

Sistem Rekomendasi Hotel Menggunakan Algoritma *Item-Based Clustering Hybrid Method*

Ngakan Putu Bagus Ananta Wijaya^{a1}, Ida Ayu Gde Suwiprabayanti Putra^{a2}, I Gusti Ngurah Anom Cahyadi Putra^{a3}, Putu Praba Santika^{a4}, I Putu Gede Hendra Suputra^{a5}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Udayana, Jimbaran, Kuta Selatan, Badung, Bali, Indonesia

¹wijaya.2208561120@student.unud.ac.id

²iagsuwiprabayantiputra@unud.ac.id

³anom.cp@unud.ac.id

⁴praba@unud.ac.id

⁵hendra.suputra@unud.ac.id

Abstrak

Penelitian ini dilaksanakan dengan tujuan untuk membangun sistem rekomendasi hotel berbasis algoritma Hybrid Recommendation Sistem (ICHM) yang menggabungkan metode Content Based Filtering dan Collaborative Filtering untuk membantu pengguna dalam menemukan hotel sesuai preferensi mereka. Metode Content Based Filtering menggunakan teknik TF-IDF dan Cosine Similarity yang berfungsi untuk menganalisis fitur konten hotel. Sedangkan metode Collaborative Filtering menggunakan matriks rating pengguna dengan proses mean centering dan Cosine Similarity yang berfungsi untuk mengelompokkan hotel berdasarkan kemiripan karakteristik data. K-Means Clustering diterapkan untuk mengelompokkan hotel berdasarkan kemiripan karakteristik data. Evaluasi menghasilkan akurasi model sebesar 0,882, MAE sebesar 0,625, precision sebesar 0,907, recall sebesar 0,790, dan F1-score sebesar 0,845 yang menunjukkan bahwa model dapat memberikan rekomendasi hotel yang relevan dan akurat serta dapat diimplementasikan ke sistem rekomendasi hotel yang memungkinkan pengguna untuk melakukan rekomendasi hotel berdasarkan preferensi masing-masing.

Kata Kunci : Sistem Rekomendasi, Hybrid Recommendation Sistem, Content Based Filtering, Collaborative Filtering, K-Means Clustering

1. Pendahuluan

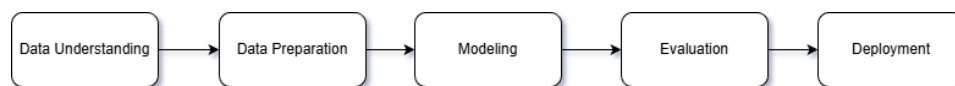
Industri perhotelan mengalami perkembangan pesat seiring meningkatnya mobilitas masyarakat dan kebutuhan akan akomodasi selama perjalanan. Data yang didapatkan dari BPS menunjukkan bahwa jumlah pemesanan hotel mengalami pertumbuhan rata-rata sekitar 15% per tahun dalam lima tahun terakhir [1]. Banyaknya pilihan hotel yang tersedia seringkali membuat konsumen pengguna / wisatawan kesulitan menentukan hotel yang sesuai dengan preferensi mereka. Oleh karena itu, sistem rekomendasi hotel menjadi solusi penting untuk membantu pengguna menemukan hotel yang relevan secara lebih cepat dan efisien sekaligus meningkatkan kepuasan pelanggan

Penelitian terkait *recommender system* telah banyak dikembangkan dengan menggunakan metode seperti *Collaborative Filtering* dan *Content-Based Filtering*. *Collaborative Filtering* mampu memberikan rekomendasi berdasarkan pola interaksi dan rating yang diberikan dari pengguna lain yang memiliki preferensi yang mirip [2]. Namun, metode ini memiliki kelemahan pada kondisi *cold-start* ketika data pengguna / item masih terbatas sehingga menyebabkan penurunan akurasi [3]. Sementara itu, *Content-Based Filtering* memanfaatkan karakteristik item seperti fasilitas, lokasi, harga dan tipe kamar pada hotel untuk menghasilkan rekomendasi berdasarkan dari kemiripan fitur [4]. Meskipun metode ini cukup efektif, rekomendasi yang dihasilkan cenderung kurang bervariasi dan hanya berfokus pada item yang karakteristiknya mirip [5]. Permasalahan tersebut menunjukkan bahwa penggunaan metode tunggal masih belum optimal dalam menghasilkan sebuah rekomendasi yang akurat dan relevan untuk pengguna

Beberapa penelitian terbaru mulai mengembangkan *hybrid recommender system* dengan menggabungkan metode *Collaborative Filtering* dan *Content-Based Filtering* untuk meningkatkan performa dari rekomendasi. Penelitian sebelumnya menunjukkan bahwa metode hybrid mampu meningkatkan kualitas rekomendasi dibandingkan dengan metode tunggal [6]. Selain itu, penelitian lainnya juga menunjukkan bahwa pendekatan hybrid dapat membantu mengatasi permasalahan *cold-start* dan meningkatkan relevansi rekomendasi [7]. Penelitian lain juga memanfaatkan *clustering* pada *recommender system* untuk mengelompokkan item berdasarkan kemiripan karakteristik sehingga proses pencarian rekomendasi menjadi lebih efektif dan terstruktur [8]. Namun, sebagian besar penelitian sebelumnya masih berfokus pada *e-learning*, *e-commerce*, berita, dan produk digital, sedangkan penerapan *hybrid recommender system* berbasis *clustering* pada sistem rekomendasi hotel masih relatif terbatas, khususnya dengan yang memanfaatkan kombinasi fitur hotel seperti fasilitas, lokasi, harga, tipe kamar dan rating pengguna

Berdasarkan permasalahan tersebut, penelitian ini mengembangkan sistem rekomendasi hotel berbasis hybrid recommendation system yang menggabungkan beberapa metode seperti *Collaborative Filtering*, *Content-Based Filtering*, dan *K-Means clustering* (ICHM). *Collaborative Filtering* digunakan untuk memanfaatkan pola rating pengguna, *Content-Based Filtering* digunakan untuk menganalisis kemiripan karakteristik hotel dan *K-Means Clustering* untuk mengelompokkan hotel berdasarkan kemiripan fitur agar menghasilkan rekomendasi yang lebih relevan. Penelitian ini bertujuan untuk mengimplementasikan sistem rekomendasi hotel menggunakan algoritma gabungan ICHM, mengetahui performa dari metode yang digunakan (*Content-Based Filtering*, *Collaborative Filtering*, dan *Hybrid Recommendation System* (ICHM)) dalam melakukan rekomendasi hotel dan mengetahui hasil pengelompokkan hotel menggunakan algoritma *K-Means Clustering*

2. Metode Penelitian



Gambar 1. CRISM-DM

2.1. Data Understanding

Tahap awal yaitu tahap data understanding yang bertujuan untuk melakukan pengenalan karakteristik data [9]. Diawali dengan pengumpulan data hotel yang diperoleh dari situs Kaggle, yang terdiri dari 6 kolom utama yaitu Hotel Name, Original Price, Price After Discount, Tax, Rating, dan Location. Selanjutnya dilakukan penambahan lima kolom fitur tambahan seperti Facility, Room Type, Room Size, Hotel Picture, dan Review berdasarkan hasil wawancara dengan pihak travel agent untuk memperkaya informasi hotel. Setelah itu dilakukan analisis awal dataset yang meliputi pengecekan jumlah data hotel, ukuran dataset, nama dan tipe kolom, serta pelaksanaan Exploratory Data Analysis (EDA) untuk memahami pola dan hubungan antar variabel data melalui visualisasi

2.2. Data Preparation

Tahap kedua yaitu tahap data preparation yang bertujuan untuk mempersiapkan dataset agar siap digunakan pada proses pemodelan. Prosesnya dimulai dari pembersihan data, pengecekan nilai duplikat, normalisasi data, serta ekstraksi fitur [10]. Proses ini penting untuk memastikan data yang digunakan memiliki kualitas yang baik dan berada pada skala yang seragam sehingga dapat meningkatkan akurasi dan stabilitas model dalam menghasilkan rekomendasi hotel

2.3. Modeling

Tahap ketiga yaitu tahap modeling yang bertujuan untuk pembuatan model yang menggunakan metode hybrid yang menggabungkan metode Content-Based Filtering, Collaborative Filtering, dan Clustering K-Means [11]. Pada tahap ini, fitur hotel seperti fasilitas, tipe kamar, lokasi, rating dan harga digunakan untuk membangun representasi karakteristik hotel melalui metode TF-IDF dan normalisasi fitur numerik. Selanjutnya, tingkat kemiripan antar hotel dihitung menggunakan cosine similarity baik berdasarkan fitur konten maupun pola rating pengguna [12]. Hasil similaritas dari kedua metode digabungkan menggunakan bobot tertentu untuk menghasilkan rekomendasi yang lebih akurat. K-Means Clustering digunakan untuk mengelompokkan hotel berdasarkan kemiripan karakteristik agar sistem dapat memberikan rekomendasi hotel yang lebih relevan sesuai preferensi pengguna [13][14]

2.4. Evaluation

Tahap keempat yaitu tahap evaluation yang bertujuan untuk menilai kinerja model sistem rekomendasi yang sudah dibangun dengan menggunakan beberapa metrik evaluasi, yaitu akurasi, presisi, recall, F1-Score, dan Mean Absolute Error (MAE) [15]

2.5. Deployment

Tahap terakhir yaitu tahap deployment yang bertujuan sebagai perancangan dan implementasi sistem rekomendasi hotel dengan *React JS* dan *Flask* [16]. Pada tahap ini, model rekomendasi yang telah dibuat, diintegrasikan ke dalam sistem rekomendasi hotel yang dapat digunakan langsung oleh pengguna. Proses ini mencakup perancangan arsitektur sistem, pembuatan antarmuka pengguna (user interface), pengelolaan basis data, serta pengujian fungsionalitas sistem agar memastikan seluruh fitur dapat berjalan dengan baik dan lancar. Pada tahap ini juga dilakukan pengaturan role pengguna (user dan admin) yang masing-masing memiliki fungsi, seperti pencarian rekomendasi hotel, pengelolaan data hotel, pemberian rating dan feedback, serta proses retraining model ketika terdapat penambahan data baru

3. Hasil Dan Pembahasan

3.1. Model

Dataset hotel yang digunakan berjumlah 1012 data dengan 11 atribut seperti hotel name, location, facility, room type, rating, original price, dan price after discount. Pada tahap awal dilakukan data cleaning untuk membersihkan data sehingga diperoleh 994 data hotel yang siap digunakan. Selanjutnya dilakukan normalisasi data untuk menyelaraskan skala fitur numerik. Model rekomendasi dibangun menggunakan Hybrid Recommendation System (ICHM) yang menggabungkan Content-Based Filtering dan Collaborative Filtering. Fitur konten hotel diubah menjadi representasi numerik menggunakan TF-IDF untuk data teks dan normalisasi untuk data numerik. Kemiripan antar hotel kemudian dihitung menggunakan cosine similarity untuk menghasilkan matriks kesamaan / similaritas hotel. Berikut hasil akurasi model dari kedua algoritma:

Tabel 1. Akurasi Model

Metode	Hasil
Content-Based Filtering	0.4045 (tingkat kemiripan sedang)
Collaborative Filtering	0.5246 (tingkat kemiripan cukup baik)

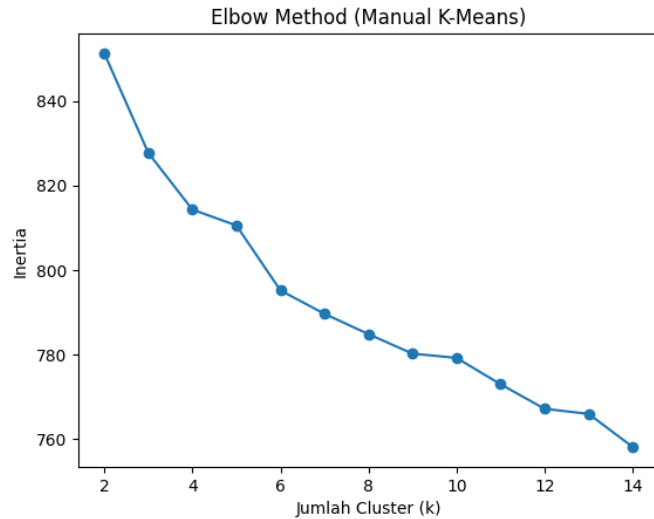
Hasil tersebut menunjukkan bahwa Collaborative Filtering memiliki nilai similaritas yang lebih tinggi dibandingkan Content-Based Filtering, sehingga keduanya kemudian digabungkan dalam metode Hybrid Recommendation System (ICHM) untuk meningkatkan akurasi rekomendasi. Berikut hasil akurasi dari model yang sudah digabungkan:

Tabel 2. Akurasi Matriks Hybrid

Metriks	Hasil
Akurasi	0.882
MAE	0.625
Precision	0.907
Recall	0.790
F1-Score	0.845

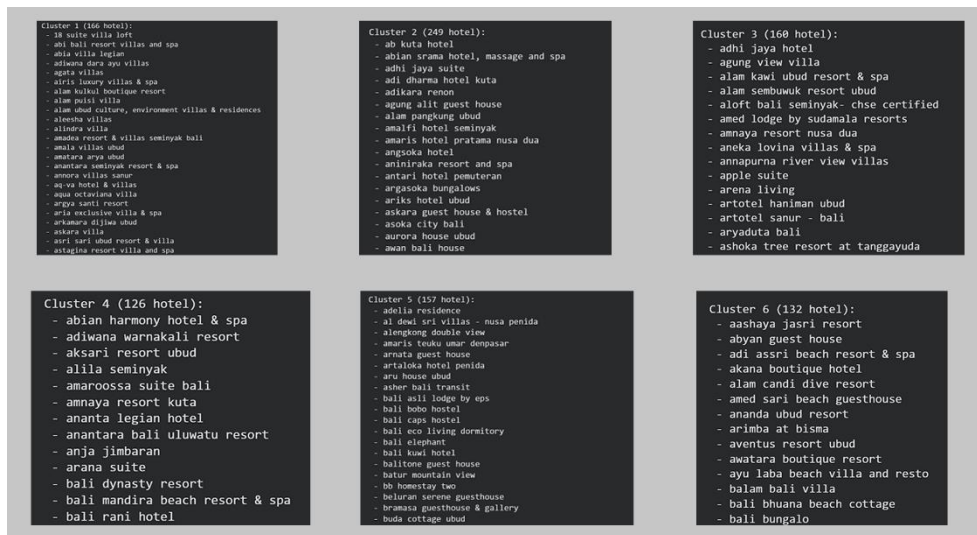
Hasil evaluasi model hybrid, menunjukkan peningkatan akurasi yang cukup baik dan kesalahan prediksi yang rendah.

Wijaya, Suwiprabayanti, Putra, Santika, Suputra
SISTEM REKOMENDASI HOTEL MENGGUNAKAN ALGORITMA ITEM-BASED CLUSTERING HYBRID METHOD



Gambar 2. Elbow Method

Selanjutnya dilakukan proses clustering menggunakan algoritma K-Means untuk mengelompokkan hotel berdasarkan kemiripan fitur. Berdasarkan dari hasil analisis Elbow Method dan Silhouette Score, diperoleh 6 cluster kelompok hotel dengan karakteristik yang serupa dan ini dapat membantu relevansi rekomendasi hotel berdasarkan preferensi pengguna



Gambar 3. Clustering Hasil

Tabel 3. Analisis Hasil Cluster

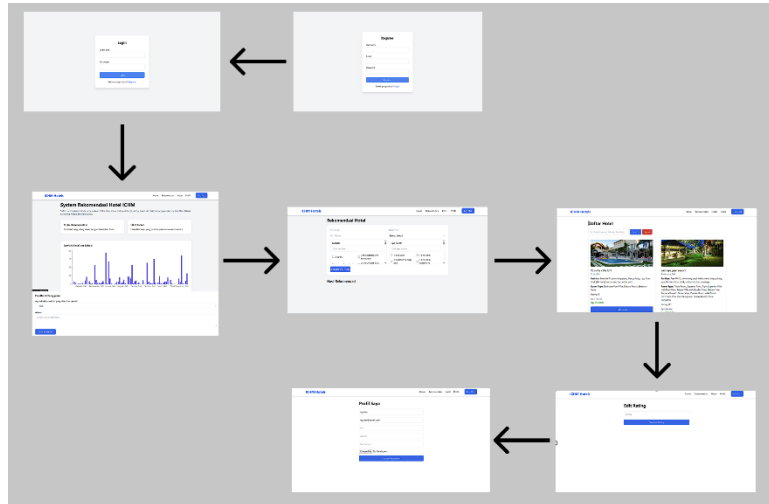
Cluster	Rata-rata Harga	Rata-rata Rating	Karakteristik Tipe Kamar	Karakteristik Fasilitas	Hotel
Cluster 1	Rp2.903.784	8,63	Villa, Suite, Private Pool Villa	Spa, private pool, beach access, resort facilities, swimming pool	Anantara Seminyak Resort, Jumeirah Bali, Six Senses Uluwatu Bali
Cluster 2	Rp336.678	8,20	Standard Room, Superior	WiFi, AC, parking, 24-hour reception	Cara Cara Inn, Kuta Central

			Room, Deluxe Room		Park Hotel, hotel kategori OYO
Cluster 3	Rp1.462.551	8,63	Deluxe Room, Superior Room, Suite	Swimming pool, gym, spa, restaurant, beach access	Grand Hyatt Bali, Padma Resort Legian, Hotel Indigo Bali
Cluster 4	Rp2.839.042	8,60	Luxury Suite, Executive Suite, Villa	Infinity pool, private beach, spa, fitness center	The Ritz-Carlton Bali, The Apurva Kempinski Bali, Hilton Bali Resort
Cluster 5	Rp198.438	8,08	Dormitory, Shared Room, Standard Room	WiFi, reception, parking	Bali Caps Hostel, Write Hostel, The Kutaya Guest House
Cluster 6	Rp764.115	8,52	Cottage, Villa, Deluxe Room	Beach access, diving activity, sea view, outdoor pool	Relax Bali Dive & Spa Resort, Mimpi Resort Tulamben, Dream Beach Cottages

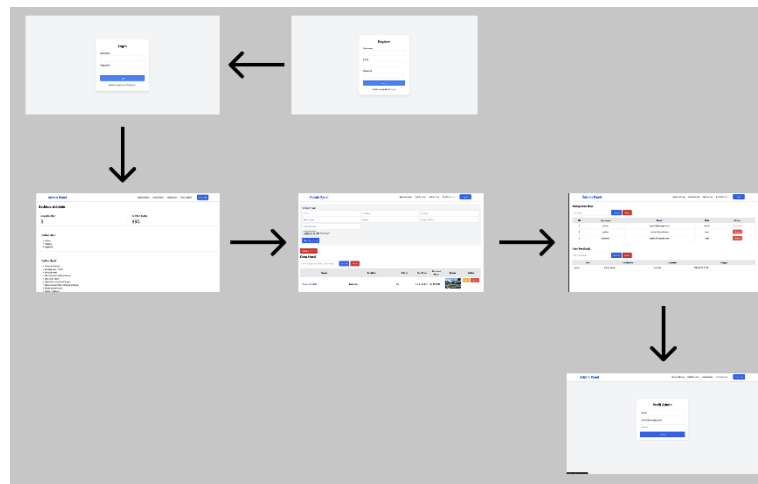
3.2. Sistem Rekomendasi

Sistem rekomendasi hotel diimplementasikan jika model sudah berhasil dibuat. Sistem ini menyediakan fitur login role (admin / user). Pengguna yang telah melakukan register dapat login ke sistem untuk mengakses menu Home, Feedback, Rekomendasi, Hotel, dan Profil. Pada bagian menu Rekomendasi, pengguna dapat memasukkan preferensi seperti minimum harga, maksimum harga, minimum rating, lokasi, fasilitas, dan tipe kamar. Berdasarkan preferensi tersebut, sistem akan memproses data menggunakan model rekomendasi yang telah dibangun dan menampilkan daftar hotel dengan tingkat kecocokan tertinggi serta informasi detail hotel seperti nama hotel, fasilitas, tipe kamar, harga, rating, dan gambar hotel. Pengguna juga dapat memberi feedback jika diperlukan

Sistem rekomendasi hotel ini juga memiliki akses admin. Admin dapat melihat jumlah pengguna dan hotel yang tersedia, menambah data hotel yang baru, menghapus data hotel, memperbarui harga hotel dan melakukan retraining model agar sistem dapat menyesuaikan dengan data terbaru. Admin juga dapat menghapus data pengguna dan melihat feedback yang diberikan pengguna. Dengan adanya sistem rekomendasi hotel ini, proses pencarian hotel menjadi lebih efisien dan mampu memberikan rekomendasi hotel yang lebih relevan.



Gambar 4. Halaman User



Gambar 5. Halaman Admin

4. Kesimpulan

Berdasarkan hasil perancangan dan implementasi, model rekomendasi hotel berhasil dibangun menggunakan pendekatan hybrid yang menggabungkan Content-Based Filtering dan Collaborative Filtering. Model Content-Based Filtering memanfaatkan fitur konten hotel seperti fasilitas, tipe kamar, lokasi, serta fitur numerik seperti rating dan harga dengan metode TF-IDF dan Cosine Similarity yang menghasilkan akurasi sebesar 0,4045 yang mampu menangkap tingkat kemiripan sedang karakteristik hotel. Sedangkan, model Collaborative Filtering menggunakan matriks rating pengguna, mean-centering, dan Cosine Similarity yang menghasilkan akurasi sebesar 0,5246 yang mampu menunjukkan kemampuan model yang cukup baik dalam mempelajari pola preferensi pengguna. Penggabungan kedua metode tersebut (hybrid) menghasilkan performa yang lebih baik dengan akurasi 0,882, MAE 0,625, precision 0,907, recall 0,790, dan F1-score 0,845, yang menunjukkan bahwa sistem mampu memberikan rekomendasi hotel yang cukup akurat dan relevan. Penerapan K-Means Clustering berhasil mengelompokkan hotel ke dalam 6 cluster berdasarkan kemiripan karakteristik hotel. Model yang dibuat juga sudah berhasil diimplementasikan ke dalam sistem rekomendasi hotel yang dapat membantu pengguna menemukan hotel sesuai preferensi mereka masing-masing.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada seluruh pihak yang telah membantu dan mendukung dalam penyelesaian penelitian Tugas Akhir ini, khususnya kepada Kaprodi Informatika, dosen pembimbing, dosen penguji, seluruh dosen Program Studi Informatika, pihak travel agent yang telah

membantu penyediaan kebutuhan data penelitian, serta orang tua dan keluarga yang selalu memberikan dukungan dan doa selama proses penelitian hingga penyusunan Tugas Akhir ini selesai

Referensi

- [1] BPS-Statistics Indonesia, "National Tourism and Transportation Development Statistics December 2019," 2020. [Online]. Available: <https://www.bps.go.id>.
- [2] B. Prasetyo, H. Haryanto, S. Astuti, E. Z. Astuti and Y. Rahayu, "Implementasi Metode Item-Based Collaborative Filtering dalam Pemberian Rekomendasi Calon Pembeli Aksesoris Smartphone," *Eksplora Informatika*, vol. 9, no. 1, pp. 17–27, 2019, doi: <https://doi.org/10.30864/eksplora.v9i1.244>.
- [3] Y. Setiawan, A. Nurwanto and A. Erlansari, "Implementasi Item Based Collaborative Filtering Dalam Pemberian Rekomendasi Agenda Wisata Berbasis Android," *Pseudocode*, vol. 6, no. 1, pp. 13–20, 2019, doi: <https://doi.org/10.33369/pseudocode.6.1.13-20>.
- [4] B. A. Larasati and H. Februariyanti, "Sistem Rekomendasi Product Emina Cosmetics Dengan Menggunakan Metode Content-Based Filtering," *Jurnal Manajemen Informatika Dan Sistem Informasi*, vol. 4, no. 1, pp. 45–54, 2021.
- [5] Z. K. A. Baizal and R. Andiety, "Content-Based Music Recommender System Using Deep Neural Network," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 2, pp. 1111–1119, 2024, doi: <https://doi.org/10.47065/bits.v6i2.5762>
- [6] K. R. Putra and M. A. Rachman, "Perbandingan Metode Content-based, Collaborative dan Hybrid Filtering pada Sistem Rekomendasi Lagu," *MIND (Multimedia Artificial Intelligent Networking Database) Journal*, vol. 9, no. 2, pp. 179–193, 2024, doi: <https://doi.org/10.26760/mindjournal.v9i2.179-19>.
- [7] A. D. A. Kusuma, A. Pramudwiatmoko and S. Suhirman, "Sistem Rekomendasi Pemilihan Film Menggunakan Metode Collaborative Filtering dan Content-Based Filtering," *INTEK*, vol. 8, no. 1, pp. 88–97, 2025.
- [8] D. Fahrizal, J. Kustija and M. A. H. Akbar, "Development Tourism Destination Recommendation Systems Using Collaborative and Content-Based Filtering Optimized with Neural Networks," *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIMD)*, vol. 7, no. 2, pp. 285–298, Sep. 2024.
- [9] F. A. Maresti, G. M. Anugraheni, R. A. Hargiyanto and K. Mustaqim, "Penerapan Exploratory Data Analysis (EDA) dan Analisis Recency, Frequency, and Monetary (RFM) untuk Segmentasi Pelanggan Ecommerce," *Competitive*, vol. 19, no. 1, pp. 14–25, 2024.
- [10] D. Septiani and I. Isabela, "Analisis Term Frequency Inverse Document Frequency (TF-IDF) dalam Temu Kembali Informasi pada Dokumen Teks," *SINTESIA: Jurnal Sistem Dan Teknologi Informasi Indonesia*, vol. 1, no. 2, pp. 81–88, 2022.
- [11] M. A. Pramono, "Implementasi Metode ICHM (Item Based Clustering Hybrid Method) untuk Rekomendasi Coffee Shop di Yogyakarta," 2020.
- [12] B. Prasetyo, H. Haryanto, S. Astuti, E. Z. Astuti and Y. Rahayu, "Implementasi Metode Item-Based Collaborative Filtering dalam Pemberian Rekomendasi Calon Pembeli Aksesoris Smartphone," *Eksplora Informatika*, vol. 9, no. 1, pp. 17–27, 2019.
- [13] Z. M. Devi, N. A. Setiawan, T. B. Adji and T. Widiyaningtyas, "Hybrid Filtering Algorithm in Event Manager Partner Recommendation System," in *Proceedings of the 7th International Conference on Sustainable Information Engineering and Technology*, 2023, pp. 182–187.
- [14] I. B. G. Sarasvananda, R. Wardoyo, and A. K. Sari, "The K-Means Clustering Algorithm With Semantic Similarity To Estimate The Cost of Hospitalization," *IJCCS (Indonesian J. Comput. Cybern. Syst.)*, vol. 13, no. 4, p. 313, 2019, doi: 10.22146/ijccs.45093.
- [15] D. M. W. Powers, "Evaluation: From Precision, Recall and F-Measure to Roc, Informedness, Markedness & Correlation," *Journal of Machine Learning Technologies*, vol. 2, pp. 37–63, 2011.
- [16] D. J. Evan and P. O. N. Saian, "Implementasi Python Framework Flask Pada Modul Transfer Out Toko di PT XYZ," *JUPI (Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika)*, vol. 8, no. 4, pp. 1121–1131, 2023.