

Analisis Pengaruh Fitur-Fitur Data Traveloka pada Kinerja Sistem Rekomendasi *Content Based Filtering* dengan *K-Nearest Neighbor* dan *Haversine Formula* pada Perhotelan di Kota Balikpapan

Aditya Akbar, Hasrullah
STMIK Borneo International, Indonesia
E-mail: adityaakbar09888@gmail.com

ABSTRAK

Dalam era digital, pengguna internet menghadapi banyak pilihan akomodasi, sehingga sistem rekomendasi menjadi solusi penting untuk membantu mereka menemukan hotel yang sesuai. Penelitian ini bertujuan untuk menganalisis pengaruh fitur-fitur data Traveloka terhadap kinerja sistem rekomendasi hotel dengan pendekatan *content-based filtering* menggunakan algoritma *K-Nearest Neighbor* (K-NN) dan Haversine Formula. Metode yang digunakan meliputi pengumpulan data dari situs Traveloka, mencakup 222 hotel bintang 3 hingga 5 di Balikpapan. Fitur yang dianalisis meliputi lokasi, harga, tipe kamar, dan fasilitas. Hasil penelitian menunjukkan bahwa penerapan metode *content-based filtering* efektif dalam memberikan rekomendasi yang relevan dan akurat sesuai dengan preferensi pengguna. Sistem ini juga mampu mengatasi masalah *cold start*, sehingga berguna bagi pengguna baru. Dengan demikian, penelitian ini berkontribusi pada pengembangan sistem rekomendasi yang lebih baik di industri perhotelan, meningkatkan kenyamanan dan kepuasan pengguna dalam merencanakan perjalanan mereka.

Kata Kunci: traveloka; kinerja sistem rekomendasi; perhotelan.

ABSTRACT

In the digital era, internet users face many accommodation choices, so that recommendation systems become an important solution to help them find suitable hotels. This study aims to analyze the effect of Traveloka data features on the performance of the hotel recommendation system with a content-based filtering approach using the K-Nearest Neighbor (K-NN) algorithm and Haversine Formula. The method used includes data collection from the Traveloka site, covering 222 3 to 5 star hotels in Balikpapan. The features analyzed include location, price, room type, and facilities. The results of the study indicate that the application of the content-based filtering method is effective in providing relevant and accurate recommendations according to user preferences. This system is also able to overcome the cold start problem, making it useful for new users. Thus, this study contributes to the development of better recommendation systems in the hospitality industry, increasing user convenience and satisfaction in planning their trips.

Keywords: traveloka; recommendation system performance; hospitality.



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

PENDAHULUAN

Dalam era digital saat ini, pengguna internet dihadapkan pada lautan informasi yang terus berkembang, termasuk dalam pencarian akomodasi seperti hotel (Wahyudi et al., 2020). Dengan banyaknya pilihan yang tersedia, seringkali pengguna merasa kesulitan

untuk menemukan hotel yang sesuai dengan kebutuhan dan preferensi mereka. Untuk mengatasi tantangan ini, sistem rekomendasi muncul sebagai solusi yang efektif, membantu pengguna menyaring informasi dan menemukan pilihan yang paling relevan (Fanny et al., 2024).

Sistem rekomendasi berbasis konten, khususnya, telah terbukti efektif dalam berbagai domain, termasuk *e-commerce* dan pariwisata (Iman & Wibisono, 2021). Metode ini memanfaatkan informasi yang tersedia tentang item, dalam hal ini hotel, untuk memberikan rekomendasi yang disesuaikan dengan preferensi pengguna (Revansa, 2024). Dengan *Content Based Filtering* menggunakan *K-Nearest Neighbor* dan Haversine Formula, sistem ini dapat menganalisis atribut-atribut penting seperti lokasi, harga, dan fasilitas hotel untuk menghasilkan rekomendasi yang lebih akurat (Obajha et al., 2023).

Metode ini bekerja dengan menganalisis interaksi pengguna sebelumnya, seperti rating atau ulasan, untuk memberikan rekomendasi yang lebih akurat. Dalam konteks ini, penelitian ini bertujuan untuk merancang dan mengimplementasikan aplikasi rekomendasi hotel di Kota Balikpapan yang memanfaatkan metode *Content Based Filtering* dengan *K-Nearest Neighbor* dan Haversine Formula (Asshiddiq & Wulandhari, 2024).

Penelitian ini menawarkan pendekatan baru dengan memanfaatkan *content-based filtering* yang dikombinasikan dengan algoritma *K-Nearest Neighbor* dan Haversine Formula, yang belum banyak diterapkan dalam konteks sistem rekomendasi hotel di Indonesia. Pendekatan ini diharapkan dapat meningkatkan akurasi rekomendasi (Maula & Hidayat, 2024).

Penelitian ini bertujuan untuk mengembangkan model sistem rekomendasi hotel yang tidak hanya memberikan saran berdasarkan data historis, tetapi juga mampu mengatasi tantangan yang dikenal sebagai cold start problem (Muliawan et al., 2022). Dengan memanfaatkan metode *content-based filtering* dan teknik pengukuran kemiripan, sistem ini diharapkan dapat memberikan rekomendasi yang akurat dan relevan, bahkan untuk pengguna baru atau hotel yang baru ditambahkan ke dalam database (Muliawan et al., 2022). Melalui pendekatan ini, diharapkan pengguna dapat menemukan akomodasi yang sesuai dengan lebih mudah dan cepat, meningkatkan pengalaman mereka dalam merencanakan perjalanan (Zahra, 2024).

Hasil dari penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam pengembangan sistem rekomendasi yang lebih baik, sehingga pengguna dapat menemukan akomodasi yang sesuai dengan lebih mudah dan cepat, meningkatkan pengalaman mereka dalam merencanakan perjalanan (Melyani, 2022).

Implikasi dari penelitian ini mencakup peningkatan kepuasan pengguna dan loyalitas terhadap platform penyedia akomodasi, serta kontribusi terhadap pengembangan teknologi informasi di sektor perhotelan, yang pada akhirnya dapat mendorong pertumbuhan industri pariwisata secara keseluruhan (Ciaputra & Hansun, 2020).

METODE PENELITIAN

Content-Based Filtering adalah metode dalam sistem rekomendasi yang berfokus pada konten atau fitur dari item yang dibandingkan dengan item yang sebelumnya disukai oleh pengguna. Metode ini tidak melibatkan pengguna lain dalam proses rekomendasi, sehingga rekomendasi yang diberikan bersifat personal dan sesuai dengan preferensi individu pengguna.

Menurut Zuhri (2024) metode dalam sistem rekomendasi yang berfokus pada analisis dan pemanfaatan informasi terkait atribut atau fitur dari item yang direkomendasikan, seperti hotel. Metode ini merekomendasikan item kepada pengguna berdasarkan kesamaan antara item yang telah mereka sukai atau interaksi sebelumnya dengan item lain yang memiliki karakteristik serupa. Dengan demikian, *content-based filtering* berusaha untuk memberikan rekomendasi yang relevan dengan preferensi individu pengguna, tanpa bergantung pada interaksi pengguna lain

Tahap Penelitian



Traveloka mengumpulkan data dari interaksi pengguna, seperti pencarian hotel, pemesanan, dan ulasan yang diberikan. Data ini mencakup informasi tentang lokasi, harga, jenis akomodasi, dan fasilitas yang dicari oleh pengguna.

Pada tahap ini, peneliti mengumpulkan data hotel dari situs traveloka.com, dengan fokus pada hotel bintang 3 hingga 5 yang berlokasi di Balikpapan. Sebanyak 222 data hotel berhasil dikumpulkan untuk dianalisis lebih lanjut.

Setiap hotel atau akomodasi di Traveloka memiliki atribut yang berbeda, seperti:

- Lokasi : Kota atau daerah tempat hotel berada.
- Fasilitas : Kolam renang, Wi-Fi, sarapan, dan lainnya.
- Harga : Rentang harga per malam.
- Tipe Kamar : Kamar standar, suite, dll.
- Rating : Penilaian dari pengguna lain

Setelah data terkumpul, langkah selanjutnya adalah melakukan pre-processing untuk membersihkan dan memformat data agar siap untuk diproses. Proses ini penting untuk mendapatkan informasi yang relevan dan dapat digunakan dalam perhitungan kemiripan (Putri & Prasetyo, 2019).

Sistem dirancang untuk menyimpan respon pengguna terhadap rekomendasi yang ditampilkan. Ketika pengguna memilih hotel, data tersebut secara otomatis disimpan ke dalam database dan digunakan sebagai data training untuk meningkatkan akurasi sistem rekomendasi di masa mendatang (Cahyani et al., 2019).

Dalam tahap ini, peneliti menggunakan metode *Haversine Formula* dan *K-Nearest Neighbor* (K-NN) untuk menghitung kemiripan antara hotel berdasarkan kriteria yang dipilih oleh pengguna. Hasil perhitungan ini akan digunakan untuk menghasilkan rekomendasi hotel yang paling sesuai dengan preferensi pengguna.

Setelah sistem rekomendasi dibangun, dilakukan evaluasi untuk menguji efektivitas dan akurasi sistem dalam memberikan rekomendasi. Uji coba ini melibatkan pengguna untuk mendapatkan umpan balik yang berguna dalam perbaikan system (Aprianto, 2022).

Pada tahap koleksi data, peneliti mengumpulkan informasi mengenai hotel-hotel yang ada di kota Balikpapan. Data yang digunakan dalam penelitian ini diambil dari situs traveloka.com, yang merupakan salah satu platform penyedia informasi akomodasi terkemuka. Peneliti berhasil mengumpulkan sebanyak 218 data hotel yang mencakup berbagai kriteria penting, seperti fasilitas yang ditawarkan, harga, tipe kamar, bintang, skor, dan lokasi. Proses pengumpulan data ini dilakukan dengan cermat untuk memastikan bahwa informasi yang diperoleh akurat dan relevan. Setelah data terkumpul, langkah selanjutnya adalah melakukan pre-processing untuk mempersiapkan data agar siap untuk dianalisis lebih lanjut.

Setelah data dikumpulkan dan diproses, langkah berikutnya adalah melakukan perhitungan untuk menentukan rekomendasi hotel yang paling sesuai dengan preferensi pengguna. Dalam penelitian ini, peneliti menggunakan dua metode utama: *Haversine Formula* dan *K-Nearest Neighbor* (K-NN).

Haversine Formula digunakan untuk menghitung jarak antara lokasi hotel dan lokasi pengguna, sehingga sistem dapat memberikan rekomendasi berdasarkan kedekatan geografis. Sementara itu, K-NN digunakan untuk menghitung kemiripan antara hotel berdasarkan kriteria yang telah ditentukan, seperti fasilitas, harga, tipe kamar, dan skor. Rumus yang digunakan menghitung kedekatan terdekat lokasi pengguna dan lokasi hotel menggunakan metode haversine formula :

$$\begin{aligned}\Delta lat &= lat2 - lat1 \\ \Delta long &= long2 - long \\ a &= \sin^2\left(\frac{\Delta lat}{2}\right) + \cos(lat1) \cdot \cos(lat2) \\ &\quad \times \sin^2\left(\frac{\Delta long}{2}\right) \\ c &= 2 \operatorname{atan2}(\sqrt{a}, \sqrt{1-a}) \\ d &= R \times c\end{aligned}$$

Keterangan :

Lat1 : Latitude asal

Lat2 : latitude tujuan

Long1 : longitude asal

Long2 : longitude tujuan

R : jari-jari bumi 6371 km

Proses perhitungan ini melibatkan pengolahan data numerik yang telah dikonversi sebelumnya, di mana setiap kriteria diberikan nilai skala tertentu. Hasil dari perhitungan ini akan menghasilkan ranking hotel yang kemudian ditampilkan kepada pengguna sebagai rekomendasi. Dengan demikian, pengguna dapat dengan mudah menemukan hotel yang paling sesuai dengan kebutuhan dan preferensi mereka. Metode KNN yang digunakan dalam mencari nilai kemiripan untuk rekomendasi hotel :

$$S(P, C) = \frac{\sum_{i=1}^n S_i \times W_i}{\sum_{i=1}^n W_i}$$

HASIL DAN PEMBAHASAN

Penelitian ini berhasil mengembangkan sistem rekomendasi hotel dengan menerapkan metode *content-based filtering*. Setelah melalui serangkaian tahapan, termasuk pengumpulan data, pre-processing, dan pemodelan, kami melakukan evaluasi terhadap model yang dihasilkan (Sengkey et al., 2022).

Hasil evaluasi menunjukkan bahwa metode *content-based filtering* mampu memberikan akurasi prediksi yang baik dalam merekomendasikan hotel berdasarkan kesamaan fitur antara hotel yang ada dan preferensi pengguna. Kami menggunakan fitur-fitur seperti lokasi, fasilitas, jenis akomodasi, dan rating untuk membangun profil pengguna dan hotel. Analisis lebih lanjut menunjukkan bahwa metode *content-based filtering* efektif dalam menangkap kesamaan antara hotel berdasarkan atribut yang relevan, sehingga meningkatkan akurasi rekomendasi

Hasil evaluasi menunjukkan bahwa metode *content-based filtering* mampu memberikan akurasi prediksi yang baik dalam merekomendasikan hotel berdasarkan kesamaan fitur antara hotel yang ada dan preferensi pengguna. Kami menggunakan fitur-fitur seperti lokasi, fasilitas, jenis akomodasi, dan rating untuk membangun profil pengguna dan hotel hasil pengumpulan data dalam penelitian ini mencakup informasi berikut:

1. Jumlah Data

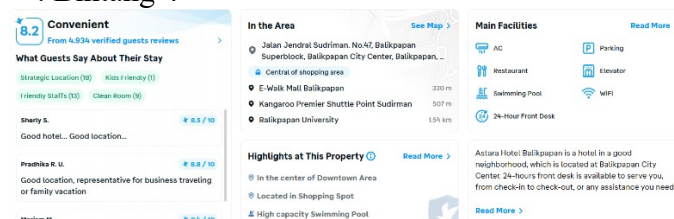
Sistem menghitung kesamaan antara pengguna berdasarkan dua fitur utama: kategori hotel dan kota. Dengan menggunakan data dari hotel yang telah dinilai oleh pengguna, sistem dapat menentukan preferensi pengguna lain yang memiliki kesamaan. Penelitian ini mengumpulkan total 222 data hotel yang terdiri dari hotel bintang 3 hingga 5. Data ini diambil dari situs traveloka.com.



2. Kriteria Data

Data yang dikumpulkan dari hotel Astara Balikpapan mencakup enam kriteria utama, yaitu:

Fasilitas hotel : Kolam renang, Fitness center, Restoran, Akses Shopping Mall
 Tarif Hotel : Mulai dari Rp 1.123.567 hingga Rp 3.058.374 per malam
 Tipe Kamar : Superior Room, Family Room, Junior suite
 Bintang : Bintang 4
 Review : Skor rata-rata 8.3/10
 Lokasi Hotel : Bintang 4



3. Proses Pre-processing

Setelah pengumpulan data, langkah selanjutnya adalah melakukan proses pre-processing. Proses ini melibatkan pembersihan dan pengolahan data untuk memastikan bahwa informasi yang diperoleh dapat diproses dengan baik dalam algoritma yang akan digunakan. Pre-processing ini sangat penting untuk mendapatkan data yang akurat dan relevan, sehingga dapat meningkatkan kualitas rekomendasi yang dihasilkan oleh sistem.

Dengan pengumpulan data yang komprehensif dan proses pre-processing yang tepat, penelitian ini bertujuan untuk membangun sistem rekomendasi hotel yang efektif dan efisien, yang dapat membantu pengguna dalam membuat keputusan yang lebih baik ketika memilih hotel yang sesuai dengan preferensi mereka. Sistem ini diharapkan dapat memberikan alternatif pilihan yang lebih baik di tengah banyaknya informasi yang tersedia di pasar.

KESIMPULAN

Penelitian ini berhasil mengembangkan sistem rekomendasi hotel menggunakan metode *Content-Based Filtering* dengan algoritma *K-Nearest Neighbor* (K-NN) dan Haversine Formula, berfokus pada hotel yang terdaftar di Traveloka di Kota Balikpapan. Hasil menunjukkan bahwa penerapan metode ini efektif dalam memberikan rekomendasi yang relevan berdasarkan preferensi pengguna. Dari analisis fitur-fitur seperti lokasi, harga, tipe kamar, dan fasilitas hotel, sistem dapat menyajikan rekomendasi yang disesuaikan dengan kebutuhan individu. Dengan menggunakan data dari 222 hotel bintang 3 hingga 5, penelitian ini menunjukkan bahwa sistem rekomendasi dapat meningkatkan pengalaman pengguna dalam memilih akomodasi yang sesuai. Hasil evaluasi mengindikasikan bahwa sistem ini tidak hanya dapat memberikan rekomendasi yang akurat, tetapi juga mampu mengatasi cold start problem, sehingga berguna bagi pengguna baru. Dengan demikian, penelitian ini diharapkan dapat berkontribusi pada pengembangan sistem rekomendasi yang lebih baik dan relevan di industri perhotelan, serta meningkatkan kenyamanan dan kepuasan pengguna dalam merencanakan perjalanan mereka.

DAFTAR PUSTAKA

- Aprianto, A. (2022). *TA: Penerapan Algoritma Content-Based Filtering untuk Rekomendasi Destinasi Wisata pada Aplikasi Picnicker*. Universitas Dinamika.
- Asshiddiq, A., & Wulandhari, L. A. (2024). Sistem Rekomendasi Hotel Dengan Ekstraksi Fitur Deskripsi Menggunakan Metode Text Mining dan Content Based Filtering. *Kesatria: Jurnal Penerapan Sistem Informasi (Komputer Dan Manajemen)*, 5(4), 1636–1646.
- Cahyani, A., Gunadi, I. M. A., & Mbulu, Y. P. (2019). Pengaruh Customer Experience Terhadap Repurchase Intention Pada Pt. Traveloka Indonesia. *Jurnal Sains Terapan Pariwisata*, 4(1), 25–36.
- Ciaputra, A. T., & Hansun, D. S. (2020). Rekomendasi Pemilihan Film Dengan Hybrid Filtering Dan K-Nearest Neighbor. *Jurnal Rekayasa Informasi*, 9(2), vol 9-2.
- Fanny, F., Susanto, R., & Nastiti, F. E. (2024). Sistem Rekomendasi Pemilihan Menu Brownies Cinta dengan Metode Content Based Filtering. *G-Tech: Jurnal Teknologi Terapan*, 8(4), 2401–2410.
- Iman, K. T. N., & Wibisono, S. (2021). Pembobotan Menggunakan Pairwise Comparison Pada Case Based Reasoning Rekomendasi Hotel. *Jurnal Manajemen Informatika Dan Sistem Informasi*, 4(1), 9–18.
- Maula, I., & Hidayat, M. N. (2024). Faktorisasi Matriks Menggunakan Stochastic Gradient Descent Untuk Optimasi Sistem Rekomendasi Hotel. *NJCA (Nusantara Journal of Computers and Its Applications)*, 9(1), 19–25.
- Melyani, C. A. (2022). Hotel recommendation system with content-based filtering approach (case study: Hotel in yogyakarta on nusatrip website). *J Statistika: Jurnal Ilmiah Teori Dan Aplikasi Statistika*, 15(1).
- Muliawan, A., Badriyah, T., & Syarif, I. (2022). Membangun Sistem Rekomendasi Hotel dengan Content Based Filtering Menggunakan K-Nearest Neighbor dan Haversine Formula. *Technomedia Journal*, 7(2 October), 231–247.
- Obajha, K., Sari, N. N. K., & Pranatawijaya, V. H. (2023). Implementasi Metode Collaborative Filtering pada Aplikasi Rekomendasi Hotel dan Wisma di Kota Palangka Raya Berbasis Website. *KONSTELASI: Konvergensi Teknologi Dan Sistem Informasi*, 3(2), 398–410.

- Putri, A. K., & Prasetyo, A. (2019). Pengaruh Citra Merek Traveloka Terhadap Loyalitas Konsumen. *EProceedings of Management*, 6(2).
- Revansa, N. H. (2024). *Sistem Rekomendasi Restoran Halal Berbasis Collaborative Filtering dan Autoencoder*. Universitas Islam Indonesia.
- Sengkey, G. T., Tampi, J. R. E., & Walangitan, O. F. C. (2022). Pengaruh Penggunaan E-Commerce Terhadap Keputusan Pembeli Jasa Traveloka. com di Kota Manado. *Productivity*, 3(2), 168–173.
- Wahyudi, K., Latupapua, J., Chandra, R., & Girsang, A. S. (2020). Hotel content-based recommendation system. *Journal of Physics: Conference Series*, 1485(1), 12017.
- Zahra, M. (2024). Pemanfaatan Sistem Rekomendasi Menggunakan Content-Based Filtering pada Hotel di Palangka Raya. *Technologia: Jurnal Ilmiah*, 15(4), 754–763.
- Zuhri, A. A. (2024). *Sistem rekomendasi komunitas akademik Teknik Informatika UIN Malang menggunakan metode Content-Based Filtering*. Universitas Islam Negeri Maulana Malik Ibrahim.