–183, Nov 2021, doi: <https://doi.org/10.33369/rekursif.v9i2.17108>.