

Performance Analysis of the Item-Based Collaborative Filtering Model in Yogyakarta Tourism Recommendations

Melany Mustika Dewi ^{1*}, Ria Andriani ^{2**}, M. Nuraminudin ^{3*}

* D3 Manajemen Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta

** D3 Teknik Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta
melany@amikom.ac.id ¹

Article Info

Article history:

Received 2024-11-11

Revised 2024-12-06

Accepted 2025-01-21

Keyword:

Item Based,
Collaborative Filtering,
Recommendation System,
Yogyakarta Tourism.

ABSTRACT

Yogyakarta is one of the most popular tourist destinations in Indonesia, offering a variety of attractions ranging from beaches and mountains to historical sites. This diversity poses a challenge for tourists in selecting destinations that match their preferences. This study employs the Item-Based Collaborative Filtering method to recommend tourist destinations based on the similarity between attractions, calculated using cosine similarity. The data analyzed includes 1,069 tourist destinations in Yogyakarta, obtained from Google Maps API, Scrapetable, and Outscraper. The results indicate that the developed recommendation model achieves high accuracy with a Mean Absolute Error (MAE) of 2.537. Compared to previous approaches, this method improves the relevance and quality of recommendations, helping tourists find destinations that suit their preferences. This study contributes to the development of more personalized and effective recommendation systems for the tourism sector.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Berdasarkan laporan Badan Pusat Statistik (BPS), Yogyakarta memiliki jumlah kunjungan wisatawan lokal sebanyak 30.761.9191 pada tahun 2023.[1] Provinsi Daerah Istimewa Yogyakarta menawarkan beragam jenis tujuan wisata, mulai dari pegunungan, pantai, situs sejarah, hingga kebudayaan yang khas. Yogyakarta, sebagai salah satu kota tujuan wisata utama di Indonesia, menyanggah reputasi sebagai destinasi yang kaya akan budaya dan sejarah.[2] Namun, keberagaman wisata ini merupakan tantangan bagi wisatawan dalam memilih destinasi yang sesuai dengan preferensi mereka yang beragam.

Perbedaan preferensi di kalangan wisatawan menuntut adanya metode yang efektif dalam membantu mereka menemukan destinasi yang tepat. Sistem rekomendasi dapat menjadi solusi yang penting untuk menghadapi tantangan ini dengan menghubungkan wisatawan ke informasi yang relevan dan personal.[3] Sistem rekomendasi adalah sistem yang menyaring informasi dan menyajikan rekomendasi berdasarkan personalisasi dan penilaian pengguna.[4] Dalam konteks ini, sistem rekomendasi menggunakan algoritma

seperti data mining dan machine learning untuk menganalisis pola perilaku pengguna dan memberikan rekomendasi yang sesuai [2].

Collaborative Filtering (CF) adalah salah satu metode Sistem Rekomendasi untuk menyaring data dan memberikan rekomendasi yang sesuai dengan preferensi pengguna target berdasarkan tingkat kemiripannya dengan pengguna lain.[5]

Teknik CF sangat bergantung pada ukuran kesamaan yang digunakan untuk menentukan tingkat hubungan antara dua pengguna atau dua item. CF memiliki beberapa metode rekomendasi yaitu Item Based Collaborative Filtering dan User Based Collaborative Filtering [6]. Metode User Based Collaborative Filtering (UCBF) mencari pola rating yang diberikan oleh satu user pada item yang disukai dan menggunakan kemiripan dari pola tersebut untuk merekomendasikan item berdasarkan kemiripan sspiralaku pengguna lain [7]. Sedangkan pada data wisata banyak destinasi yang tidak dikunjungi oleh user yang sama dan tidak diberikan rating sehingga menghasilkan data yang sparsity tinggi yang kurang cocok untuk digunakan pada metode UCBF.

Pendekatan Item-Based Collaborative Filtering (IBCF) merupakan salah satu metode populer dalam sistem rekomendasi. Metode ini bekerja dengan memberikan rekomendasi berdasarkan kemiripan antar item yang telah dipilih oleh pengguna lain dengan preferensi serupa. Dengan demikian, rekomendasi tidak hanya bergantung pada pengguna dengan preferensi yang sama, tetapi juga pada item-item yang sering dipilih bersama oleh pengguna lain [8] Item-Based Collaborative Filtering memiliki beberapa keunggulan, di antaranya kemampuan untuk memberikan rekomendasi yang lebih stabil dibandingkan dengan User-Based Collaborative Filtering, terutama ketika data pengguna tidak lengkap atau terbatas [9]. Metode item mengurangi cold-start dan mempertahankan fitur yang sesuai dengan CF [10]. Dengan menggunakan pendekatan ini, sistem dapat mengidentifikasi item yang memiliki kesamaan, sehingga meningkatkan relevansi rekomendasi yang diberikan kepada pengguna [11].

Penelitian serupa mengenai sistem rekomendasi wisata telah berkembang pesat dalam beberapa tahun terakhir, dengan fokus pada penggunaan teknik Item Collaborative Filtering (CF) untuk memberikan rekomendasi yang relevan dan akurat. Studi oleh Mahendra et al, berfokus pada implementasi teknik ini untuk menghasilkan rekomendasi film berdasarkan kesamaan antar item. Menggunakan dataset dari Kaggle, penelitian ini menerapkan cosine similarity sebagai metode untuk menghitung kesamaan antar film dan mengevaluasi kinerja model dengan metrik Mean Absolute Error (MAE) dan Root Mean Square Error (RMSE). Hasil penelitian menunjukkan bahwa model ini mencapai nilai MAE sebesar 3.05 dan RMSE sebesar 3.26, yang menandakan bahwa model tersebut memberikan akurasi yang baik dalam merekomendasikan film, membuktikan bahwa Item-Based Collaborative Filtering dapat menghasilkan rekomendasi yang relevan dan personal bagi pengguna. [12]

Dalam studi Asad dan Zakariyah, sistem rekomendasi wisata dikembangkan dengan tujuan membantu wisatawan menemukan destinasi wisata di Jawa Tengah yang sesuai dengan preferensi mereka. Teknik Item Collaborative Filtering yang digunakan dalam penelitian ini bekerja dengan menemukan kesamaan antar pengguna berdasarkan preferensi yang telah mereka tunjukkan, seperti harga tiket, jarak dari lokasi pengguna, dan fasilitas yang tersedia. Pendekatan ini memungkinkan sistem untuk memberikan rekomendasi yang dipersonalisasi, yang pada gilirannya dapat membantu wisatawan membuat keputusan yang lebih tepat terkait destinasi wisata. Selain manfaat personalisasi, sistem ini juga berpotensi meningkatkan perekonomian lokal di daerah wisata dengan cara memudahkan wisatawan dalam merencanakan perjalanan mereka, yang pada akhirnya dapat meningkatkan arus wisatawan ke daerah tersebut. Penelitian ini mencatat bahwa nilai rata-rata akurasi dan Mean Absolute Error (MAE) masing-masing adalah 80,63% dan 0,9689. [13]

Penelitian yang dilakukan oleh Wibisono dkk pada tahun 2021 berjudul "Sistem Rekomendasi Suku Cadang Berdasarkan Item Based Filtering". Fokus utama penelitian

ini adalah membangun sistem rekomendasi suku cadang untuk bengkel motor, dengan tujuan meningkatkan efisiensi dalam memilih suku cadang yang tepat. Metode utama yang digunakan dalam penelitian ini adalah Singular Value Decomposition (SVD), yang terbukti memberikan hasil yang lebih baik dibandingkan metode lain seperti K-Nearest Neighbors (KNN). Penelitian ini menggunakan data transaksi servis bengkel motor selama enam bulan, dengan hasil yang menunjukkan nilai MAE sebesar 1.2752, RMSE sebesar 1.4882, dan FCP sebesar 0.4947. Validasi dilakukan menggunakan data rating acak, meskipun salah satu keterbatasannya adalah data rating yang digunakan merupakan data buatan, sehingga tidak sepenuhnya merepresentasikan kondisi nyata. Evaluasi performa menunjukkan bahwa metode SVD lebih efektif dibandingkan dengan KNN. [14]

Berdasarkan penelitian Agustian dkk, dapat disimpulkan bahwa penggunaan metode Collaborative Filtering terbukti efektif dalam membangun sistem rekomendasi tempat kos yang sesuai dengan kebutuhan dan preferensi pengguna. Dengan akurasi sebesar 29,68% berdasarkan Mean Absolute Error (MAE) dan tingkat penerimaan pengguna sebesar 83,2% dari hasil User Acceptance Testing (UAT), sistem ini mampu memberikan prediksi yang cukup akurat dan diterima dengan baik oleh para pengguna. Hasil ini menunjukkan bahwa sistem rekomendasi ini berguna bagi mahasiswa dalam menemukan tempat kos yang sesuai. [15]

Sistem rekomendasi berbasis Item-Based Collaborative Filtering digunakan untuk membantu pengguna menemukan buku yang sesuai dengan preferensi mereka. Collaborative Filtering bekerja dengan menggunakan data historis berupa rating yang diberikan oleh pengguna terhadap item, dalam hal ini buku, untuk mengidentifikasi pola kesukaan. Metode Item-Based Collaborative Filtering dalam sistem rekomendasi buku ini menghasilkan nilai Mean Absolute Error (MAE) sebesar 1.2912 berdasarkan hasil rating dari 100 anggota yang dikumpulkan melalui kuesioner pada periode April hingga Juni. Oleh karena itu, metode Item-Based Collaborative Filtering telah terbukti efektif dalam konteks perpustakaan, membantu pengguna menemukan buku baru yang relevan dan meningkatkan pengalaman membaca mereka. [16]

Penelitian Kurniawan dkk, berhasil mengembangkan sistem rekomendasi Channel YouTube bertema resep masakan menggunakan metode Collaborative Filtering yang menggabungkan teknik Cosine Similarity dan Weighted Sum. Cosine Similarity digunakan untuk mengukur kemiripan Channel berdasarkan histori tontonan pengguna, sementara Weighted Sum menghitung bobot dari setiap resep yang direkomendasikan guna meningkatkan akurasi rekomendasi. Hasil pengujian menunjukkan sistem mampu memberikan rekomendasi dengan nilai rata-rata MAE sebesar 2,103, menggunakan 1000 data Channel dari API YouTube dan 100 channel untuk data latih. Dengan demikian, sistem ini efektif dalam memberikan rekomendasi Channel masakan yang sesuai dengan preferensi pengguna. [17]

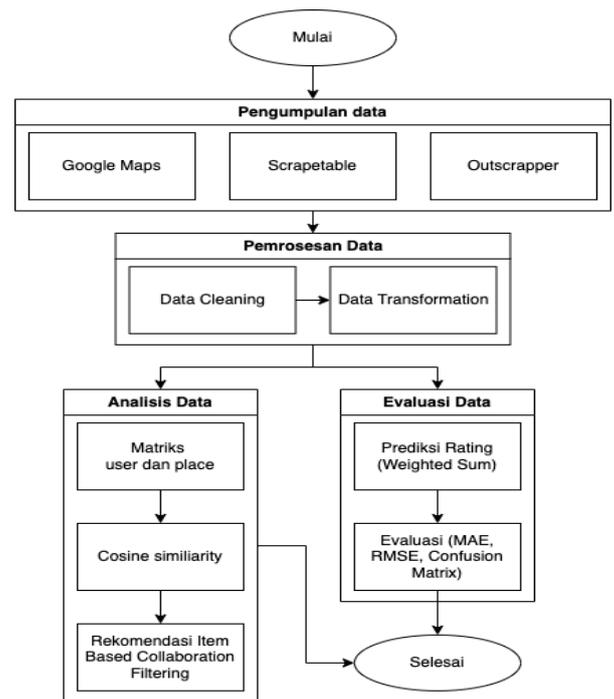
Penelitian yang dilakukan oleh Ratih Yulia Hayuningtyas & Retno Sari pada tahun 2022 dengan judul “Implementasi Data Mining Dengan Algoritma Multiple Linear Regression Untuk Memprediksi Penyakit Diabetes” Penelitian ini mengaplikasikan metode *multiple linear regression* untuk memprediksi risiko diabetes dan mengevaluasi akurasi model menggunakan *Root Mean Square Error* (RMSE). Hasil dari penelitian ini menunjukkan bahwa model regresi linear berganda berhasil memberikan prediksi dengan RMSE sebesar 0,403, yang diperoleh melalui pengujian *cross-validation* dengan variasi jumlah validasi. Hasil ini menunjukkan bahwa model yang diterapkan dapat menghasilkan prediksi yang cukup akurat dalam mendeteksi diabetes. [18]

Berdasarkan penelitian Hartatik dkk pada tahun 2021 dengan judul “Sistem Rekomendasi Wisata Kuliner di Yogyakarta dengan Metode Item-Based Collaborative Filtering” Pada penelitian ini, digunakan Item-based Collaborative Filtering untuk merekomendasikan makanan kepada pengguna, berdasarkan rating yang diberikan. Dalam implementasinya, prediksi rating makanan dihitung menggunakan weighted sum berdasarkan nilai kemiripan antar makanan. Penelitian ini menggunakan dataset yang terdiri dari 23 jenis makanan dan 22 pengguna. Setelah melalui tahapan pengumpulan data, perancangan sistem, dan implementasi, dilakukan pengujian menggunakan beberapa metrik evaluasi, termasuk Mean Absolute Error (MAE), Confusion Matrix, dan F1 Score. Hasil pengujian menunjukkan bahwa sistem ini memberikan prediksi yang cukup akurat dengan akurasi sebesar 83% menggunakan 6 tetangga terdekat (neighbor). Penerapan Collaborative Filtering dalam sistem rekomendasi ini terbukti efektif dalam memberikan rekomendasi yang relevan, serta menunjukkan pentingnya evaluasi performa sistem untuk memastikan kualitas rekomendasi yang diberikan. [19]

Penelitian sebelumnya menunjukkan bahwa Item-Based Collaborative Filtering dapat meningkatkan kualitas rekomendasi [6]. Penelitian ini bertujuan untuk mengevaluasi kinerja metode Item-Based Collaborative Filtering dalam memberikan rekomendasi destinasi wisata di Yogyakarta. Dengan menganalisis tingkat akurasi dan relevansi rekomendasi yang dihasilkan menggunakan metode evaluasi model MAE dengan dataset yang berbeda yaitu dari API Google Maps dan Scrapetable untuk daerah wisata di Yogyakarta, diharapkan penelitian ini dapat memberikan kontribusi signifikan terhadap pengembangan sistem rekomendasi yang lebih baik.

II. METODE

Adapun metode penelitian yang akan digunakan dapat dilihat pada Gambar 1. Gambar 1 menggambarkan metode penelitian yang telah diimplementasikan dalam studi ini. Berikut adalah penjelasan rinci mengenai setiap tahapan penelitian.



Gambar 1. Metode penelitian.

A. Pengumpulan Data

Pengumpulan data dilakukan melalui API Google Maps, situs Scrapetable (<https://www.scrapetable.com>) dan situs outscraper (<https://outscraper.com/>) mengambil data destinasi wisata yang berlokasi di wilayah Yogyakarta.

B. Pemrosesan Data

Data yang dikumpulkan akan dilakukan pemrosesan data dengan dua tahapan yaitu Data cleaning dan Data Transformation. Data cleaning yaitu membersihkan data seperti menghapus atau mengisi nilai yang hilang. Data Transformation yaitu menggabungkan ketiga dataset dari API Google Maps, Scrapetable, dan Outscraper. Selanjutnya melakukan pengecekan dan penyamaan format pada inkonsistensi data pemilihan fitur yang digunakan. Pada tahapan pemrosesan data hingga tahapan evaluasi menggunakan Bahasa pemrograman python dan aplikasi Google Colaboratory.

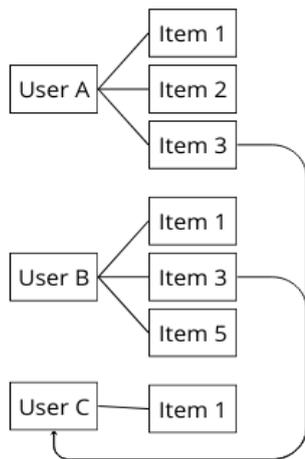
C. Analisis Data

Item-Based Collaborative Filtering adalah salah satu metode dalam sistem rekomendasi yang menggunakan kesamaan antar item untuk memberikan rekomendasi kepada pengguna. Berbeda dengan User-Based Collaborative Filtering yang fokus pada kesamaan antar pengguna, metode ini menghitung kesamaan antara item berdasarkan pola rating yang diberikan oleh pengguna. Algoritma ini pertama-tama membangun matriks kesamaan antar item menggunakan metrik cosine similarity. Ketika seorang pengguna memberikan rating pada suatu item, sistem akan merekomendasikan item lain yang serupa berdasarkan kesamaan tersebut. Keunggulan utama dari Item-Based

Collaborative Filtering adalah kemampuannya dalam mengatasi masalah sparsity dan cold start, karena item biasanya memiliki lebih sedikit variasi dibandingkan pengguna. Metode ini sangat cocok digunakan dalam aplikasi yang memiliki banyak item dengan fitur yang relevan dan cukup konsisten. Berikut ini adalah tahapan dari item collaborative filtering yang pertama yaitu melakukan proses perhitungan nilai kemiripan antar item dengan metode Cosine Similarity dengan rumus yang tertera pada persamaan 1.

$$\text{Cos}\alpha = \frac{\sum_{i=1}^n a_i \times b_i}{\sqrt{\sum_{i=1}^n a_i^2} \times \sqrt{\sum_{i=1}^n b_i^2}} \quad (1)$$

Dimana a dan b merupakan fitur yang akan dihitung nilai kemiripannya dan n merupakan jumlah keseluruhan data. Selanjutnya dari hasil kemiripan item tersebut akan dilakukan pencarian nilai yang tertinggi untuk rekomendasi berdasarkan item based collaborative filtering. Cara kerja dari item based collaborative filtering adalah sebagai contoh pada Gambar 2, user A memiliki Item 1, Item 2 dan Item 3. User B memiliki Item 1, Item 3 dan Item 5. User C akan diberikan rekomendasi Item 3 karena user A dan B sama-sama memiliki Item 1 dan Item 3. Apabila beberapa user memiliki rangkaian Item yang sama dan sesuai dengan Item yang dimiliki maka item tersebut yang akan direkomendasikan ke user lain.



Gambar 2. Ilustrasi item based collaborative filtering.

Selanjutnya adalah prediksi rating dilakukan untuk memberikan nilai rating pada rating yang kosong menggunakan metode weighted sum. Berikut adalah persamaan metode weighted sum.

$$P(a, b) = \frac{\sum_{i \in I} (R_{a,i} \times S_{b,i})}{\sum_{i \in I} |S_{b,i}|} \quad (2)$$

P merupakan prediksi dari fitur a dan b . Himpunan dari fitur yang mirip dengan disimbolkan dengan $i \in I$. R merupakan simbol dari rating dan S merupakan nilai kemiripan antara fitur a dan b .

D. Evaluasi data

Tahapan evaluasi data menggunakan data prediksi rating dari tahapan sebelumnya. Metode yang digunakan untuk

evaluasi kinerja dari model yaitu Mean Absolute Error (MAE). Berikut pada persamaan x merupakan rumus MAE.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (3)$$

Dimana (\hat{y}_i) merupakan rating yang aktual, y_i merupakan rating dari hasil prediksi dan n merupakan jumlah keseluruhan data.

Menggunakan metode evaluasi confusion matrix untuk membantu mengidentifikasi kelemahan model, sehingga dapat meningkatkan akurasi dan relevansi rekomendasi. [20] Confusion matrix adalah metode evaluasi model klasifikasi yang menggambarkan performa model berdasarkan empat parameter utama: True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN).

Evaluasi confusion matrix dihitung dengan rumus sebagai berikut :

$$\text{Accuracy} = \frac{TP+TN}{(TP+FP+FN+TN)} \quad (4)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (5)$$

$$\text{Recall} = \left(\frac{TP}{(TP+FN)} \right) \quad (6)$$

$$\text{F1 Score} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad (7)$$

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Dalam studi ini, pengumpulan data destinasi wisata di Yogyakarta dilakukan pada bulan Agustus 2024 menggunakan tiga sumber, yaitu melalui API Google Maps, situs *scrapetable.com* dan situs *outscraper.com* dengan penyaringan data untuk memastikan hanya destinasi wisata di Provinsi Daerah Istimewa Yogyakarta seperti pada Tabel 1. Data dari sumber API Google Maps untuk selanjutnya disebut dataset B menghasilkan 455 baris data dengan 14 kolom yang mencakup `id_tempat`, `nama_tempat`, `alamat`, `koordinat`, `waktu_operasional_harian`, `waktu_tutup`, `jumlah_ulasan`, `peringkat`, `waktu_paling_populer`, `waktu_operasional_populer`, `nama_pengulas`, `peringkat_ulasan`, `teks_ulasan`, dan `tanggal_ulasan`. Sumber *scrapetable.com* untuk selanjutnya disebut dataset A, diperoleh 608 baris data yang berisi sebanyak 17 atribut terkait destinasi wisata, seperti `id_bisnis`, `nomor_telepon`, `nama_tempat`, `alamat_lengkap`, `lintang`, `bujur`, `jumlah_ulasan`, `peringkat`, `zona_waktu`, `situs_web`, `id_tempat`, `tautan_tempat`, `jenis_tempat`, `jadwal_operasional`, `kota`, `status_verifikasi` dan `negara_bagian`. Sumber ketiga, dari *outscraper* memiliki 493 baris data dengan 42 kolom berikut query, `name`, `google_id`, `place_id`, `location_link`, `reviews_link`, `reviews`, `rating`, `review_id`, `user_id`, `review_pagination_id`, `author_link`, `author_title`, `author_id`, `author_image`, `author_reviews_count`, `author_ratings_count`, `review_text`, `review_img_urls`, `review_img_url`, `review_questions`, `review_photo_ids`, `owner_answer`, dan `owner_answer_timestamp`.

owner_answer_timestamp_datetime_utc, review_link, review_rating, review_timestamp, review_datetime_utc, review_likes, reviews_id, reviews_per_score_1, reviews_per_score_2, reviews_per_score_3, reviews_per_score_4, reviews_per_score_5, review_questions_Waktu kunjungan, review_questions_Waktu antrean, review_questions_Sebaliknya buat reservasi, review_questions_Area piknik, review_questions_None, reviews_per_score. Ketiga sumber data ini akan diintegrasikan pada tahapan selanjutnya untuk menghasilkan informasi yang lebih lengkap dan akurat, yang kemudian digunakan dalam analisis sistem rekomendasi destinasi wisata berbasis item-based collaborative filtering.

TABEL I
SUMBER DATA

Nama	Sumber data	Jumlah Kolom	Jumlah Baris
Dataset A	API Google Maps	14	455
Dataset B	situs scrapetable.com	17	608
Dataset C	situs outscraper.com	42	493

B. Pemrosesan Data

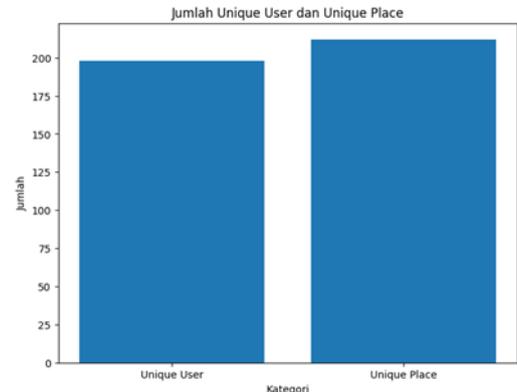
Tahapan integrasi data dilakukan dengan menggabungkan kedua sumber data berdasarkan kolom yang sama sebagai kunci yaitu *place_id* sehingga mendapatkan 1069 baris data. Proses preprocessing diterapkan untuk memastikan kualitas dan konsistensi data. Langkah awal adalah data cleaning, di mana nilai yang hilang dihapus atau diisi, seperti data *name* dan *rating* yang kosong akan dihapus. Selanjutnya dilakukan pengecekan terhadap format penulisan, panjang data, dan konsistensi atribut untuk menjaga integritas data. Beberapa penyesuaian dilakukan, seperti memberikan nilai pada *place_id* yang kosong berupa *string* pada *place_id* dengan *name* yang sama. Proses pengkodean dilakukan pada *place_id* untuk merubah dari *string* ke integer agar dapat mempermudah tahapan analisis seperti pada Tabel 2.

TABEL II
PENGKODEAN

Nama	place_id (string)	place_id (integer)
Sendang Ngajaran	ChIJ7fEnvrX_ei4R5_4hFQ03pOI	158
Taman Puspa Gading Tegaldowo Bantul Yogyakarta	ChIJcd37dqn5ei4RwEyVzTojEG0	159
Laguna View Depok	ChIJBcWI68kBey4RcOffSwZ8vyl	160
Bukit Pandawa	ChIJW89xXvb2ei4REkjDHXyvUSA	162

Pada tahap data *transformation*, dilakukan pemilihan fitur atau kolom. *Kolom place_id, name, user_id* dan *rating* merupakan fitur utama yang sesuai dengan kebutuhan analisis

dan evaluasi sistem rekomendasi berbasis *item-based collaborative filtering*. Fitur *place_id* digunakan untuk kunci dari destinasi wisata, *name* untuk menampilkan nama dari destinasi wisata, *user_id* sebagai kunci dari user dan *rating* adalah nilai rating yang diberikan oleh user. Total data yang digunakan adalah 198 untuk data *user_id* dan 212 data *place_id* seperti pada Gambar 3.



Gambar 2. Jumlah *user* dan *place*.

Tingkat sparsity atau kelonggaran data cukup tinggi dengan total user 198 dan sebanyak 218 total destinasi wisata. dimana dalam satu destinasi wisata memiliki jumlah review yang beragam dari pengguna seperti pada Tabel II.

TABEL II
PENGKODEAN

Nama	jumlah review
Agro Tainment Wisata Kampung Markisa	2
BANYU WIYOSO DAGEN (Kolam Ikan)	4
Bendhung Lepen	1366
Bendungan Irigasi Jalan Samas	1
Bukit Kleresede	5
Candi Donotirto	284
Cepuri Parangkusumo	176
Curug Banyunibo	306

Tahap penghitungan kemiripan antar tempat dilakukan dengan menggunakan *cosine similarity*, yang memungkinkan model untuk mengidentifikasi tempat wisata yang paling mirip berdasarkan rating dari pengguna. Setiap *user_id* dianalisis untuk menghitung kemiripan preferensi menggunakan perhitungan cosine similarity pada data rating, yang memungkinkan identifikasi hubungan antar pengguna dengan preferensi yang serupa. Pada contoh perhitungan manual, nilai cosine similarity antara dua tempat wisata—misalnya, tempat wisata A dan B—adalah sebesar 0.98, yang menunjukkan kemiripan yang sangat tinggi. Nilai ini dihasilkan dengan menghitung dot product antara vektor rating dari kedua tempat wisata dan membaginya dengan hasil kali panjang vektor masing-masing. Nilai cosine similarity yang tinggi seperti ini menunjukkan bahwa tempat wisata A dan B memiliki pola rating yang mirip, sehingga B akan

direkomendasikan kepada pengguna yang telah memberi rating tinggi untuk A.

Berikut adalah langkah perhitungan cosine similiarity

1) *Matriks user_id dan place_id*: Langkah pertama dalam metode ini adalah membentuk matriks *user-id* dan *place-id*, di mana setiap baris mewakili destinasi wisata (*place_id*) dan setiap kolom mewakili pengguna (*user_id*) yang memberikan rating terhadap destinasi tersebut. Misalnya, data sederhana yang digunakan adalah *Place_id* 0 dengan sssuser 0,1,2,3 melakukan *rating* dengan nilai 4.0, 3.0, 5.0 dan 0.0, *Place_id* 1 dengan *user* yang sama mendapatkan *rating* 3.0, 4.0, 2.0 dan 0.0 hasil dari matrik digambarkan pada Tabel 3.

TABEL III
Matrik *USER_ID* DAN *PLACE_ID*

<i>user_id</i>	Rating <i>Place_id</i>	
	0	1
0	4.0	3.0
1	3.0	4.0
2	5.0	2.0
3	0.0	0.0

2) *Perhitungan Cosine Similarity*: Langkah selanjutnya adalah menghitung kemiripan (*similarity*) antar destinasi wisata menggunakan rumus cosine *similarity*. Berikut hitung cosine *similarity* antara destinasi wisata 0 dan 1 dengan vektor rating sebagai berikut:

Place_id 0: [4, 5, 3],

Place_id 1: [3, 4, 2]

Perhitungan untuk dot product A.B

$$A \cdot B = (4 \times 3) + (5 \times 4) + (3 \times 2) = 12 + 20 + 6 = 38$$

Perhitungan vector untuk *Place_id* 0

$$\|A\| = \sqrt{4^2 + 5^2 + 3^2} = \sqrt{50} = 7.07$$

Perhitungan vector untuk *Place_id* 1

$$\|B\| = \sqrt{3^2 + 4^2 + 2^2} = \sqrt{29} = 5.39$$

Perhitungan cosine similiarity antara *Place_id* 0 dan *Place_id* 1

$$cosine_{similarity(0,1)} = \frac{A \cdot B}{\|A\| \|B\|} = \frac{38}{38.11} = 0.997$$

Nilai cosine *similarity* sebesar 0.997 menunjukkan bahwa kedua destinasi memiliki kesamaan yang sangat tinggi. Hasil dari *similarity* ini kemudian digunakan untuk menyusun rekomendasi destinasi wisata, dengan mempertimbangkan preferensi dan pola perilaku pengguna. Menggunakan metode item collaborative filtering dengan menggunakan inputan *place_id* dibangun model sistem rekomendasi. Sebagai contoh, *place_id* 185 dengan nama History of Java Museum mendapatkan 10 rekomendasi tempat favorit yaitu Museum Wayang Beber Sekartaji (Sanggar Bhuana Alit) (*place_id*:

171) dengan skor *similarity* 0.45, Jogja National Museum (*place_id*: 162) dengan skor *similarity* 0.32, Sendang Ngajaran (*place_id*: 158) dengan skor *similarity* 0.32, Kampung Flory Jogja (*place_id*: 176) dengan skor *similarity* 0.26, Laguna View Depok (*place_id*: 163) dengan skor *similarity* 0.18, JOGLO JOSS (*place_id*: 207) dengan skor *similarity* 0.18, Museum Radio AURI PC 2 (*place_id*: 194) dengan skor *similarity* 0.18, dan Tugu Pal Gamping (*place_id*: 208) dengan skor *similarity* 0.17.

Langkah berikutnya adalah memprediksi rating yang akan diberikan oleh pengguna untuk destinasi tertentu. Prediksi ini didasarkan pada rating tempat wisata serupa, dengan bobot yang diberikan sesuai dengan tingkat kemiripan (*similarity*) antara tempat wisata tersebut. Prediksi rating dihitung menggunakan metode weighted sum, di mana bobot dari pengguna dengan preferensi serupa diakumulasikan untuk memprediksi rating destinasi yang belum dinilai. Berikut adalah contoh perhitungan prediksi rating *user_id* 1 untuk *place_id* 2. Menggunakan dua *place_id* teratas (0 dan 1) dengan nilai *similarity* 0.8 untuk *place_id* 0 dan 0.9 untuk *place_id* 1.

$$prediksi\ rating = \frac{0.8 \times 4 + 0.9 \times 3}{0.8 + 0.9} = 3.47$$

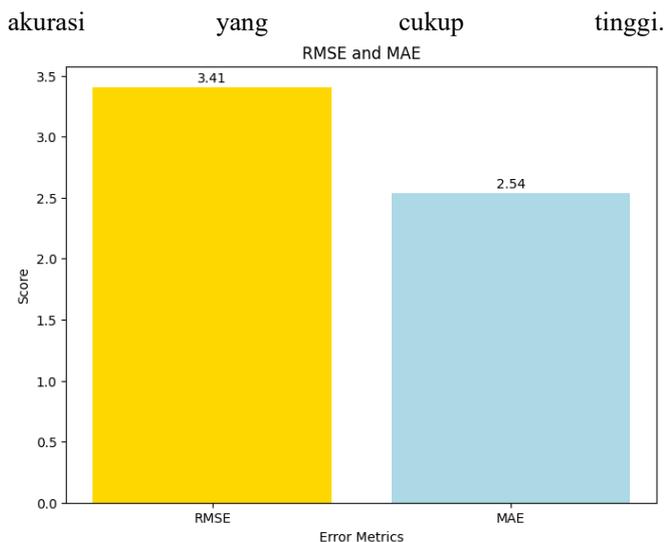
Dari perhitungan tersebut didapatkan nilai 3.47 untuk hasil prediksi rating *place_id* 2 oleh *user_id* 1.

Hasil prediksi rating berdasarkan kode program menggunakan bahasa python dijelaskan pada Tabel 4.

TABEL IV
PREDIKSI RATING

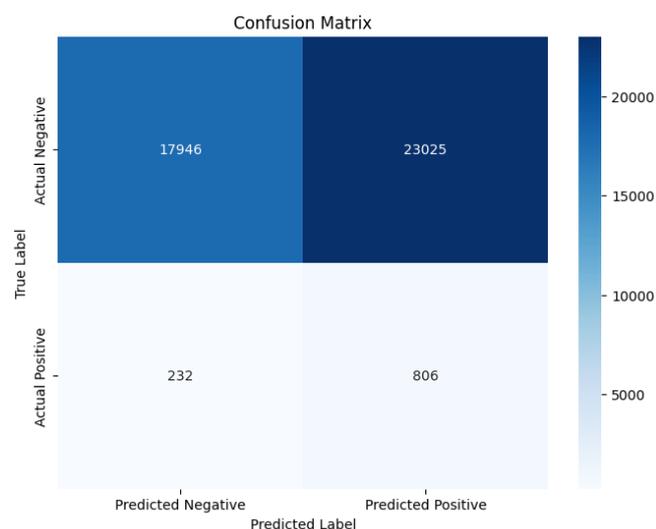
<i>User_id</i>	<i>Place_id</i>	Rating	Prediksi rating
0	0	0.0	4.4
0	1	0.0	5.0
0	2	0.0	4.3
0	3	0.0	5.0
.	.	.	.
94	71	0.0	0.0

Tahap terakhir dalam proses ini adalah mengevaluasi kinerja model. Metode MAE dan RMSE dihitung dengan membandingkan rating prediksi dengan rating aktual yang diberikan oleh pengguna. Berdasarkan hasil evaluasi model rekomendasi yang menggunakan metode Item-Based Collaborative Filtering, nilai MAE yang diperoleh adalah 2.54 dan RMSE sebesar 3.41 pada Gambar 3. Nilai MAE dan RMSE ini menunjukkan tingkat kesalahan rata-rata antara rating prediksi dan rating aktual pada dataset yang digunakan. Nilai MAE dan RMSE yang mendekati nol mengindikasikan bahwa prediksi rating yang dihasilkan oleh model memiliki



Gambar 3. Hasil evaluasi RMSE dan MAE

Hasil evaluasi pada Gambar 4 menggunakan Confusion Matrix menunjukkan bahwa model memiliki True Negative (TN) sebanyak 17.946, yaitu prediksi rating negatif yang benar-benar sesuai dengan rating sebenarnya. Namun, model juga menghasilkan False Positive (FP) yang sangat tinggi, sebanyak 23.025, yang berarti banyak rating negatif salah diprediksi sebagai positif. Hal ini menunjukkan bahwa model sering salah dalam mengklasifikasikan rating negatif. Sebaliknya, jumlah False Negative (FN) relatif kecil, yaitu hanya 232, yang menunjukkan bahwa hanya sedikit rating positif yang salah diprediksi sebagai negatif, sesuai dengan nilai recall yang tinggi. Sementara itu, model berhasil mengidentifikasi True Positive (TP) sebanyak 806, yaitu prediksi rating positif yang akurat. Evaluasi ini menunjukkan bahwa meskipun model cukup baik dalam mengenali rating positif, performanya dalam membedakan rating negatif masih memerlukan perbaikan. Model mengalami ketidakseimbangan antara precision dan recall, di mana recall yang tinggi menunjukkan kemampuan model dalam mengenali banyak data positif, namun precision yang rendah mengindikasikan bahwa model sering salah memprediksi data negatif sebagai positif. Hal ini kemungkinan disebabkan oleh dataset yang tidak seimbang, sehingga model lebih memprioritaskan recall tanpa memperhatikan akurasi prediksi.



Gambar 4. Hasil evaluasi Confusion Matrix

IV. KESIMPULAN

Kesimpulan dari penelitian ini menegaskan bahwa penerapan item-based collaborative filtering pada sistem rekomendasi destinasi wisata di Yogyakarta efektif dalam menghasilkan rekomendasi yang sesuai dengan preferensi pengguna. Dengan menghitung kemiripan antar destinasi menggunakan cosine similarity, sistem ini mampu mengidentifikasi kesamaan pola penilaian dan memberikan rekomendasi destinasi berdasarkan preferensi pengguna yang serupa. Prediksi rating untuk destinasi yang belum diulas dilakukan melalui metode weighted sum, yang menghasilkan akurasi prediksi dengan nilai Mean Absolute Error (MAE) sebesar 2.54 dan RMSE sebesar 3.41. Nilai MAE dan RMSE ini mengindikasikan bahwa rata-rata kesalahan prediksi relatif rendah, yang menunjukkan tingkat akurasi yang memadai dalam rekomendasi destinasi. Penelitian selanjutnya dapat menyeimbangkan data menggunakan teknik oversampling atau undersampling. Selain itu, mencoba algoritma yang lebih cocok terhadap data yang tidak seimbang, seperti Random Forest atau XGBoost, dapat memberikan hasil yang lebih baik.

UCAPAN TERIMA KASIH

Terima kasih kepada Direktorat Akademik Pendidikan Tinggi Vokasi, Direktorat Jenderal Pendidikan Vokasi yang telah memberikan pendanaan penelitian ini melalui program Hibah Penelitian dan Pengabdian kepada Masyarakat (PPM) Perguruan Tinggi Vokasi Tahun 2024 dengan nomor kontrak 419/SPK/D.D4/PPK.01.APTV/VIII/2024 tertanggal 26 Agustus 2024.

DAFTAR PUSTAKA

- [1] Badan Pusat Statistik, "Statistik Pariwisata Indonesia 2022," Badan Pusat Statistik.
- [2] Alsa Yuda Putri, Usep Suhud, and Rahmi, "Peran Destination Attributes dalam Meningkatkan Revisit Intention: Kasus pada Turis di Keraton Ngayogyakarta Hadiningrat," *JEMSI (Jurnal Ekonomi,*

- Manajemen, dan Akuntansi*, vol. 9, no. 4, pp. 1668–1678, Aug. 2023, doi: 10.35870/jemsi.v9i4.1414.
- [3] F. T. Abdul Hussien, A. M. S. Rahma, and H. B. Abdul Wahab, “Recommendation Systems For E-commerce Systems An Overview,” *J Phys Conf Ser*, vol. 1897, no. 1, p. 012024, May 2021, doi: 10.1088/1742-6596/1897/1/012024.
- [4] H. Ko, S. Lee, Y. Park, and A. Choi, “A Survey of Recommendation Systems: Recommendation Models, Techniques, and Application Fields,” *Electronics (Basel)*, vol. 11, no. 1, p. 141, Jan. 2022, doi: 10.3390/electronics11010141.
- [5] 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, Oct. 2022, doi: 10.1016/j.jksuci.2021.09.014.
- [6] M. M. Dewi, “Optimasi Pearson Correlation untuk Sistem Rekomendasi menggunakan Algoritma Firefly,” *JURNAL INFORMATIKA*, vol. 9, no. 1, pp. 1–5, Apr. 2022.
- [7] D. Nugraha, T. W. Purboyo, and R. A. Nugrahaeni, “Sistem Rekomendasi Film Menggunakan Metode User Based Collaborative Filtering (Movie Recommendation System Using User Based Collaborative Filtering Method).”
- [8] L. Gang, “Personalized Recommendation of Tourist Attractions Based on Collaborative Filtering,” *IEEE Xplore*, 2020.
- [9] A. H. Ardiansyah, “Implementation of Item-Based Collaborative Filtering on Web-Based Culinary Tourism System,” *Journal Unimma*, 2022.
- [10] C. Ajaegbu, “An optimized item-based collaborative filtering algorithm,” *J Ambient Intell Humaniz Comput*, vol. 12, no. 12, pp. 10629–10636, Dec. 2021, doi: 10.1007/s12652-020-02876-1.
- [11] A. M. Ninan and J. E. Rajan, “An Item-Based Collaborative Filtering on Recommendation of Travel Route,” *Academia*, 2019.
- [12] Rayhan Rizal Mahendra, Fetty Tri Anggraeny, and Henni Endah Wahanani, “Implementasi Item-Based Collaborative Filtering Untuk Rekomendasi Film,” *Repeater: Publikasi Teknik Informatika dan Jaringan*, vol. 2, no. 3, pp. 213–221, Jul. 2024, doi: 10.62951/repeater.v2i3.140.
- [13] Ibrahim Asad and Muhammad Zakariyah, “Aplikasi Rekomendasi Pemesanan Paket Wisata Menggunakan Metode Collaborative Filtering,” *METIK JURNAL*, vol. 7, no. 2, pp. 76–84, Dec. 2023, doi: 10.47002/metik.v7i2.639.
- [14] R. Marappan, “Movie Recommendation System using an Item-based Collaborative Filtering,” *International Journal of Mathematical, Engineering, Biological and Applied Computing*, vol. 1, no. 1, pp. 42–43, Jun. 2022, doi: 10.31586/ijmebac.2022.340.
- [15] A. Agustian, S. P. A. Alkadri, and I. Istikomah, “Penerapan Metode Collaborative Filtering untuk Rekomendasi Tempat Kos di Sekitar Kampus UM Pontianak,” *Jurnal Informatika Polinema*, vol. 10, no. 3, pp. 333–340, May 2024, doi: 10.33795/jip.v10i3.5085.
- [16] S. Rosyad, D. Mahendra, and N. Azizah, “Sistem Rekomendasi Buku Di Perpustakaan Daerah Jepara Menggunakan Metode Item-Based Collaborative Filtering,” *Biner: Jurnal Ilmiah Informatika dan Komputer*, vol. 2, no. 2, pp. 76–81, Jul. 2023, doi: 10.32699/biner.v2i2.3934.
- [17] F. Kurniawan, A. Kania Ningsih, and A. Komarudin, “Sistem Rekomendasi Channel Youtube Resep Masakan Menggunakan Collaborative Filtering,” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 4, pp. 5849–5855, Jun. 2024, doi: 10.36040/jati.v8i4.10107.
- [18] R. Y. Hayuningtyas and R. Sari, “Implementasi Data Mining Dengan Algoritma Multiple Linear Regression Untuk Memprediksi Penyakit Diabetes,” doi: 10.31294/jtk.v4i2.
- [19] H. Hartatik, S. D. Nurhayati, and W. Widayani, “Sistem Rekomendasi Wisata Kuliner di Yogyakarta dengan Metode Item-Based Collaborative Filtering,” *Journal Automation Computer Information System*, vol. 1, no. 2, pp. 55–63, Nov. 2021, doi: 10.47134/jacis.v1i2.8.
- [20] B. Zuraeni and Fitrianiingsih, “EKOSPHERE: Jurnal Ekonomi Pembangunan dan Manajemen Volume Analisis Ramalan Cuaca di Sekupang, Kota Batam Menggunakan Algoritma Decision Tree dan Confusion Matrix.” [Online]. Available: <https://ibnusunapublisher.org/index.php/EKOSPHERE>