

Perbandingan Model Rekomendasi Wisata Yogyakarta Berbasis Rating dan Segmentasi Usia

Comparison of Tourism Recommender Models in Yogyakarta Using Ratings and Age Segmentation

¹Fiyas Mahananing Puri*

¹Program Studi Sistem Informasi, Universitas Amikom Yogyakarta, Indonesia
fiyas@amikom.ac.id

Received:
16 Desember 2026

Revised:
11 Januari 2026

Accepted:
12 Januari 2026

Published:
6 Februari 2026

ABSTRAK

Ledakan informasi pada platform digital pariwisata membuat wisatawan kesulitan memilih destinasi yang paling relevan dengan preferensi mereka, sehingga sistem rekomendasi menjadi semakin penting untuk menyaring pilihan berdasarkan data rating pengguna. Penelitian ini bertujuan menganalisis dan membandingkan kinerja tiga model rekomendasi destinasi wisata di Yogyakarta, yaitu *baseline* global berbasis rata-rata rating, item-based *collaborative filtering*, dan *demographic filtering* berbasis segmentasi usia. Penelitian menggunakan pendekatan kuantitatif dengan studi kasus pada dataset pariwisata Yogyakarta yang berisi 2.870 interaksi antara 300 pengguna dan 126 destinasi. Data dipra-proses melalui integrasi tiga berkas, pembersihan data, pembentukan kelompok usia, serta pembagian data menjadi 80 persen latih dan 20 persen uji. Kinerja model dievaluasi menggunakan *Mean Absolute Error* (MAE), baik secara global maupun per kelompok usia. Hasil menunjukkan bahwa *baseline* global memberikan MAE terendah secara keseluruhan, yaitu 1,222, sedangkan *demographic filtering* dan *collaborative filtering* memperoleh MAE masing-masing 1,283 dan 1,301. Pada seluruh kelompok usia, *baseline* tetap menjadi model dengan MAE terkecil, sehingga pendekatan sederhana berbasis rata-rata rating terbukti lebih stabil pada dataset yang relatif kecil dan jarang. Meskipun demikian, *demographic filtering* tetap bermanfaat untuk memetakan preferensi destinasi per segmen usia. Penelitian selanjutnya disarankan menggunakan dataset yang lebih besar, menambahkan atribut kontekstual, menguji model hibrida, serta melengkapi evaluasi dengan metrik berbasis ranking dan studi kepuasan pengguna.

Kata Kunci : sistem rekomendasi, pariwisata Yogyakarta, *collaborative filtering*, *demographic filtering*, MAE.

ABSTRACT

The rapid growth of digital tourism platforms has made it difficult for travelers to identify destinations that truly match their preferences, which increases the importance of recommender systems to filter options based on user rating data. This study aims to analyze and compare the performance of three recommendation models for tourist destinations in Yogyakarta, namely a global baseline based on average rating, item-based collaborative filtering, and demographic filtering based on age segmentation. A quantitative approach is applied to the Yogyakarta tourism dataset containing 2,870 interactions between 300 users and 126 destinations. The data are preprocessed through integration of three files, data cleaning, age group formation, and an 80 percent training and 20 percent test split. Model performance is evaluated using Mean Absolute Error (MAE), both globally and for each age group. The results show that the global baseline achieves the lowest overall MAE of 1.222, while demographic filtering and collaborative filtering obtain MAE values of 1.283 and 1.301 respectively. For all age groups, the baseline remains the best model, indicating that a simple average-rating approach is more stable on a relatively small and sparse dataset. Nevertheless, demographic filtering is still valuable for mapping destination preferences by age segment. Future work is recommended to use larger datasets, incorporate additional contextual attributes, explore hybrid models, and complement evaluation with ranking-based metrics and user satisfaction studies.

Keywords : recommender system, Yogyakarta tourism, *collaborative filtering*, *demographic filtering*, MAE.

PENDAHULUAN

Perkembangan teknologi informasi dan pesatnya penetrasi internet telah mengubah cara wisatawan merencanakan perjalanan mereka. Calon wisatawan kini tidak lagi bergantung pada brosur fisik atau agen perjalanan konvensional, melainkan mengakses ratusan ulasan, foto, dan rating melalui berbagai platform digital. Meskipun informasi yang melimpah ini memperkaya pilihan, kenyataannya justru menimbulkan beban kognitif karena pengguna harus menyaring sendiri destinasi yang paling relevan dengan preferensi dan keterbatasan mereka. Sistem rekomendasi hadir sebagai solusi penting untuk membantu pengguna menavigasi informasi tersebut dengan cara yang lebih personal dan efisien, melalui pemanfaatan data historis perilaku maupun profil pengguna (Solano-Barliza et al., 2024) (Fahrizal et al., 2024).

Dalam beberapa tahun terakhir, penelitian mengenai sistem rekomendasi pada sektor pariwisata menunjukkan kemajuan yang signifikan. Kajian sistematis terkini menunjukkan bahwa sistem rekomendasi pariwisata tidak hanya meningkatkan akurasi dan personalisasi, tetapi juga memperkaya pengalaman wisatawan serta mendukung pengelolaan destinasi dan promosi layanan wisata (Arregoces-Julio et al., 2025). Beragam pendekatan telah diusulkan, mulai dari content-based filtering, *collaborative filtering*, hingga hybrid models yang memanfaatkan deep learning dan teknik optimasi lainnya (Fahrizal et al., 2024) (Fahrizal et al., 2024). Pada destinasi wisata dengan ketersediaan data yang masih terbatas, pendekatan berbasis profil pengguna dan data demografis menjadi alternatif yang menjanjikan karena tidak memerlukan riwayat interaksi yang panjang (Arregoces-Julio et al., 2025) (Generosi et al., 2025).

Daerah Istimewa Yogyakarta merupakan salah satu destinasi wisata unggulan di Indonesia, yang menawarkan beragam atraksi mulai dari wisata budaya, sejarah, kuliner, hingga wisata alam. Variasi ini membuat Yogyakarta sangat menarik bagi wisatawan, namun juga menghadirkan tantangan tersendiri dalam memilih destinasi yang sesuai dengan preferensi pribadi, terutama bagi wisatawan yang baru pertama kali berkunjung. Beberapa studi lokal telah mengembangkan sistem rekomendasi untuk wisata Yogyakarta, umumnya menggunakan pendekatan item-based *collaborative filtering* (Puri et al., 2024), dan menunjukkan relevansi rekomendasi berdasarkan kemiripan pola rating antar objek wisata (Hartatik et al., 2021) (Dewi et al., 2025). Namun demikian, fokus penelitian tersebut lebih menekankan pada akurasi global model tanpa memperhatikan variasi kinerja berdasarkan karakteristik pengguna seperti kelompok usia.

Collaborative filtering (CF) merupakan salah satu pendekatan yang paling banyak digunakan dalam sistem rekomendasi karena mengandalkan pola kesamaan antar pengguna atau antar item untuk memprediksi preferensi pengguna (Generosi et al., 2025) (Solano-Barliza et al., 2024). Namun, literatur terbaru menekankan pentingnya integrasi data demografis dalam desain sistem rekomendasi, terutama dalam kondisi data interaksi yang terbatas atau ketika rekomendasi ditujukan untuk kelompok pengguna tertentu. Studi profile-based *recommender* menunjukkan bahwa pengelompokan pengguna berdasarkan atribut sosio-demografis dan karakteristik perjalanan mampu menghasilkan rekomendasi yang cukup akurat meskipun data rating relatif sedikit (Arregoces-Julio et al., 2025).

Demographic filtering memanfaatkan atribut seperti usia, jenis kelamin, dan lokasi tinggal pengguna untuk membentuk segmen yang homogen, dan mengolah statistik preferensi per segmen untuk membangun sistem rekomendasi. Penelitian terbaru dalam bidang layanan digital dan perjalanan menunjukkan bahwa integrasi informasi demografis dapat meningkatkan relevansi rekomendasi, terutama dalam domain yang dipengaruhi oleh faktor usia dan konteks sosial pengguna (Huda, 2023) (Generosi et al., 2025). Dalam konteks pariwisata, pendekatan ini dimanfaatkan untuk menyusun paket perjalanan atau aktivitas wisata yang sesuai dengan preferensi segmen tertentu, seperti wisata keluarga, wisatawan muda, atau wisatawan senior (Arregoces-Julio et al., 2025).

Di sisi lain, banyak studi sistem rekomendasi pariwisata yang cenderung langsung mengusulkan model yang kompleks seperti kombinasi hybrid CF–CBF atau model berbasis jaringan saraf tanpa secara eksplisit membandingkan dengan *baseline* sederhana seperti rata-rata rating global atau model MostPop. Padahal, *baseline* sederhana sering kali memberikan kinerja yang kompetitif dan menjadi tolok ukur penting untuk menilai nilai tambah dari model yang lebih kompleks (Klimashevskaja et al., 2024). Tanpa adanya perbandingan yang jelas terhadap *baseline*, sulit untuk menilai efektivitas praktis model yang diusulkan.

Selain itu, masih sedikit penelitian yang secara eksplisit membahas kinerja model rekomendasi berdasarkan kelompok demografis tertentu. Literatur tentang fairness dan bias dalam sistem rekomendasi menunjukkan bahwa kualitas rekomendasi dapat bervariasi signifikan antar kelompok usia, dan bahwa metrik agregat saja tidak cukup untuk menangkap variasi ini (Zhao et al., 2025) (Chizari et al., 2024). Namun demikian, sebagian besar penelitian di bidang pariwisata, terutama dalam konteks lokal Indonesia, masih melaporkan kinerja sistem pada level agregat tanpa mengevaluasi performa antar segmen pengguna secara terpisah (Hartatik et al., 2021) (Sispianyala, 2024).

Berdasarkan tinjauan tersebut, penelitian ini bertujuan untuk menganalisis dan membandingkan kinerja tiga model rekomendasi destinasi wisata di Yogyakarta, yaitu *baseline* global berbasis rata-rata rating, item-based *collaborative filtering*, dan *demographic filtering* berbasis usia pengguna. Penelitian ini secara

khusus mengevaluasi apakah segmentasi usia dapat menurunkan kesalahan prediksi rating dibandingkan *baseline* global, serta sejauh mana pendekatan CF dan DF memberikan manfaat praktis dalam konteks dataset lokal yang terbatas. Evaluasi dilakukan dengan menggunakan metrik *Mean Absolute Error* (MAE), baik secara global maupun per kelompok usia, guna memperoleh pemahaman yang lebih mendalam terhadap performa dan relevansi masing-masing pendekatan.

METODE

1. Desain Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan studi kasus pada sistem rekomendasi tempat wisata di Yogyakarta. Fokus penelitian adalah membandingkan kinerja tiga pendekatan rekomendasi, yaitu *baseline* global top-N, *collaborative filtering*, dan *demographic filtering* berbasis segmentasi usia, dalam memprediksi rating pengguna terhadap destinasi wisata (Casillo et al., 2023). Secara umum, alur penelitian terdiri dari tahapan: pengumpulan dan integrasi dataset, pra-pemrosesan data, pembangunan model rekomendasi, serta evaluasi kinerja menggunakan metrik *Mean Absolute Error* (MAE) secara global dan per kelompok usia pengguna (Rajesh & Kumar, 2025) (Bauer et al., 2024).

2. Dataset

Dataset yang digunakan dalam penelitian ini berasal dari tiga berkas utama <https://www.kaggle.com/datasets/taufan/dataset-pariwisata-yogyakarta> yaitu *tour.csv*, *tour_rating.csv*, dan *user.csv*. Berkas *tour.csv* memuat informasi dasar tempat wisata, seperti ID destinasi (*tour_id*) dan nama tempat wisata (*tour_name*).

- a. Berkas *tour_rating.csv* memuat data rating yang diberikan pengguna terhadap tempat wisata, dengan atribut utama ID pengguna (*user_id*), ID destinasi (*tour_id*), dan nilai rating (*rating*).
- b. Berkas *user.csv* memuat informasi pengguna, terutama ID pengguna (*user_id*) dan usia (*age*).

Ketiga berkas tersebut digabungkan (*join*) berdasarkan pasangan kunci (*tour_id*) dan (*user_id*) sehingga menghasilkan satu tabel terintegrasi yang memuat minimal atribut: *user_id*, *age*, *tour_id*, *tour_name*, dan *rating*.

Setelah proses integrasi, diperoleh sebanyak [*N_total*] baris data interaksi pengguna–destinasi yang siap digunakan untuk tahap pra-pemrosesan dan pemodelan (Carnovalini et al., 2025).

3. Pra-Pemrosesan Data

Tahap pra-pemrosesan data dilakukan untuk memastikan data yang digunakan bersih dan konsisten (Carnovalini et al., 2025) (Deldjoo et al., 2024). Langkah-langkah yang dilakukan adalah sebagai berikut:

1. Seleksi wilayah Yogyakarta. Jika dataset mengandung destinasi dari beberapa wilayah, maka pada tahap ini dilakukan penyaringan sehingga hanya destinasi yang berada di wilayah Yogyakarta yang disertakan dalam penelitian. Penyaringan dilakukan berdasarkan atribut [misal: *region/city*] pada berkas *tour.csv*.
2. Penanganan nilai hilang (*missing values*). Data dengan nilai hilang pada atribut penting, seperti *age* atau *rating*, dihapus dari dataset karena dapat memengaruhi proses pembelajaran model dan evaluasi kinerja. Setelah proses ini, jumlah data yang digunakan dalam penelitian menjadi [*N_bersih*] baris.
3. Segmentasi usia pengguna. Usia pengguna pada atribut *age* kemudian dikelompokkan ke dalam empat segmen usia yang didefinisikan sebagai berikut: <18 tahun, 18–25 tahun, 26–35 tahun, >35 tahun
4. Hasil segmentasi disimpan dalam atribut baru *age_group*. Segmentasi ini digunakan pada tahap pembangunan model *demographic filtering* guna menghasilkan rekomendasi yang spesifik untuk tiap kelompok usia.
5. Pembentukan dataset akhir. Setelah seluruh tahap pra-pemrosesan dilakukan, diperoleh dataset akhir dengan struktur minimal: *user_id*, *age*, *age_group*, *tour_id*, *tour_name*, dan *rating*. Dataset inilah yang selanjutnya digunakan pada proses pelatihan dan pengujian model rekomendasi.

4. Pembagian Data Latih dan Uji

Untuk mengevaluasi kinerja model secara obyektif, dataset akhir dibagi menjadi dua subset, yaitu data latih (*train set*) dan data uji (*test set*) (Carnovalini et al., 2025). Dataset dibagi menjadi **data latih** dan **data uji** dengan proporsi masing-masing sebesar **80%** dan **20%** menggunakan teknik **pembagian acak (*random split*)**

Data latih digunakan untuk membangun ketiga model rekomendasi (*baseline*, *collaborative filtering*, dan *demographic filtering*), sedangkan data uji digunakan hanya untuk menghitung kinerja model tanpa dilibatkan dalam proses pelatihan. Pendekatan ini bertujuan untuk mengurangi bias dan memberikan estimasi kinerja model yang lebih representatif terhadap data baru.

5. Model Rekomendasi

Penelitian ini membandingkan tiga model rekomendasi, yaitu *baseline* global top-N, *collaborative filtering* berbasis item, dan *demographic filtering* berbasis usia. Ketiga model dibangun menggunakan data latih yang sama sehingga perbandingan kinerja dapat dilakukan secara adil.

1. *Baseline* Global Top-N

Model *baseline* digunakan sebagai pembanding dasar yang tidak mempertimbangkan karakteristik individu pengguna. Pada model ini, rata-rata rating setiap destinasi wisata dihitung berdasarkan seluruh data latih. Nilai rata-rata rating destinasi i dinyatakan sebagai:

$$\bar{r}_i = \frac{1}{N_i} \sum_{u \in U_i} r_{u,i} \quad (1)$$

\bar{r}_i adalah rata-rata rating destinasi i ,

U_i adalah himpunan pengguna yang memberikan rating terhadap destinasi i ,

N_i adalah jumlah pengguna dalam U_i ,

$r_{u,i}$ adalah rating yang diberikan pengguna u terhadap destinasi i .

Prediksi rating pada data uji untuk destinasi i dilakukan dengan menggunakan nilai \bar{r}_i tersebut, terlepas dari identitas pengguna. Rekomendasi top-N destinasi diperoleh dengan mengurutkan destinasi berdasarkan nilai rata-rata rating tertinggi.

2. *Collaborative filtering* Berbasis Item

Model *collaborative filtering* (CF) yang digunakan dalam penelitian ini adalah item-based CF. Intuisi utama metode ini adalah bahwa destinasi yang dinilai serupa oleh banyak pengguna akan cenderung direkomendasikan bersama (Fkih, 2022). Langkah-langkah utama item-based CF adalah sebagai berikut:

a. Pembentukan Matriks User-item

Dari data latih dibentuk matriks R berukuran $|U| \times |I|$, dengan $|U|$ adalah jumlah pengguna dan $|I|$ adalah jumlah destinasi wisata. Setiap elemen $r_{u,i}$ merepresentasikan rating pengguna u terhadap destinasi i .

b. Perhitungan Kemiripan Antar Item

Kemiripan antar destinasi wisata dihitung menggunakan ukuran kemiripan, misalnya cosine similarity. Kemiripan antara item i dan j dinyatakan sebagai:

$$\text{sim}(i, j) = \frac{\sum_{u \in U_{i,j}} r_{u,i} \cdot r_{u,j}}{\sqrt{\sum_{u \in U_{i,j}} r_{u,i}^2} \cdot \sqrt{\sum_{u \in U_{i,j}} r_{u,j}^2}} \quad (2)$$

dengan $U_{i,j}$ adalah himpunan pengguna yang memberikan rating terhadap kedua item i dan j .

c. Prediksi Item

Prediksi rating pengguna u terhadap destinasi i dihitung sebagai rata-rata berbobot rating pengguna tersebut terhadap item-item lain yang mirip dengan i :

$$\hat{r}_{u,i} = \frac{\sum_{j \in N_i(u)} \text{sim}(i, j) \cdot r_{u,j}}{\sum_{j \in N_i(u)} |\text{sim}(i, j)|} \quad (3)$$

dengan $N_i(u)$ adalah himpunan item yang mirip dengan i dan telah dirating oleh pengguna u .

3. *Demographic filtering* Berbasis Usia

Model utama dalam penelitian ini adalah *demographic filtering* yang memanfaatkan segmentasi usia pengguna (Deldjoo et al., 2024). Pada pendekatan ini, prediksi rating untuk suatu destinasi bergantung pada kelompok usia pengguna.

a. Pengelompokan berdasarkan usia

Data latih dikelompokkan berdasarkan kombinasi (age_group, tour_id). Untuk setiap kelompok usia g dan destinasi i , dihitung rata-rata rating:

$$\bar{r}_{g,i} = \frac{1}{N_{g,i}} \sum_{u \in U_{g,i}} r_{u,i} \quad (4)$$

$\bar{r}_{g,i}$ adalah rata-rata rating destinasi i pada kelompok usia g ,

$U_{g,i}$ adalah himpunan pengguna pada kelompok usia g yang memberikan rating pada destinasi i ,

$N_{g,i}$ adalah jumlah pengguna dalam $U_{g,i}$.

b. Prediksi Rating

Prediksi rating untuk pengguna pada kelompok usia g terhadap destinasi i menggunakan nilai $\bar{r}_{g,i}$. Jika kombinasi (g, i) belum pernah muncul di data latih (cold-start lokal), maka dapat digunakan nilai fallback, misalnya rata-rata rating global destinasi i .

c. Rekomendasi top-N per kelompok usia

Untuk pengguna baru dalam kelompok usia g , daftar rekomendasi top-N disusun dengan mengurutkan destinasi berdasarkan nilai $\bar{r}_{g,i}$ tertinggi pada kelompok usia tersebut.

6. Evaluasi Kinerja

Evaluasi kinerja dilakukan menggunakan metrik *Mean Absolute Error* (MAE) pada data uji. MAE mengukur rata-rata selisih absolut antara rating aktual dan rating prediksi (Rajesh & Kumar, 2025), yang dirumuskan sebagai:

$$MAE = \frac{1}{N} \sum_{k=1}^N |r_k - \hat{r}_k| \quad (5)$$

N adalah jumlah pasangan rating aktual dan prediksi pada data uji,

r_k adalah rating aktual ke- k ,

\hat{r}_k adalah rating prediksi ke- k .

HASIL DAN PEMBAHASAN

A. Hasil

1. Statistik Deskriptif Dataset

Dataset yang digunakan dalam penelitian ini merupakan DATASET PARIWISATA YOGYAKARTA yang terdiri dari tiga berkas, yaitu `tour.csv`, `tour_rating.csv`, dan `user.csv`. Setelah proses integrasi berdasarkan pasangan kunci `Place_Id` dan `User_Id`, diperoleh sebanyak 2.870 interaksi pengguna–destinasi yang memuat informasi `User_Id`, `Age`, `Place_Id`, `Place_Name`, dan `Place_Ratings`.

Secara keseluruhan, dataset ini mencakup 300 pengguna unik dan 126 destinasi wisata di Kota Yogyakarta. Nilai rating yang digunakan berada pada skala 1 sampai 5, dengan rata-rata rating keseluruhan sebesar 3,10. Hal ini menunjukkan kecenderungan penilaian yang relatif positif terhadap destinasi wisata yang ada.

Distribusi pengguna berdasarkan kelompok usia ditunjukkan pada Tabel 1. Segmentasi usia mengikuti pembagian: <18, 18–25, 26–35, dan >35 tahun. Berdasarkan data pengguna (`user.csv`), kelompok usia 18–25 mencakup 102 pengguna, 26–35 sebanyak 143 pengguna, dan >35 sebanyak 55 pengguna. Tidak terdapat pengguna dengan usia di bawah 18 tahun pada dataset ini, sehingga kelompok <18 tidak terisi. Jika dilihat dari jumlah interaksi rating (bukan jumlah pengguna), distribusi per kelompok usia pada data terintegrasi adalah sebagai berikut:

a. 18-25 tahun = 937 interaksi rating

b. 26-35 tahun = 1.385 interaksi rating

c. >35 tahun = 548 interaksi rating

Hal ini menunjukkan bahwa kelompok usia 26–35 tahun merupakan segmen yang paling aktif dalam memberikan rating pada destinasi wisata. Berdasarkan analisis rata-rata rating per destinasi, lima tempat wisata dengan rata-rata rating tertinggi (tanpa filter jumlah minimal rating, namun seluruh lima destinasi ini memiliki sedikitnya 17 rating) ditunjukkan pada Tabel 2.

Tabel 1. Lima destinasi dengan rata-rata rating tertinggi

No	Nama Destinasi	Jumlah Rating	Rata-rata Rating
1	Puncak Gunung Api Purba – Nglanggeran	17	3,88
2	Monumen Kembali	22	3,77
3	Bukit Bintang Yogyakarta	17	3,77
4	Pantai Baron	23	3,70
5	Desa Wisata Gamplong	30	3,63

Secara umum, destinasi dengan nilai rata-rata rating tertinggi didominasi oleh objek wisata alam dan sejarah, yang mengindikasikan preferensi wisatawan terhadap kombinasi daya tarik pemandangan dan nilai historis/budaya. Jika ditinjau per kelompok usia, pola preferensi yang muncul juga menarik. Dengan batas minimal 5 rating per destinasi untuk mengurangi bias dari

jumlah rating yang terlalu sedikit, tiga destinasi dengan rata-rata rating tertinggi di setiap kelompok usia adalah sebagai berikut:

1. Kelompok usia 18–25 tahun
 - a. Ledok Sambi (rata-rata rating 4,50; 6 rating)
 - b. Sumur Gumuling (4,40; 5 rating)
 - c. Kampung Wisata Rejowinangun (3,86; 7 rating)
2. Kelompok usia 26–35 tahun
 - a. Pantai Baron (4,20; 10 rating)
 - b. Air Terjun Kedung Pedut (4,11; 9 rating)
 - c. Kampung Wisata Taman Sari (4,10; 10 rating)
3. Kelompok usia >35 tahun
 - a. Sumur Gumuling (4,50; 6 rating)
 - b. Monumen Yogya Kembali (4,40; 5 rating)
 - c. Pantai Ngrawe (Mesra) (4,20; 5 rating)

Dari hasil tersebut tampak bahwa kelompok usia yang lebih muda (18–25 tahun) relatif tertarik pada destinasi bernuansa alam dan area rekreasi terbuka yang mendukung aktivitas eksplorasi dan swafoto, seperti Ledok Sambi. Sementara itu, kelompok usia >35 tahun cenderung memberikan rating lebih tinggi pada destinasi yang memiliki nilai historis dan edukatif, seperti Monumen Yogya Kembali dan Sumur Gumuling.

2. Hasil Evaluasi Model

Evaluasi model dilakukan dengan membagi dataset menjadi data latih (80%) dan data uji (20%) secara acak. Data latih digunakan untuk membangun tiga model rekomendasi, yaitu:

1. *Baseline Global Top-N*: memprediksi rating berdasarkan rata-rata rating global setiap destinasi. Lakukan perhitungan yang sama pada semua data, dengan contoh perhitungan seperti berikut

- a. Data 1 - User 1, tempat 154 (Pantai Ngrawe), Usia 18-25
 Di data latih, destinasi Place_Id = 154 punya 20 rating dengan nilai: (5, 5, 3, 5, 2, 4, 2, 1, 5, 3, 2, 4, 3, 5, 4, 2, 5, 4, 3, 3)

Jumlah dan rata-ratanya:

$$\begin{aligned} \sum r_{154} &= 71 \\ N_{154} &= 20 \\ \bar{r}_{154} &= \frac{71}{20} = 3,55 \end{aligned}$$

Jadi prediksi Baseline untuk user 1: $\hat{r}_{1,154}^{(baseline)} = 3,55$

- b. Data 2 - User 14, tempat 167 (Geoforest Watu Payung), usia 26-35
 Di data latih, Place_Id = 167 punya 26 rating, dengan rata-rata:
 $\bar{r}_{167} = 3,0769$
 Sehingga $\hat{r}_{14,167}^{(baseline)} = 3,0769$

- c. Data 3 – User 7, tempat 163 (Jogja Bay), usia > 35
 Di data latih, Place_Id = 183 punya 26 rating, rata-ratanya:
 $\bar{r}_{183} = 3,3462$
 Sehingga $\hat{r}_{7,183}^{(baseline)} = 3,3462$

2. *Collaborative filtering berbasis item (CF)*: memanfaatkan kemiripan antar destinasi berdasarkan pola rating pengguna. Lakukan perhitungan yang sama pada semua data, dengan contoh perhitungan seperti berikut :

- a. Data 1 - User 1, tempat 154 (Pantai Ngrawe), Usia 18-25

Tabel 2. Riwayat rating user 1 di data latih

No	Place Id	Place Name	Rating
1.	103	Tugu Pal Putih Jogja	3
2.	101	Kampung Wisata Sosro Menduran	4
3.	208	Taman Sungai Mudal	5

Dibutuhkan kemiripan $\text{sim}(154, j)$ untuk tiap item yang dirating user 1. Dari perhitungan cosine similarity (berdasarkan user-user lain yang merating pasangan destinasi tersebut di train):

$$\text{sim}(154,103) = 0,5806$$

$$\text{sim}(154,101) = 0,8321$$

$$\text{sim}(154,208) = 1,0$$

$$\text{Sehingga prediksi CF : } \hat{r}_{1,154}^{(CF)} = 4,17$$

- b. Data 2 - User 14, tempat 167 (Geoforest Watu Payung), usia 26-35
 Untuk prediksi rating di test, user 14 sekarang memberi rating aktual 4 untuk Place 167.
 Dari cosine similarity antar item (di train). Sehingga prediksi CF : $\hat{r}_{14,167}^{(CF)} = 3,33$
- c. Data 3 – User 7, tempat 163 (Jogja Bay), usia > 35
 Kita prediksi rating user 7 untuk 183 di data uji (rating aktual: 5). Dihitung cosine similarity antar item (berdasarkan rating user lain) Sehingga prediksi CF : $\hat{r}_{7,183}^{(CF)} = 3,82$

3. Demographic filtering (DF): memanfaatkan rata-rata rating per kombinasi kelompok usia (age_group) dan destinasi (Place_Id), dengan *fallback* ke rata-rata global destinasi jika kombinasi usia–destinasi belum pernah muncul pada data latih. Lakukan perhitungan yang sama pada semua data, dengan contoh perhitungan seperti berikut :

- a. Data 1 - User 1, tempat 154 (Pantai Ngrawe), Usia 18-25
 Kita lihat **data latih** untuk Place_Id = 154 dan Age_group = "18–25". Rating yang masuk kelompok ini: [5, 5, 2, 2, 4]

$$\sum r_{18-25,154} = 18$$

$$N_{18-25,154} = 5$$

$$\text{Sehingga prediksi DF: } \hat{r}_{1,154}^{(DF)} = 3,6$$

- b. Data 2 - User 14, tempat 167 (Geoforest Watu Payung), usia 26-35
 Di data latih, untuk kombinasi (Age_group = 26–35, Place_Id = 167, rating-nya: [1, 4, 3, 5, 4, 1, 3, 5, 3, 5, 3])

$$\sum r_{26-35,167} = 37$$

$$N_{26-35,167} = 11$$

$$\text{Sehingga prediksi DF: } \hat{r}_{14,167}^{(DF)} = 3,36$$

- c. Data 3 – User 7, tempat 163 (Jogja Bay), usia > 35
 Di data latih, (Age_group = ">35", Place_Id = 183) punya rating: [5, 5, 3, 4, 4, 3, 4]

$$\sum r_{>35,183} = 28$$

$$N_{>35,183} = 7$$

$$\text{Sehingga prediksi DF: } \hat{r}_{7,183}^{(DF)} = 4,0$$

Jika suatu kombinasi (Age_group, Place_Id) **tidak ada** di train, maka:

$$\hat{r}_{u,i}^{(DF)} = \bar{r}_i \text{ (sama seperti Baseline).}$$

4. Evaluasi Kinerja 3 Model

- a. Hitung Error Absolut per Baris (mini 3 baris)
 Pakai 3 baris contoh tsebelumny dan tiga model:
- Baris 1 – User 1, Place 154, usia 18–25, rating aktual 2
1. Baseline: $\hat{r} = 3,55 \rightarrow |error| = |2 - 3,55| = 1,55$
 2. DF : $\hat{r} = 3,60 \rightarrow |error| = |2 - 3,60| = 1,60$
 3. CF : $\hat{r} = 4,17 \rightarrow |error| = 2,17$
- Baris 2 – User 14, Place 167, usia 26–35, rating aktual 4
1. Baseline: $\hat{r} = 3,08 \rightarrow |error| = 0,92$
 2. DF : $\hat{r} = 3,36 \rightarrow |error| = 0,64$
 3. CF : $\hat{r} = 3,33 \rightarrow |error| = 0,67$
- Baris 3 – User 7, Place 183, usia >35, rating aktual 5
1. Baseline: $\hat{r} = 3,35 \rightarrow |error| = 1,65$
 2. DF : $\hat{r} = 4,0 \rightarrow |error| = 1,0$
 3. CF : $\hat{r} = 3,82 \rightarrow |error| = 1,18$
- b. Menghitung MAE Global (mini 3 baris)
 Dalam perhitungan menggunakan 547 baris data uji, dengan contoh perhitungan beberapa datanya sebagai berikut :

Baseline

$$MAE_{\text{mini}}^{(baseline)} = \frac{1}{3}(1,55 + 0,92 + 1,65) = \frac{4,12}{3} = 1,38$$

DF

$$MAE_{\text{mini}}^{(DF)} = \frac{1}{3}(1,60 + 0,64 + 1,00) = \frac{3,24}{3} = 1,08$$

CF

$$MAE_{\text{mini}}^{(CF)} = \frac{1}{3}(2,17 + 0,67 + 1,18) = \frac{4,02}{3} = 1,34$$

Dengan cara kerja MAE kumpulkan *error* absolut tiap baris → jumlah → bagi jumlah baris. MAE per kelompok usia (mini) sama dengan *error* absolut di baris itu:

1. Kelompok 18–25:
 - a. Baseline = 1,55
 - b. DF = 1,60
 - c. CF = 2,17
2. Kelompok 26–35:
 - a. Baseline = 0,92
 - b. DF = 0,64
 - c. CF = 0,67
3. Kelompok >35:
 - a. Baseline = 1,65
 - b. DF = 1,00
 - c. CF = 1,18

Kinerja ketiga model dievaluasi menggunakan metrik *Mean Absolute Error* (MAE) pada data uji. Hasil evaluasi MAE secara global ditunjukkan pada Tabel 3.

Tabel 3. MAE global pada data uji

No	Model	MAE Global
1.	Baseline Global Top-N	1,222
2.	<i>Collaborative filtering</i> (item-based)	1,301
3.	<i>Demographic filtering</i> (berbasis usia)	1,283

Dari Tabel 3, terlihat bahwa model *baseline* global justru memberikan nilai MAE paling rendah (1,222), diikuti oleh *Demographic filtering* (1,283) dan *Collaborative filtering* (1,301). Pada dataset ini, pendekatan sederhana berbasis rata-rata rating global destinasi masih cukup kompetitif dan bahkan sedikit lebih baik dibandingkan dua model berbasis personalisasi.

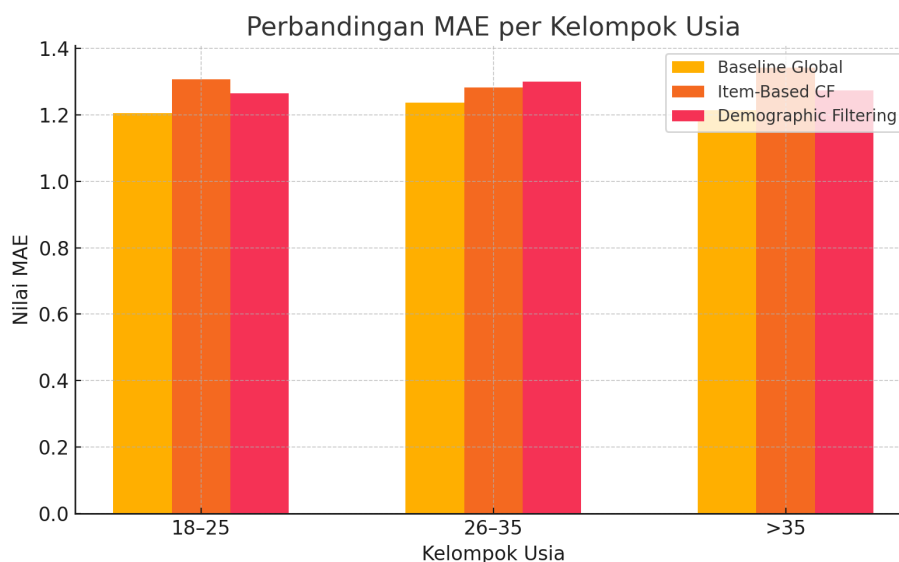
Untuk memperoleh gambaran yang lebih rinci, MAE juga dihitung secara terpisah untuk setiap kelompok usia yang hasilnya disajikan pada Tabel 3.

Tabel 4. MAE per kelompok usia pada data uji

No	Kelompok Usia	Baseline Global	CF (item-based)	DF (usia)
1.	18–25	1,205	1,307	1,264
2.	26–35	1,236	1,282	1,300
3.	>35	1,213	1,342	1,273

Interpretasi awal dari Tabel 4 adalah sebagai berikut:

1. Pada kelompok usia 18–25 tahun, model Baseline Global memberikan MAE terkecil (1,205), sehingga masih menjadi pendekatan paling akurat. *Demographic filtering* (DF) berada di urutan kedua dengan MAE 1,264, sedangkan *Collaborative filtering* (CF) memiliki MAE tertinggi (1,307) pada segmen ini.
2. Pada kelompok usia 26–35 tahun, Baseline Global kembali menghasilkan MAE paling rendah (1,236). *Collaborative filtering* sedikit lebih baik dibandingkan *Demographic Filtering*, dengan MAE masing-masing 1,282 (CF) dan 1,300 (DF). Namun, keduanya tetap belum mampu mengungguli akurasi *baseline*.
3. Pada kelompok usia >35 tahun, ketiga model menunjukkan pola yang serupa: Baseline Global tetap memberikan MAE terkecil (1,213), diikuti oleh *Demographic filtering* dengan MAE 1,273, sementara *Collaborative filtering* memiliki MAE terbesar (1,342). Hal ini mengindikasikan bahwa, bahkan pada segmen usia lebih tua, pendekatan rata-rata global destinasi masih lebih stabil dibanding dua model berbasis personalisasi.



Gambar 1. Perbandingan MAE per kelompok usia

Pada gambar 1 menunjukkan perbandingan nilai MAE dari tiga model rekomendasi berdasarkan kelompok usia pengguna. Secara umum, hasil ini menunjukkan bahwa personalisation tidak otomatis menghasilkan *error* yang lebih rendah dibandingkan model popularitas global, khususnya ketika ukuran dataset relatif terbatas dan pola preferensi antar pengguna belum cukup kaya atau konsisten.

B. Pembahasan

1. Analisis Kinerja Model

Hasil evaluasi menunjukkan bahwa secara global model Baseline Global (rata-rata rating per destinasi) masih menjadi pendekatan dengan kinerja terbaik. Nilai MAE global Baseline tercatat sekitar 1,222, lebih rendah dibandingkan *Collaborative filtering* (CF) yang mencapai 1,301 dan *Demographic filtering* (DF) berbasis usia dengan MAE sekitar 1,283.

Secara praktis, nilai MAE $\approx 1,2$ pada skala rating 1–5 berarti bahwa rata-rata kesalahan prediksi berada di kisaran ± 1 poin rating dari nilai aktual. Dengan kata lain, ketika pengguna memberikan rating 4, model rata-rata memprediksi berada di sekitar rentang 3–5. Untuk konteks sistem rekomendasi sederhana berbasis rating, performa ini masih dapat diterima, namun jelas belum “tajam”. Fakta bahwa model paling sederhana justru memiliki MAE terkecil mengindikasikan beberapa hal:

1. Kepadatan dan struktur data

Dataset hanya berisi 2.870 interaksi dari 300 pengguna terhadap 126 destinasi. Dengan kepadatan seperti ini, pola ko-rating yang dibutuhkan oleh CF untuk membangun kemiripan item yang kuat belum terbentuk secara optimal. Cosine similarity yang dihitung antar destinasi cenderung dibangun dari jumlah user yang terbatas per pasangan, sehingga estimasi kemiripannya rentan noise.

2. Distribusi rating yang relatif “ramai di tengah”

Rata-rata rating global berada di sekitar 3,10 dengan simpangan baku sekitar 1,39. Ini mengindikasikan kecenderungan pengguna memberi rating di kisaran tengah–tinggi. Dalam distribusi seperti ini, model yang selalu “menembak” mendekati rata-rata per destinasi sudah cukup kompetitif. Upaya memodifikasi prediksi dengan informasi tambahan (usia atau pola ko-rating) tidak serta-merta meningkatkan akurasi, dan bahkan berpotensi menambah varians kesalahan.

3. Keterbatasan sinyal personalisasi

Baik CF maupun DF pada dasarnya mencoba memanfaatkan sinyal personalisasi: CF dari pola perilaku pengguna lain, DF dari segmentasi usia. Namun, dengan ukuran data yang terbatas dan tanpa fitur tambahan lain (misalnya kategori destinasi, preferensi tematik, atau riwayat kunjungan yang lebih panjang), sinyal personalisasi tersebut belum cukup kuat untuk mengalahkan *baseline* berbasis popularitas destinasi.

Secara keseluruhan, dari sudut pandang optimasi MAE global, sistem rekomendasi yang hanya menggunakan rata-rata rating per destinasi sudah memberikan benchmark yang sulit dilampaui oleh dua model personalisasi yang diuji dalam penelitian ini.

2. Implikasi Terhadap Sistem Rekomendasi Wisata

Meskipun secara numerik model *baseline* global masih menjadi yang terbaik dalam hal MAE, *Demographic filtering* tetap memberikan nilai tambah yang penting dari sisi interpretabilitas dan personalisasi:

1. Interpretabilitas preferensi per segmen usia

Dengan menggunakan DF, pengelola sistem rekomendasi dapat melihat dengan jelas destinasi apa saja yang paling disukai oleh kelompok usia tertentu. Hal ini tercermin dari tren berikut :

- a. Kelompok 18–25 tahun cenderung memberikan rating tinggi pada Ledok Sambu dan Sumur Gumuling, yang menunjukkan preferensi terhadap wisata alam dan spot foto yang unik.
- b. Kelompok >35 tahun memberikan rating tinggi pada Monumen Yogya Kembali dan Sumur Gumuling, yang mengindikasikan apresiasi yang lebih besar terhadap destinasi bernuansa sejarah dan religi.

Informasi ini penting bagi pengelola pariwisata untuk merancang strategi promosi yang tersegmentasi berdasarkan usia.

2. Peran *baseline* dalam integrasi model

Hasil bahwa *baseline* global memiliki MAE yang paling rendah tidak berarti pendekatan berbasis demografi atau CF harus ditinggalkan. Sebaliknya, *baseline* dapat dijadikan komponen fallback dalam desain sistem rekomendasi hibrida, misalnya:

- a. Menggunakan DF sebagai prioritas untuk segmen usia yang datanya cukup padat (misal 18–25 dan 26–35),
- b. Menggunakan *baseline* global ketika kombinasi usia–destinasi kurang data atau ketika model DF justru meningkatkan *error*, seperti pada kelompok usia >35 tahun.

3. Arah pengembangan lanjutan

Beberapa pengembangan yang dapat dilakukan pada penelitian berikutnya antara lain:

- a. Menambah fitur demografis lain, seperti jenis kelamin atau asal daerah, untuk melihat apakah kombinasi multi-demografi dapat menurunkan MAE lebih jauh.
- b. Mengadopsi pendekatan hybrid yang menggabungkan skor CF dan DF, misalnya melalui kombinasi berbobot antara skor kemiripan perilaku dan rata-rata rating per kelompok usia.
- c. Menggunakan metrik evaluasi yang berorientasi pada kualitas rekomendasi top-N (misalnya Precision@K atau Recall@K) untuk melengkapi analisis berbasis MAE.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa:

- a. Model popularitas global masih menjadi *baseline* yang kuat untuk dataset pariwisata Yogyakarta yang relatif terbatas.
- b. *Demographic filtering* berbasis usia memberikan kinerja yang kompetitif pada segmen tertentu (khususnya usia 18–25 tahun) dan sangat berguna untuk memahami pola preferensi wisata per kelompok usia.
- c. *Collaborative filtering* pada skala dataset ini belum mampu mengungguli *baseline*, sehingga perlu dukungan data yang lebih kaya atau teknik regularisasi yang lebih kuat untuk meningkatkan performa.

Secara umum, hasil menunjukkan bahwa model *baseline* global berbasis rata-rata rating per destinasi mampu memberikan prediksi rating yang lebih akurat dibandingkan dua pendekatan personalisasi. Nilai MAE yang berada pada kisaran 1,2 berarti rata-rata kesalahan prediksi masih dalam rentang satu poin dari nilai aktual, yang secara praktis masih dapat diterima dalam sistem rekomendasi skala kecil.

Kinerja superior *baseline* ini dapat dijelaskan oleh beberapa faktor. Pertama, struktur data yang relatif jarang membuat metode *collaborative filtering* kesulitan membentuk estimasi kemiripan yang andal. Kedua, distribusi rating yang cenderung menumpuk pada nilai tengah menguntungkan model yang memprediksi nilai rata-rata. Ketiga, sinyal personalisasi dari usia pengguna belum cukup kuat untuk memperbaiki prediksi secara signifikan, terutama pada kelompok usia yang datanya tidak terlalu padat.

Meskipun demikian, *demographic filtering* tetap memberikan manfaat dari sisi interpretabilitas. Model ini memungkinkan pemetaan preferensi wisata berdasarkan kelompok usia secara eksplisit, yang sangat bermanfaat dalam konteks perancangan strategi promosi destinasi wisata. Misalnya, destinasi seperti Monumen Yogya Kembali dapat diprioritaskan untuk promosi kepada segmen usia dewasa, sementara destinasi terbuka seperti Ledok Sambu lebih cocok untuk segmen usia muda.

Temuan ini sejalan dengan literatur sebelumnya yang menunjukkan bahwa sistem rekomendasi tidak selalu memperoleh peningkatan akurasi dari kompleksitas model, terutama ketika data terbatas (Klimashevskaja et al., 2024)(Zhao et al., 2025). Sebaliknya, *baseline* yang kuat dapat menjadi fondasi penting bagi pengembangan model hibrida di masa mendatang.

KESIMPULAN

Penelitian ini menyimpulkan bahwa pada konteks dataset pariwisata Yogyakarta yang berskala relatif kecil dan bersifat rating-only, model *baseline* global berbasis rata-rata rating destinasi masih menjadi pendekatan yang paling akurat untuk tugas prediksi rating, dengan nilai MAE global yang lebih rendah dibandingkan model item-based *collaborative filtering* dan *demographic filtering* berbasis usia. Hasil evaluasi per kelompok usia menunjukkan pola yang konsisten, yaitu *baseline* global tetap memberikan kesalahan prediksi paling kecil pada segmen usia 18–25, 26–35, maupun >35, sedangkan dua model personalisasi belum mampu melampaui kinerja *baseline*. Temuan ini mengindikasikan bahwa kompleksitas model tidak secara otomatis menjamin peningkatan akurasi, terutama ketika data interaksi pengguna terbatas dan pola ko-rating belum cukup kaya. Meskipun demikian, *demographic filtering* tetap memberikan nilai tambah dari sisi pemahaman preferensi wisata per kelompok usia, sehingga dapat dimanfaatkan untuk merancang strategi promosi yang lebih tersegmentasi. Kontribusi utama penelitian ini terletak pada penyajian analisis komparatif yang eksplisit antara *baseline* global, *collaborative filtering* standar, dan *demographic filtering* usia, baik secara global maupun per segmen usia, serta penegasan batasan generalisasi hasil pada jenis data rating-only, model dasar tanpa regularisasi lanjutan, dan metrik MAE yang berfokus pada kesalahan prediksi rating, bukan kualitas ranking top-N. Penelitian lanjutan disarankan untuk mengeksplorasi dataset yang lebih besar dan lebih padat, menambahkan atribut kontekstual lain seperti kategori destinasi atau variabel demografis tambahan, menguji model hibrida yang menggabungkan *baseline*, CF, dan DF, serta melengkapi evaluasi dengan metrik berbasis ranking dan studi kepuasan pengguna sehingga gambaran kinerja sistem rekomendasi wisata dapat dinilai secara lebih menyeluruh.

DAFTAR PUSTAKA

- Arregoces-Julio, I., Solano-Barliza, A., Valls, A., Moreno, A., Castillo-Palacio, M., Acosta-Coll, M., & Escorcia-Gutierrez, J. (2025). A Flexible Profile-Based Recommender System for Discovering Cultural Activities in an Emerging Tourist Destination. *Informatics*, 12(3), 81. <https://doi.org/10.3390/informatics12030081>
- Bauer, C., Zangerle, E., & Said, A. (2024). Exploring the Landscape of Recommender Systems Evaluation: Practices and Perspectives. *ACM Transactions on Recommender Systems*, 2(1), 11–31. <https://doi.org/10.1145/3629170>
- Carnovalini, F., Roda, A., & Wiggins, G. A. (2025). Popularity Bias in Recommender Systems: The Search for Fairness in the Long Tail. *Information*, 16(2), 151. <https://doi.org/10.3390/info16020151>
- Casillo, M., Colace, F., Conte, D., Lombardi, M., Santaniello, D., & Valentino, C. (2023). Context-aware recommender systems and cultural heritage: a survey. *Journal of Ambient Intelligence and Humanized Computing*, 14, 3109–3127. <https://doi.org/10.1007/s12652-021-03438-9>
- Chizari, N., Tajfar, K., & Moreno-García, M. N. (2024). Assessing Unfairness in GNN-Based Recommender Systems: A Focus on Metrics for Demographic Sub-Groups. *Proceedings of the 20th International Conference on Web Information Systems and Technologies (WEBIST 2024)*, 433–440. <https://doi.org/10.5220/0013069400003825>
- Deldjoo, Y., Jannach, D., Bellogin, A., Difonzo, A., & Zanzonelli, D. (2024). Fairness in recommender systems: research landscape and future directions. *User Modeling and User-Adapted Interaction*, 34, 59–108. <https://doi.org/10.1007/s11257-023-09364-z>
- Dewi, M. M., Andriani, R., & Nuraminudin, M. (2025). Performance Analysis of the Item-Based Collaborative filtering Model in Yogyakarta Tourism Recommendations. *Journal of Applied Informatics and Computing*, 9(2), 534–541.
- Fahrizal, D., Kustija, J., & Akbar, M. A. H. (2024). Development Tourism Destination Recommendation Systems using Collaborative and Content-Based Filtering Optimized with Neural Networks. *Indonesian Journal of Artificial Intelligence and Data Mining*, 7(2).
- Fkih, F. (2022). Similarity measures for Collaborative filtering-based Recommender Systems: Review and experimental comparison. *Journal of King Saud University - Computer and Information Sciences*, 34(9), 7645–7669. <https://doi.org/10.1016/j.jksuci.2021.09.014>
- Generosi, A., Villafan, J. Y., Ferretti, M., & Mengoni, M. (2025). A -based web platform to boost tourism in marginal territories. *Information Technology & Tourism*, 27(3), 797–831. <https://doi.org/10.1007/s40558-025-00327-1>
- Hartatik, H., Nurhayati, S. D., & Widayani, W. (2021). Sistem Rekomendasi Wisata Kuliner di Yogyakarta dengan Metode Item-Based Collaborative filtering. *Journal Automation Computer Information System (JACIS)*, 1(2), 55–63. <https://doi.org/10.47134/jacis.v1i2.8>
- Huda, M. T. (2023). Recommendation System for Mobile Applications Tour Guide and Travel Services using Demographic filtering and Content-Based Filtering Methods based on Android. *Jurnal Web*

Sains.

- Klimashevskaja, A., Jannach, D., Elahi, M., & Trattner, C. (2024). A survey on popularity bias in recommender systems. *User Modeling and User-Adapted Interaction*, 34(5), 1777–1834. <https://doi.org/10.1007/s11257-024-09406-0>
- Puri, F. M., Ramadhani, S. T. A. R., Istiqomah, D. A., Windarni, V. A., & Hidayat, K. R. (2024). Analisis Rekomendasi Pembuatan Produk Menggunakan RStudio Dan Twitter (Studi Kasus : Git Solution). *JURNAL TECNOSCIENZA*, 9(1), 97–112. <https://doi.org/https://doi.org/10.51158/c9z0w740>
- Rajesh, D. B., & Kumar, A. (2025). Collaborative filtering models an experimental and detailed comparative study. *Scientific Reports*, 15(1), 31667. <https://doi.org/10.1038/s41598-025-15096-4>
- Sispianygala, A. (2024). Pengembangan Aplikasi Sistem Rekomendasi Tempat Wisata di Jakarta Menggunakan Collaborative filtering. *Jurnal Progresif*, 20(2), 828–838.
- Solano-Barliza, A., Arregoces-Julio, I., Valls, A., & Moreno, A. (2024). Recommender systems applied to the tourism industry: a literature review. *Cogent Business & Management*, 11(1), 2367088. <https://doi.org/10.1080/23311975.2024.2367088>
- Zhao, Y., Wang, Y., Liu, Y., Cheng, X., Aggarwal, C. C., & Derr, T. (2025). Fairness and Diversity in Recommender Systems: A Survey. *ACM Transactions on Intelligent Systems and Technology*, 16(1), 1–28. <https://doi.org/10.1145/3664928>